(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 June 2022 (16.06.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/125968 A1

(51) International Patent Classification:
C12N 9/22 (2006.01)
C12N 9/78 (2006.01)
C12N 15/10 (2006.01)
C12N 15/62 (2006.01)
C12N 15/113 (2010.01)
A61K 9/127 (2006.01)
A61K 31/7088 (2006.01)

(21) International Application Number:
PCT/US2021/062922

(22) International Filing Date:
10 December 2021 (10.12.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

63/124,060	11 December 2020 (11.12.2020)	US
63/130,104	23 December 2020 (23.12.2020)	US
63/165,636	24 March 2021 (24.03.2021)	US
63/275,424	03 November 2021 (03.11.2021)	US

(71) Applicant: **INTELLIA THERAPEUTICS, INC.** [US/US]; 40 Erie Street, Cambridge, Massachusetts 02139 (US).

(72) Inventors: **DOMBROWSKI, Christian**; 40 Erie Street, Cambridge, Massachusetts 02139 (US). **HARRINGTON, William, Frederick**; 40 Erie Street, Cambridge, Massachusetts 02139 (US). **OLIVEIRA, Ruan**; 40 Erie Street, Cambridge, Massachusetts 02139 (US).

(74) Agent: **BAUR, Amelia, Feulner** et al.; McNeill Baur PLLC, 125 Cambridge Park Drive, Suite 301, Cambridge, Massachusetts 02140 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: POLYNUCLEOTIDES, COMPOSITIONS, AND METHODS FOR GENOME EDITING INVOLVING DEAMINATION

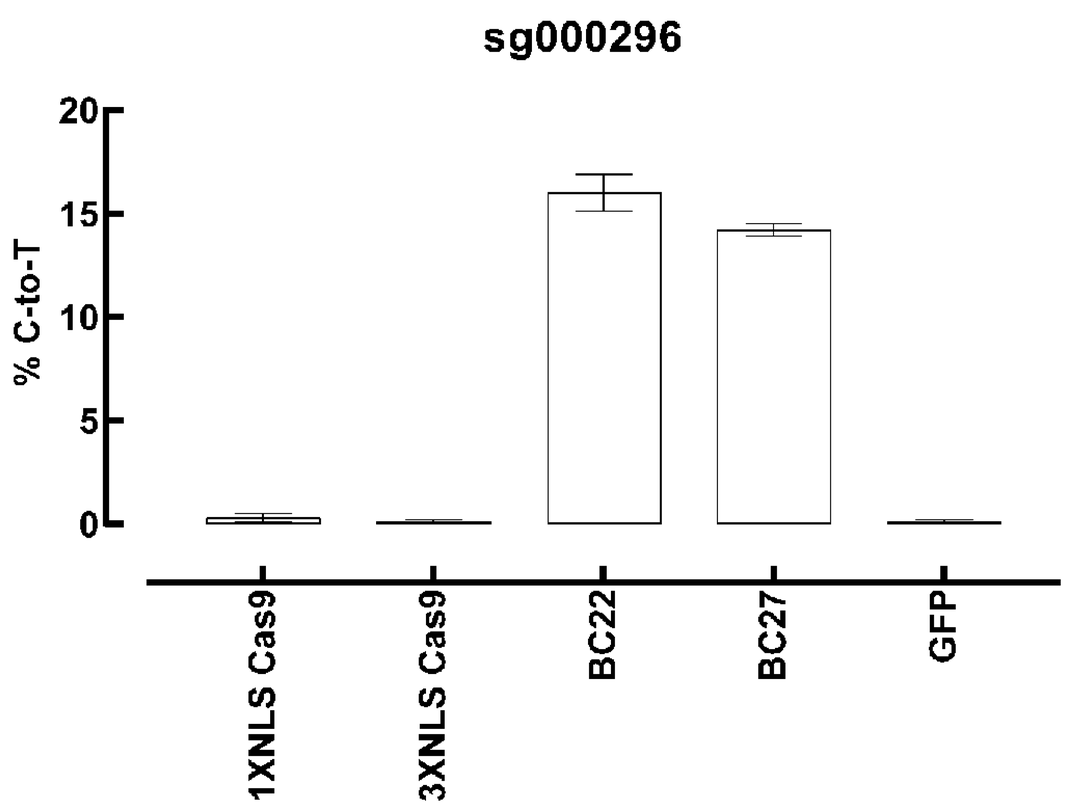


FIG. 1A

(57) Abstract: Polynucleotides, polypeptides, compositions, and methods for genome editing using deamination are provided. An mRNA containing an open reading frame (ORF) encoding a polypeptide is provided herein. The polypeptide includes a cytidine deaminase and an RNA-guided nickase, and does not include a uracil glycosylase inhibitor (UGI). A composition provided herein may include two different mRNAs. The first mRNA includes an ORF encoding a cytidine deaminase and an RNA-guided nickase, and the second mRNA includes an ORF encoding uracil glycosylase inhibitor (UGI).

[Continued on next page]

WO 2022/125968 A1

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

POLYNUCLEOTIDES, COMPOSITIONS, AND METHODS FOR GENOME EDITING INVOLVING DEAMINATION

[0001] This application claims the benefit under 35 U.S.C. 119(e) of U.S. Provisional Application No. 63/124,060, filed December 11, 2020; U.S. Provisional Application No. 63/130,104, filed December 23, 2020; U.S. Provisional Application No. 63/165,636, filed March 24, 2021; and U.S. Provisional Application No. 63/275,424, filed November 3, 2021, each of which is herein incorporated by reference in its entirety.

[0002] This application is filed with a Sequence Listing in electronic format. The Sequence Listing is provided as a file entitled "2021-12-08_01155-0016-00PCT_ST25.txt" created on December 8, 2021, which is 1,557,107 bytes in size. The information in the electronic format of the sequence listing is incorporated herein by reference in its entirety.

INTRODUCTION AND SUMMARY

[0003] The present disclosure relates to polynucleotides, compositions, and methods for genomic editing involving deamination.

[0004] Base editing is a genome editing method that directly generates point mutations within a specific region of the genomic DNA without causing double-stranded breaks (DSB). DNA base editors (BEs) comprise fusions between a catalytically impaired Cas nuclease and a base-modification enzyme. Currently, effectors for cytidine-to-thymidine (C-to-T) editing fuse a cytidine deaminase with a nickase and a uracil glycosylase inhibitor (UGI). For example, base editor 3 (BE3) consists of a Cas9 nuclease bearing a mutation that converts it into a nickase (nCas9), fused to an APOBEC1 (apolipoprotein mRNA editing enzyme, catalytic polypeptide 1) deaminase and a UGI (e.g., Wang et al. Cell Research 27:1289-1292 (2017)), and it was reported that an "nCas9-fused UGI domain is still important for achieving high fidelity of base editing, even when high levels of free UGI is present." As another example, an engineered human APOBEC3A (A3A or apolipoprotein mRNA editing enzyme, catalytic polypeptide 3A) deaminase has been investigated as a replacement for rat APOBEC1 deaminase (RAPO1) in the original BE3 (Gehrke et al., Nature Biotechnology, 36: 977-982 (2018)), but it was noted that the ability of base editors "to edit all Cs within their editing window can potentially have deleterious effects" and that "mutation of the N57 residue in the human A3A deaminase was critical to restoring its native target sequence precision in the context of a base editor and also to lowering its off-target editing activity." Indeed, APOBEC3A-Class 2 Cas nickase (D10A) base editors have been reported as having a "high degree of mutagenicity" and as showing Cas9-independent off-

target base editing (Doman et al., Nature Biotechnology 38:620-628 (2020)). Accordingly, improved compositions and methods for targeted C-to-T base editing using cytidine deaminases (e.g., an APOBEC3A deaminase) and RNA-guided nickase are needed.

[0005] Accordingly, the present disclosure provides polynucleotides, compositions, and methods for genomic editing involving a cytidine deaminase (e.g., an APOBEC3A deaminase) and an RNA-guided nickase that induce C-to-T conversions at target nucleotides with greater fidelity and may minimize bystander mutations. The present disclosure is based in part on the findings that by pairing a cytidine deaminase (e.g., an APOBEC deaminase) and an RNA-guided nickase system with UGI *in trans* (e.g., as a separate mRNA), it is possible to lower the amount of other base editing (C-to-A/G conversions, insertions, or deletions) and increase the purity of C-to-T editing.

[0006] Accordingly, the following embodiments are provided.

[0007] In some embodiments, a composition is provided, the composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA. In some embodiments, the first open reading frame does not comprise a sequence encoding a UGI. In some embodiments, the composition comprises a first composition and a second composition, wherein the first composition comprises a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase and does not comprise a uracil glycosylase inhibitor (UGI), and the second composition comprises a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA. In some embodiments, the composition comprises lipid nanoparticles.

[0008] In some embodiments, a method of modifying a target gene is provided, the method comprising delivering to a cell a first mRNA comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase and an RNA-guided nickase, a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, and at least one guide RNA (gRNA).

[0009] In some embodiments, a method of modifying at least one cytidine within a target gene in a cell is provided, the method comprising expressing in the cell or contacting the cell with: (i) a first polypeptide comprising a cytidine deaminase and an RNA-guided

nickase, wherein the first polypeptide does not comprise a uracil glycosylase inhibitor (UGI); (ii) a UGI polypeptide; and (iii) at least one guide RNA (gRNA) wherein the first polypeptide and gRNA form a complex with the target gene and modify the at least one cytidine in the target gene. In some embodiments, the ratio of the UGI polypeptide to the first polypeptide is from 10:1 to 50:1.

[0010] In some embodiments, an mRNA containing an open reading frame (ORF) encoding a polypeptide is provided, the polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI). The polypeptide encoded by the mRNA is also provided. In some embodiments, a method of modifying a target gene is provided, the method comprising delivering an mRNA or a polypeptide described herein to a cell.

[0011] In some embodiments, a composition comprises two different mRNAs in which the first mRNA comprises an ORF encoding a cytidine deaminase (e.g., A3A) and an RNA-guided nickase, and the second mRNA comprises an ORF encoding uracil glycosylase inhibitor (UGI). In some embodiments, the first mRNA in the composition does not comprise an ORF encoding UGI. In some embodiments, the molar ratio of the second mRNA to the first mRNA is from 1:1 to 30:1, from 2:1 to 30:1, from 7:1 to 22:1. In some embodiments, the molar ratio of the second mRNA to the first mRNA is 22:1, 7:1, 2:1, or 1:1, 1:4, 1:11, or 1:33.

[0012] Further embodiments are provided throughout and described in the claims and Figures.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] Figs. 1A-1E show C-to-T conversion activity deaminase editors profiled against 5 different guide targeting sequences, respectively.

[0014] Fig. 2A shows the percentage of edited reads with the 4 target cytosines converted to thymidines for deaminase editors profiled using sg000296.

[0015] Fig. 2B shows the percentage of edited reads with the 5 target cytosines converted to thymidines for deaminase editors profiled using sg0001373.

[0016] Fig. 2C shows the percentage of edited reads with the 4 target cytosines converted to thymidines for deaminase editors profiled using sg001400.

[0017] Fig. 2D shows the percentage of edited reads with the 6 target cytosines converted to thymidines for deaminase editors profiled using sg003018.

[0018] Fig. 2E shows the percentage of edited reads with at least 6 of 8 target cytosines converted to thymidines for deaminase editors profiled using sg005883.

[0019] Fig. 3A-3E show the percentage of C-to-T conversion by base position for 5 different guide targeting sequences, respectively.

[0020] Figs. 4A-4B show editing profile as a percentage of total reads in U-2OS cells (Fig. 4A) and HuH-7 cells (Fig. 4B).

[0021] Figs. 5A-5B show editing profile as a percentage of edited reads in U-2OS cells (Fig. 5A) and HuH-7 cells (Fig. 5B).

[0022] Figs. 6A-6B represent editing profile as a percentage of total reads upon titrating UGI mRNAs (SEQ ID NOs: 25 and 34).

[0023] Fig. 7 shows TTR editing levels in CD-1 mice treated with different LNP combinations with mRNA constructs and sgRNAs. Fig. 7 shows % editing of C-to-T conversions, C-to-A/G conversions, and indels of DNA sequences (NGS sequencing) extracted from liver tissue samples harvested from the CD-1 mice.

[0024] Fig. 8 shows TTR editing levels in liver tissue harvested from CD-1 mice treated with different LNP combinations with mRNA constructs and sgRNAs when the UGI sequence was delivered *in trans* (a separate mRNA).

[0025] Fig. 9A-9C show scatter plots representing statistically significant (p. adj. < 0.05) differential gene expression events (black dots) in liver samples from mice treated with sgRNA G000282 and BC22n mRNA in the absence of UGI mRNA *in trans* (Fig. 9A), sgRNA G000282 and BC22n mRNA in the presence of UGI mRNA *in trans* (Fig. 9B) or Cas9 mRNA in the presence of UGI mRNA *in trans* (Fig. 9C).

[0026] Fig. 10 shows the editing profile in T cells following treatment with different mRNA constructs and CIITA-targeting sgRNAs.

[0027] Fig. 11 shows MHC class II negative cells assessed by flow cytometry analysis of T cells treated with different mRNA constructs and CIITA guide RNAs.

[0028] Figs. 12A-12B show scatter plots showing statistically significant (* = p. adj. < 0.05) differential gene expression events (black dots) in T cells treated with sgRNA G018076, UGI mRNA and either Cas9 mRNA (Fig. 12A) or BC22n mRNA (Fig. 12B).

[0029] Figs. 13A-13B show scatter plots showing statistically significant (* = p. adj. < 0.05) differential gene expression events (black dots) in T cells treated with sgRNA G018117, UGI mRNA and either Cas9 mRNA (Fig. 13A) or BC22n mRNA (Fig. 13B).

[0030] Figs. 14A-14B show protein-protein interaction networks enriched among the list of differentially expressed genes in T cells treated with sgRNA G018076, UGI mRNA and either Cas9 mRNA (Fig. 14A) or BC22n mRNA (Fig. 14B).

[0031] Figs. 15A-15B show protein-protein interaction networks enriched among the list of differentially expressed genes in T cells treated with sgRNA G018117, UGI mRNA and either Cas9 mRNA (Fig. 15A) or BC22 mRNA (Fig. 15B).

[0032] Figs. 16A-16C show the editing profiles of T cells. Editing profiles in the on-target TRAC locus (Fig. 16A) in addition to 10 loci previously described as mutational hotspots in APOBEC-positive tumors (Figs. 16B-16C).

[0033] Figs. 17A-17E show the editing profiles of T cells when treated with varying levels of BC22n mRNA and Cas9 mRNAs. Cells were edited using sgRNAs G015995 (Fig. 17A), G016017 (Fig. 17B), G016206 (Fig. 17C), G018117 (Fig. 17D), or G016086 (Fig. 17E).

[0034] Figs. 18A-18D show the editing profiles for T cells edited with four guides simultaneously using varying levels of BC22n mRNA or Cas9 mRNAs. The editing profile at each edited locus is represented separately: G015995 (Fig. 18A), G016017 (Fig. 18B), G016206 (Fig. 18C), G018117 (Fig. 18D).

[0035] Figs. 19A-19I show phenotyping results as percent of cells negative for antibody binding with increasing total RNA for BC22n and Cas9 samples. Fig. 19A shows the percentage of B2M negative cells when B2M guide G015995 was used for editing. Fig. 19B shows the percentage of B2M negative cells when multiple guides were used for editing. Fig. 19C shows the percentage of CD3 negative cells when TRAC guide G016017 was used for editing. Fig. 19D shows the percentage of CD3 negative cells when TRBC guide G016206 was used for editing. Fig. 19E shows the percentage of CD3 negative cells when multiple guides were used for editing. Fig. 19F shows the percentage of MHC Class II negative cells when CIITA guide G018117 was used for editing. Fig. 19G shows the percentage of MHC Class II negative cells when multiple guides were used for editing. Fig. 19H shows the percentage of triple (B2M, CD3, MHC II) negative cells when multiple guides were used for editing. FIG. 19I shows the percentage of MHC class II negative T cells when CIITA guide G016086 was used for editing.

[0036] Figs. 20A-20C show the editing profile for T cells while using varying levels of UGI mRNA *in trans* with various editor mRNAs. Fig. 20A shows the percent editing with BC22n mRNA at 27.3 nM. Fig. 20B shows the percent editing with BC22-2xUGI mRNA at 24.7 nM. Fig. 20C shows the percent editing with BE4Max mRNA at 24.0 nM.

[0037] Fig. 21 shows C-to-T conversion as a percentage of edited reads while using varying levels of UGI mRNA *in trans* with various editor mRNAs (BC22n at 27.3 nM, BC22-2xUGI at 24.7 nM, BE4Max at 24.0 nM).

[0038] Fig. 22 shows the mean percentage of T cells negative for cell surface expression of MHC II (% MHC II negative") for several guides using Cas9 or BC22 in relation to the distance between the cut site boundary nucleotide shown as base pairs ("bp"). Positive numerical values indicate a splice site boundary nucleotide 3' of the cut site, whereas the negative numerical values indicate a splice site boundary nucleotide 5' of the cut site.

[0039] Fig. 23A shows an exemplary sgRNA (SEQ ID NO: 141, methylation not shown) in a possible secondary structure with labels designating individual nucleotides of the conserved region of the sgRNA, including the lower stem, bulge, upper stem, nexus (the nucleotides of which can be referred to as N1 through N18, respectively, in the 5' to 3' direction), and the hairpin region which includes hairpin 1 and hairpin 2 regions. A nucleotide between hairpin 1 and hairpin 2 is labeled n. A guide region may be present on an sgRNA and is indicated in this figure as "(N)x" preceding the conserved region of the sgRNA.

[0040] Fig. 23B labels the 10 conserved region YA sites in an exemplary sgRNA sequence (SEQ ID NO: 141, methylation not shown) from 1 to 10. The numbers 25, 45, 50, 56, 64, 67, and 83 indicate the position of the pyrimidine of YA sites 1, 5, 6, 7, 8, 9, and 10 in an sgRNA with a guide region indicated as $(N)_x$, e.g., wherein x is optionally 20.

[0041] Figs. 24A-B show results for efficiency of three CIITA guides (G016086, G016092, and G016067) for editing T cells with BC22. FIG. 24A shows the percent C-to-T conversion. FIG. 24B shows the percentage of MHC class II negative T cells.

[0042] Fig. 25 shows the percentage of B2M negative T cells after editing with different mRNA combinations. EP indicates electroporation.

[0043] Fig. 26 shows the percentage of single nucleotide variants (SNVs) that are C-to-U conversions in the transcriptome of T cells edited at B2M with different mRNA combinations (n.s. = not significant).

[0044] Fig. 27 shows the percentage of B2M negative T cells after-treatment with different mRNA combinations.

[0045] Fig. 28 shows the editing profiles at B2M locus in T cells after-treatment with different mRNA combinations.

[0046] Fig. 29 shows the percentage of SNVs that are C-to-T conversions in amplified genomic DNA from single T cells edited at B2M with different mRNA combinations (n.s. = not significant).

[0047] Fig. 30 shows the percentage of B2M negative eHap1 cells following editing with different mRNA combinations.

[0048] Fig. 31 shows the editing profiles of the B2M locus in eHap1 cells following treatment with different mRNA combinations.

[0049] Fig. 32 shows the percentage of SNVs that are C-to-T conversions in clonally expanded eHap1 cells editing at B2M with different mRNA combinations (n.s. = not statistically significant).

[0050] Fig. 33 shows the cell viability relative to untreated cells following electroporation or LNP delivery of BC22n or Cas9 editors and single or multiple guides.

[0051] Fig. 34 shows the total γH2AX spot intensity per nuclei following electroporation or LNP delivery of BC22n or Cas9 editors and single or multiple guides.

[0052] Fig. 35 shows the percentage editing at loci of interest following LNP delivery of BC22n or Cas9 editors and single or multiple guides.

[0053] Fig. 36 shows the percentage of negative cells for stated surface proteins following LNP delivery of BC22n or Cas9 editors and single or multiple guides.

[0054] Fig. 37 shows the percentage of interchromosomal translocations among total unique molecules following LNP delivery of BC22n or Cas9 editors and multiple guides.

[0055] Fig. 38 shows mean percent editing in mouse liver following treatment with various base editors.

[0056] Fig. 39 shows mean percent C-to-T conversion activity for base editor constructs designed with various deaminase domains.

[0057] Figs. 40A-40K show percent of total reads containing at least 1 C to T conversion for base editing constructs designed with various linkers.

[0058] Fig. 41 shows mean percent editing at SERPINA1 in Huh-7 after treatment with base editor constructs designed using various linkers.

[0059] Fig. 42 shows EC90 for base editing mRNA designed with various linkers. The 95% confidence interval (CI) for each EC90 value is also shown.

[0060] Fig. 43 shows mean percent editing at the ANAPC5 locus in PHH using base editing constructs designed with various linkers.

[0061] Fig. 44 shows EC95 for base editing mRNA designed with various linkers. The 95% confidence interval (CI) for each EC95 value is also shown.

[0062] Fig. 45 shows mean percent editing at the TRAC locus in PHH using base editing constructs designed with various linkers.

[0063] Fig. 46 shows Mass of base editor mRNAs designed with various linkers that leads to 90% of the maximum (EC90) knockdown of CD3. The 95% confidence interval for each EC50 value is also shown.

[0064] Fig. 47 shows the mass of BC22n mRNAs designed with various linkers that leads to 90% of the maximum (EC90) knockdown of CD3, HLA-A3 and HLA-DR, DP, DQ (EC50). The 95% confidence interval (CI) for each EC50 value is also shown.

[0065] Fig. 48 shows mean percent editing at the TRAC locus in T cells treated with sgRNA in the 100-mer or 91-mer formats.

[0066] Fig. 49A shows mean percent editing at the TRBC1 locus in T cells treated with sgRNA in the 100-mer or 91-mer formats.

[0067] Fig. 49B shows mean percent editing at the TRBC2 locus in T cells treated with sgRNA in the 100-mer or 91-mer formats.

[0068] Fig. 50 shows mean percent editing at the CIITA locus in T cells treated with sgRNA in the 100-mer or 91-mer formats.

[0069] Fig. 51 shows mean percent editing at the B2M locus in T cells treated with sgRNA in the 100-mer or 91-mer formats.

[0070] Fig. 52 shows mean percent editing at the CD38 locus in T cells treated with sgRNA in the 100-mer or 91-mer formats.

[0071] Fig. 53A shows the mean percentage of CD8+ T cells that are negative for CD3 surface receptors following treatment with sgRNAs in the 100-mer or 91-mer formats targeting TRAC.

[0072] Fig. 53B shows the mean percentage of CD8+ T cells that are negative for CD3 surface receptors following treatment with sgRNAs in the 100-mer or 91-mer formats targeting TRBC.

[0073] Fig. 54A shows the mean percentage of CD8+ T cells that are negative for HLA-DR, DP, DQ surface receptors following treatment with sgRNAs in the 100-mer or 91-mer formats targeting CIITA.

[0074] Fig. 54B shows the mean percentage of CD8+ T cells that are negative for HLA-A surface receptors following treatment with sgRNAs in the 100-mer or 91-mer formats targeting HLA-A.

[0075] Fig. 55A shows the mean percentage of CD8+ T cells that are negative for B2M surface receptors following treatment with sgRNAs in the 100-mer or 91-mer formats targeting B2M.

[0076] Fig. 55B shows the mean percentage of CD8+ T cells that are negative for CD38 surface receptors following treatment with sgRNAs in the 100-mer or 91-mer formats targeting CD38.

[0077] Fig. 56 shows mean percent editing at the ANAPC5 locus in mouse liver with increasing amounts of UGI mRNA.

[0078] Fig. 57 shows percent C to T editing purity at the ANAPC5 locus in mouse liver following editing with increasing amounts of UGI mRNA.

[0079] Fig. 58A shows total editing at different BC22n-HiBiT mRNA concentrations at B2M in PHH cells.

[0080] Fig. 58B shows C-to-T purity at different UGI-HiBiT mRNA concentrations at B2M in PHH cells.

[0081] Fig. 58C shows total editing at different BC22-2xUGI-HiBiT mRNA concentrations at B2M in PHH cells.

[0082] Fig. 58D shows C-to-T purity at different BC22-2xUGI-HiBiT mRNA concentrations at B2M in PHH cells.

[0083] Fig. 59A shows total editing at different BC22n-HiBiT mRNA concentrations in T cells.

[0084] Fig. 59B shows C-to-T purity at different UGI-HiBiT mRNA concentrations in T cells.

[0085] Fig. 59C shows total editing at different BC22-2xUGI-HiBiT mRNA concentrations in T cells.

[0086] Fig. 59D shows C-to-T purity at different BC22-2xUGI-HiBiT mRNA concentrations in T cells.

[0087] Fig. 60 shows editing in liver tissue harvested from CD-1 mice treated with LNPs with fixed doses of sgRNA and base editor mRNA and different doses of UGI mRNA.

[0088] Fig. 61 shows C-to-T purity in liver tissue harvested from CD-1 mice treated with LNPs with fixed doses of sgRNA and base editor mRNA and different doses of UGI mRNA.

[0089] Fig. 62 shows the percent lysis of T cells targeted by NK cells at different effector:target (E:T) ratios treated with sgRNA and base editor and UGI mRNAs.

[0090] Fig. 63 shows conversion rate at each guide nucleotide position for high activity guides edited with a Spy base editor.

[0091] Fig. 64 shows conversion rate at each guide nucleotide position for high activity guides edited with an Nme2 base editor.

BRIEF DESCRIPTION OF DISCLOSED SEQUENCES

SEQ ID NO	Description
1	mRNA encoding BC22n
2	Open reading frame for BC22n
3	Amino acid sequence for BC22n
4	mRNA encoding BC22n with HiBit tag
5	Open reading frame for BC22n with HiBit tag
6	Amino acid sequence for BC22n with HiBit tag
7	Not used
8	Open reading frame for Cas9
9	Amino acid sequence for Cas9
10	Not used
11	Open reading frame for Cas9
12	Amino acid sequence for Cas9
13	mRNA encoding BE3
14	Open reading frame for BE3
15	Amino acid sequence for BE3
16	mRNA encoding BE3
17	Open reading frame for BE3
18	Amino acid sequence for BE3
19	mRNA encoding BC22
20	Open reading frame for BC22
21	Amino acid sequence for BC22
22	Not used
23	Open reading frame for Cas9 with HiBit tag
24	Amino acid sequence for Cas9 with HiBit tag
25	mRNA encoding UGI
26	Open reading frame for UGI
27	Amino acid sequence for UGI
28	mRNA encoding BC22 with 2x UGI
29	Open reading frame for BC22 with 2x UGI
30	Amino acid sequence for BC22 with 2x UGI
31	mRNA encoding BE4MAX protein
32	Open reading frame for BE4MAX protein
33	Amino acid sequence for BE4MAX protein
34	mRNA sequence encoding UGI
35	Open reading frame for UGI
36	Amino acid sequence for recombinant Cas9

37	Not used
38	Not used
39	Not used
40	Amino acid sequence of *H. sapiens* APOBEC3A deaminase (A3A)
41	Amino acid sequence of *R. norvegicus* APOBEC1
42	Exemplary coding sequence for UGI (SEQ ID NO. 43)
43	Amino acid sequence for exemplary UGI
44	Exemplary coding sequence for XTEN (SEQ ID NO. 46)
45	Exemplary coding sequence for XTEN (SEQ ID NO. 46)
46	Amino acid sequence for exemplary XTEN
47	Amino acid sequence for exemplary XTEN
48	Amino acid sequence for exemplary XTEN
49	Amino acid sequence for exemplary linker
50	Amino acid sequence for exemplary linker
51	Amino acid sequence for exemplary linker
52	Amino acid sequence for exemplary linker
53	Amino acid sequence for exemplary linker
54	Amino acid sequence for exemplary linker
55	Amino acid sequence for exemplary linker
56	Amino acid sequence for exemplary linker
57	Amino acid sequence for exemplary linker
58	Amino acid sequence for exemplary linker
59	Amino acid sequence for exemplary linker
60	Nucleic acid sequence for exemplary linker
61	Amino acid sequence for exemplary linker
62	Nucleic acid sequence for SV40 NLS
63	Amino acid sequence for SV40 NLS
64	pCl-Neo
65	Screening plasmid - invariant sequence
66	pUC19
67	U6 promoter
68	CMV promoter
69	3' UTR from human albumin gene
70	Amino acid sequence of Spy Cas9 nickase (D10A) with 1x NLS as the C-terminal 7 amino acids
71	Spy Cas9 nickase (D10A) ORF encoding SEQ ID NO: 70 using minimal uridine codons as listed in Table 3, with start and stop codons
72	Spy Cas9 nickase (D10A) ORF coding sequence using minimal uridine codons as listed in Table 3 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)
73	Amino acid sequence of Spy Cas9 nickase (without NLS)
74	Spy Cas9 nickase ORF encoding SEQ ID NO: 73 using minimal uridine codons as listed in Table 3, with start and stop codons
75	Spy Cas9 nickase coding sequence encoding SEQ ID NO: 73 using minimal

	uridine codons as listed in Table 3 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)
76	Amino acid sequence of Spy Cas9 nickase with two nuclear localization signals as the C-terminal amino acids
77	Spy Cas9 nickase ORF encoding SEQ ID NO: 76 using minimal uridine codons as listed in Table 3, with start and stop codons
78	Spy Cas9 nickase coding sequence encoding SEQ ID NO: 76 using minimal uridine codons as listed in Table 3 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)
79	Spy Cas9 nickase ORF using low A codons of Table 4, with start and stop codons
80	Spy Cas9 nickase ORF using low A codons of Table 4, with start and stop codons and no NLS
81	Spy Cas9 nickase ORF using low A codons of Table 4, with two C-terminal NLS sequences and start and stop codons
82	Spy Cas9 nickase ORF using low A/U codons of Table 4, with start and stop codons
83	Spy Cas9 nickase ORF using low A/U codons of Table 4, with two C-terminal NLS sequences and start and stop codons
84	Spy Cas9 nickase ORF using low A/U codons of Table 4, with start and stop codons and no NLS
85	Spy Cas9 nickase ORF using low A codons of Table 4 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)
86	Spy Cas9 nickase ORF using low A codons of Table 4 (no NLS and no start or stop codons; suitable for inclusion in fusion protein coding sequence)
87	Spy Cas9 nickase ORF using low A codons of Table 4, with two C-terminal NLS sequences (no start or stop codons; suitable for inclusion in fusion protein coding sequence)
88	Spy Cas9 nickase ORF using low A/U codons of Table 4 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)
89	Spy Cas9 nickase ORF using low A/U codons of Table 4, with two C-terminal NLS sequences (no start or stop codons; suitable for inclusion in fusion protein coding sequence)
90	Spy Cas9 nickase ORF using low A/U codons of Table 4 (no NLS and no start or stop codons; suitable for inclusion in fusion protein coding sequence)
91	Exemplary 5’ UTR
92	Exemplary 5’ UTR
93	Exemplary 5’ UTR
94	Exemplary 5’ UTR
95	Exemplary 5’ UTR
96	Exemplary 5’ UTR
97	Exemplary 5’ UTR
98	Exemplary 5’ UTR
99	Exemplary 3’ UTR
100	Exemplary 3’ UTR
101	Exemplary 3’ UTR
102	Exemplary 3’ UTR

103	Exemplary 3’ UTR
104	Exemplary 3’ UTR
105	Exemplary 3’ UTR
106	Exemplary 3’ UTR
107	Exemplary Kozak sequence
108	Exemplary Kozak sequence
109	Exemplary poly-A sequence
110	Exemplary NLS 1
111	Exemplary NLS 2
112	Exemplary NLS 3
113	Exemplary NLS 4
114	Exemplary NLS 5
115	Exemplary NLS 6
116	Exemplary NLS 7
117	Exemplary NLS 8
118	Exemplary NLS 9
119	Exemplary NLS 10
120	Exemplary NLS 11
121	Alternative SV40 NLS
122	Nucleoplasmin NLS
123	Exemplary coding sequence for SV40 NLS
124	Exemplary coding sequence for NLS1
125	Exemplary coding sequence for NLS2
126	Exemplary coding sequence for NLS3
127	Exemplary coding sequence for NLS4
128	Exemplary coding sequence for NLS5
129	Exemplary coding sequence for NLS6
130	Exemplary coding sequence for NLS7
131	Exemplary coding sequence for NLS8
132	Exemplary coding sequence for NLS9
133	Exemplary coding sequence for NLS10
134	Exemplary coding sequence for NLS11
135	Exemplary coding sequence for alternate SV40 NLS
136-138	Not used
139	Exemplary nucleotide sequence following the 3’ end of the guide sequence to form a crRNA
140	Conserved Portion of a spyCas9 sgRNA
141	Modified sgRNA pattern, where N are nucleotides encoding a guide sequence
142	Exemplary guide constant region modification pattern (G282-C)
143	Exemplary guide modification pattern (G282-mN3Nx)
144	Exemplary guide modification pattern (G282-Nx)
145	Exemplary guide modification pattern (G282-N20)
146-150	Not used

151	Exemplary guide sequence
152	Exemplary guide sequence for B2M gene
153	Exemplary guide sequence for TTR gene
154	Exemplary guide sequence for TRAC gene
155	Exemplary guide sequence for TRBC1/2 gene
156	Exemplary guide sequence
157	Exemplary guide sequence for SERPINA1 gene
158	Exemplary guide sequence for SERPINA1 gene
159	Exemplary guide sequence
160	Exemplary guide sequence for CIITA gene
161	Exemplary guide sequence for CIITA gene
162	Exemplary guide sequence for CIITA gene
163	Exemplary guide sequence for CIITA gene
164	Exemplary guide sequence for CIITA gene
165	Exemplary guide sequence for CIITA gene
166	Exemplary guide sequence for CIITA gene
167	Exemplary guide sequence for CIITA gene
168	Exemplary guide sequence for CIITA gene
169	Exemplary guide sequence for CIITA gene
170	Exemplary guide sequence for CIITA gene
171	Exemplary guide sequence for CIITA gene
172	Exemplary guide sequence for CIITA gene
173	Exemplary guide sequence for CIITA gene
174-176	not used
177	G013009 guide RNA targeting TRAC
178	G016016 guide RNA targeting TRAC
179	G015991 guide RNA targeting B2M
180	G015996 guide RNA targeting B2M
181	G000297 guide RNA
182	G015995 guide RNA targeting B2M with guide sequence SEQ ID NO: 152
183	G000282 guide RNA targeting TTR with guide sequence SEQ ID NO: 153
184	G016017 guide RNA targeting TRAC with guide sequence SEQ ID NO: 154
185	G016206 guide RNA targeting TRBC1/2 with guide sequence SEQ ID NO: 155
186	SG000296 guide RNA
187	SG001373 guide RNA targeting SERPINA1 with guide sequence SEQ ID NO: 157
188	SG001400 guide RNA targeting SERPINA1 with guide sequence SEQ ID NO: 158
189	SG005883 guide RNA
190	SG003018 guide RNA targeting CIITA with guide sequence SEQ ID NO: 160
191	G018075 guide RNA targeting CIITA with guide sequence SEQ ID NO: 161
192	G018076 guide RNA targeting CIITA with guide sequence SEQ ID NO: 162

193	G018077 guide RNA targeting CIITA with guide sequence SEQ ID NO: 163
194	G018078 guide RNA targeting CIITA with guide sequence SEQ ID NO: 164
195	G018081 guide RNA targeting CIITA with guide sequence SEQ ID NO: 165
196	G018082 guide RNA targeting CIITA with guide sequence SEQ ID NO: 166
197	G018084 guide RNA targeting CIITA with guide sequence SEQ ID NO: 167
198	G018085 guide RNA targeting CIITA with guide sequence SEQ ID NO: 168
199	G018091 guide RNA targeting CIITA with guide sequence SEQ ID NO: 169
200	G018100 guide RNA targeting CIITA with guide sequence SEQ ID NO: 170
201	G018117 guide RNA targeting CIITA with guide sequence SEQ ID NO: 171
202	G018118 guide RNA targeting CIITA with guide sequence SEQ ID NO: 172
203	G018120 guide RNA targeting CIITA with guide sequence SEQ ID NO: 173
204-210	Not used
211	Amino acid sequence for exemplary linker
212	Amino acid sequence for exemplary linker
213	Amino acid sequence for exemplary linker
214	Amino acid sequence for exemplary linker
215	Amino acid sequence for exemplary linker
216	Amino acid sequence for exemplary linker
217	Amino acid sequence for exemplary linker
218	Amino acid sequence for exemplary linker
219	Amino acid sequence for exemplary linker
220	Amino acid sequence for exemplary linker
221	Amino acid sequence for exemplary linker
222	Amino acid sequence for exemplary linker
223	Amino acid sequence for exemplary linker
224	Amino acid sequence for exemplary linker
225	Amino acid sequence for exemplary linker
226	Amino acid sequence for exemplary linker
227	Amino acid sequence for exemplary linker
228	Amino acid sequence for exemplary linker
229	Amino acid sequence for exemplary linker
230	Amino acid sequence for exemplary linker
231	Amino acid sequence for exemplary linker
232	Amino acid sequence for exemplary linker
233	Amino acid sequence for exemplary linker
234	Amino acid sequence for exemplary linker
235	Amino acid sequence for exemplary linker
236	Amino acid sequence for exemplary linker
237	Amino acid sequence for exemplary linker
238	Amino acid sequence for exemplary linker
239	Amino acid sequence for exemplary linker
240	Amino acid sequence for exemplary linker

241	Amino acid sequence for exemplary linker
242	Amino acid sequence for exemplary linker
243	Amino acid sequence for exemplary linker
244	Amino acid sequence for exemplary linker
245	Amino acid sequence for exemplary linker
246	Amino acid sequence for exemplary linker
247	Amino acid sequence for exemplary linker
248	Amino acid sequence for exemplary linker
249	Amino acid sequence for exemplary linker
250	Amino acid sequence for exemplary linker
251	Amino acid sequence for exemplary linker
252	Amino acid sequence for exemplary linker
253	Amino acid sequence for exemplary linker
254	Amino acid sequence for exemplary linker
255	Amino acid sequence for exemplary linker
256	Amino acid sequence for exemplary linker
257	Amino acid sequence for exemplary linker
258	Amino acid sequence for exemplary linker
259	Amino acid sequence for exemplary linker
260	Amino acid sequence for exemplary linker
261	Amino acid sequence for exemplary linker
262	Amino acid sequence for exemplary linker
263	Amino acid sequence for exemplary linker
264	Amino acid sequence for exemplary linker
265	Amino acid sequence for exemplary linker
266	Amino acid sequence for exemplary linker
267	Amino acid sequence for exemplary linker
268	Amino acid sequence for exemplary linker
269	Amino acid sequence for exemplary linker
270	Amino acid sequence for exemplary linker
271	Amino acid sequence for exemplary linker
272	Amino acid sequence for exemplary linker
273-300	Not Used
301	Exemplary mRNA encoding APOBEC3A-Nme2D16A
302	Exemplary open reading frame for APOBEC3A-Nme2D16A
303	Exemplary amino acid sequence for APOBEC3A-Nme2D16A
304	Exemplary mRNA encoding APOBEC3A-Nme2D16A
305	Exemplary open reading frame for APOBEC3A-Nme2D16A
306	Exemplary amino acid sequence for APOBEC3A-Nme2D16A
307	Exemplary mRNA encoding APOBEC3A-Nme2D16A
308	Exemplary open reading frame for APOBEC3A-Nme2D16A
309	Exemplary amino acid sequence for APOBEC3A-Nme2D16A

310	Exemplary mRNA encoding APOBEC3A-Nme2D16A
311	Exemplary open reading frame for APOBEC3A-Nme2D16A
312	Exemplary amino acid sequence for APOBEC3A-Nme2D16A
313	Exemplary amino acid sequence for NLS-NLS-APOBEC3A-L070-Nme2D16A
314	mRNA encoding BC22-2XUGI with a C-terminal HiBiT tag (BC22-2XUGI-HibIT)
315	mRNA encoding BC22-Nme2D16A (Nme2 BC22n)
316	mRNA encoding UGI with a C-terminal HiBiT tag (UGI-HiBiT)
317-319	Not Used
320	Amino acid sequence for Nme2Cas9
321-337	Exemplary amino acid sequences for base editor with linker
340	mRNA encoding Nme2Cas9
341-357	Exemplary mRNA sequences for base editor with linker
360	Open reading frame encoding Nme2Cas9
361-377	Exemplary open reading frame sequences for base editor with linker
378-386	Not Used
387	Exemplary amino acid sequence for D16A Nme2Cas9 nickase
388	Exemplary coding sequence for D16A Nme2Cas9 nickase
389	Exemplary coding sequence for D16A Nme2Cas9 nickase
390	Exemplary coding sequence for D16A Nme2Cas9 nickase
391	Exemplary open reading frame for D16A Nme2Cas9 nickase
392	Exemplary open reading frame for D16A Nme2Cas9 nickase
393	Exemplary open reading frame for D16A Nme2Cas9 nickase
394-400	Not Used
401-416	Exemplary guide RNAs targeting HLA-A
417	Not Used
418-422	Exemplary guide RNAs targeting HLA-A
423	Not Used
424	Exemplary guide RNAs targeting HLA-A
425-426	Not Used
429-435	Exemplary guide RNAs targeting HLA-A
436	Not Used
437-443	Exemplary guide RNAs targeting HLA-A
444	Not Used
445-453	Exemplary guide RNAs targeting HLA-A
454	Not Used
455-495	Exemplary guide RNAs targeting HLA-A
496	G023522 Exemplary guide RNA Targeting CD38 gene
497	G019427 Exemplary guide RNA Targeting ANAPC5
498	G023519 Exemplary guide RNA Targeting B2M
499	G023520 Exemplary guide RNA Targeting TRAC
500	G023521 Exemplary guide RNA Targeting CIITA

501	G023523 Exemplary guide RNA Targeting HLA-A
502	G023524 Exemplary guide RNA Targeting TRBC1/2
503	G020073 Exemplary Nme2Cas9 guide RNA
504	G020927 Exemplary Nme2Cas9 guide RNA
505	G020928 Exemplary Nme2Cas9 guide RNA
506	G020929 Exemplary Nme2Cas9 guide RNA
507	G021237 Exemplary Nme2Cas9 guide RNA
508	G021249 Exemplary Nme2Cas9 guide RNA
509	G021321 Exemplary Nme2Cas9 guide RNA
510	G021844 Exemplary Nme2Cas9 guide RNA with non-peptide linker
511	G000502 Exemplary guide RNA targeting TTR
512	Exemplary NmeCas9 sgRNA
513	Conserved region of exemplary shortened NmeCas9 sgRNA
514	Conserved region of Exemplary shortened NmeCas9 sgRNA pattern
515	Conserved region of Exemplary shortened NmeCas9 sgRNA pattern
516	Conserved region of Exemplary shortened/modified NmeCas9 sgRNA pattern (Mod-77)
517	Conserved region of Exemplary shortened/modified NmeCas9 sgRNA comprising linkers (Mod-78)
518	Exemplary shortened NmeCas9 sgRNA comprising linkers
519	Exemplary shortened/modified NmeCas9 guide RNA comprising linkers
520	Exemplary shortened/modified SpyCas9 guide RNA
521	Exemplary shortened SpyCas9 guide RNA
522	Exemplary shortened/modified NmeCas9 guide RNA
523	Exemplary shortened/modified SpyCas9 guide RNA comprising linkers
524	Exemplary shortened SpyCas9 guide RNA comprising linkers
525-528	Not Used
529	G016788 Exemplary guide RNA
530-617	Exemplary guide RNAs targeting CIITA
618-705	Exemplary guide sequence for TRBC gene
706-746	Exemplary guide sequence for TRAC gene
747-762	Exemplary guide RNA sequence for TRAC gene
763-800	Not Used
801-852	Exemplary guide RNA sequences for TRBC gene
853-868	Exemplary guide RNA sequences for TRAC gene
869-956	Exemplary guide RNA for CD38 gene
957-959	Not used
960-1023	Exemplary amino acid sequences of cytidine deaminases
1024-1075	Exemplary guide RNA sequences for TRBC gene
1076-1163	Exemplary guide RNA sequences for CIITA gene

See the Sequence Table below for the sequences themselves. Transcript sequences may generally include GGG as the first three nucleotides for use with ARCA or AGG as the first three nucleotides for use with CleanCap™. Accordingly, the first three nucleotides can be modified for use with other capping approaches, such as *Vaccinia* capping enzyme. Promoters and poly-A sequences are not included in the transcript sequences. A promoter such as a U6 promoter (SEQ ID NO: 67) or a CMV Promotor (SEQ ID NO: 68) and a poly-A sequence such as SEQ ID NO: 109 can be appended to the disclosed transcript sequences at the 5' and 3' ends, respectively. Most nucleotide sequences are provided as DNA but can be readily converted to RNA by changing Ts to Us.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

[0092] Reference will now be made in detail to certain embodiments of the invention, examples of which are illustrated in the accompanying drawings. While the invention will be described in conjunction with the illustrated embodiments, it will be understood that they are not intended to limit the invention to those embodiments. On the contrary, the invention is intended to cover all alternatives, modifications, and equivalents, which may be included within the invention as defined by the appended claims.

[0093] Before describing the present teachings in detail, it is to be understood that the disclosure is not limited to specific compositions or process steps, as such may vary. It should be noted that, as used in this specification and the appended claims, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. Thus, for example, reference to "a conjugate" includes a plurality of conjugates and reference to "a cell" includes a plurality of cells and the like.

[0094] Numeric ranges are inclusive of the numbers defining the range. Measured and measurable values are understood to be approximate, taking into account significant digits and the error associated with the measurement. Also, the use of "comprise", "comprises", "comprising", "contain", "contains", "containing", "include", "includes", and "including" are not intended to be limiting. It is to be understood that both the foregoing general description and detailed description are exemplary and explanatory only and are not restrictive of the teachings.

[0095] The term "about" or "approximately" means an acceptable error for a particular value as determined by one of ordinary skill in the art, which depends in part on how the value is measured or determined, or a degree of variation that does not substantially affect the properties of the described subject matter, or within the tolerances accepted in the

art, *e.g.*, within 10%, 5%, 2%, or 1%. Accordingly, unless indicated to the contrary, the numerical parameters set forth in the following specification and attached claims are approximations that may vary depending upon the desired properties sought to be obtained. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits and by applying ordinary rounding techniques.

[0096] Unless specifically noted in the above specification, embodiments in the specification that recite "comprising" various components are also contemplated as "consisting of" or "consisting essentially of" the recited components; embodiments in the specification that recite "consisting of" various components are also contemplated as "comprising" or "consisting essentially of" the recited components; and embodiments in the specification that recite "consisting essentially of" various components are also contemplated as "consisting of" or "comprising" the recited components (this interchangeability does not apply to the use of these terms in the claims).

[0097] The section headings used herein are for organizational purposes only and are not to be construed as limiting the desired subject matter in any way. In the event that any literature incorporated by reference contradicts the express content of this specification, including but not limited to a definition, the express content of this specification controls. While the present teachings are described in conjunction with various embodiments, it is not intended that the present teachings be limited to such embodiments. On the contrary, the present teachings encompass various alternatives, modifications, and equivalents, as will be appreciated by those of skill in the art.

I. Definitions

[0098] Unless stated otherwise, the following terms and phrases as used herein are intended to have the following meanings:

[0099] The term "or combinations thereof" as used herein refers to all permutations and combinations of the listed terms preceding the term. For example, "A, B, C, or combinations thereof" is intended to include at least one of: A, B, C, AB, AC, BC, or ABC, and if order is important in a particular context, also BA, CA, CB, ACB, CBA, BCA, BAC, or CAB. Continuing with this example, expressly included are combinations that contain repeats of one or more item or term, such as BB, AAA, AAB, BBC, CBBA, CABA, and so forth. The skilled artisan will understand that typically there is no limit on the number of items or terms in any combination, unless otherwise apparent from the context.

[00100] As used herein, the term "kit" refers to a packaged set of related components, such as one or more polynucleotides or compositions and one or more related materials such as delivery devices (e.g., syringes), solvents, solutions, buffers, instructions, or desiccants.

[00101] "Or" is used in the inclusive sense, i.e., equivalent to "and/or," unless the context requires otherwise.

[00102] "Polynucleotide" and "nucleic acid" are used herein to refer to a multimeric compound comprising nucleosides or nucleoside analogs which have nitrogenous heterocyclic bases or base analogs linked together along a backbone, including conventional RNA, DNA, mixed RNA-DNA, and polymers that are analogs thereof. A nucleic acid "backbone" can be made up of a variety of linkages, including one or more of sugar-phosphodiester linkages, peptide-nucleic acid bonds ("peptide nucleic acids" or PNA; PCT No. WO 95/32305), phosphorothioate linkages, methylphosphonate linkages, or combinations thereof. Sugar moieties of a nucleic acid can be ribose, deoxyribose, or similar compounds with substitutions, e.g., 2' methoxy or 2' halide substitutions. Nitrogenous bases can be conventional bases (A, G, C, T, U), analogs thereof (e.g., modified uridines such as 5-methoxyuridine, pseudouridine, or N1-methylpseudouridine, or others); inosine; derivatives of purines or pyrimidines (e.g., N^4-methyl deoxyguanosine, deaza- or aza-purines, deaza- or aza-pyrimidines, pyrimidine bases with substituent groups at the 5 or 6 position (e.g., 5-methylcytosine), purine bases with a substituent at the 2, 6, or 8 positions, 2-amino-6-methylaminopurine, O^6-methylguanine, 4-thio-pyrimidines, 4-amino-pyrimidines, 4-dimethylhydrazine-pyrimidines, and O^4-alkyl-pyrimidines; US Pat. No. 5,378,825 and PCT No. WO 93/13121). For general discussion see *The Biochemistry of the Nucleic Acids* 5-36, Adams et al., ed., 11th ed., 1992). Nucleic acids can include one or more "abasic" residues where the backbone includes no nitrogenous base for position(s) of the polymer (US Pat. No. 5,585,481). A nucleic acid can comprise only conventional RNA or DNA sugars, bases and linkages, or can include both conventional components and substitutions (e.g., conventional bases with 2' methoxy linkages, or polymers containing both conventional bases and one or more base analogs). Nucleic acid includes "locked nucleic acid" (LNA), an analogue containing one or more LNA nucleotide monomers with a bicyclic furanose unit locked in an RNA mimicking sugar conformation, which enhance hybridization affinity toward complementary RNA and DNA sequences (Vester and Wengel, 2004, *Biochemistry* 43(42):13233-41). RNA and DNA have different sugar moieties and can differ by the presence of uracil or analogs thereof in RNA and thymine or analogs thereof in DNA.

[00103] "Polypeptide" as used herein refers to a multimeric compound comprising amino acid residues that can adopt a three-dimensional conformation. Polypeptides include but are not limited to enzymes, enzyme precursor proteins, regulatory proteins, structural proteins, receptors, nucleic acid binding proteins, antibodies, etc. Polypeptides may, but do not necessarily, comprise post-translational modifications, non-natural amino acids, prosthetic groups, and the like.

[00104] As used herein, a "cytidine deaminase" means a polypeptide or complex of polypeptides that is capable of cytidine deaminase activity, that is catalyzing the hydrolytic deamination of cytidine or deoxycytidine, typically resulting in uridine or deoxyuridine. Cytidine deaminases encompass enzymes in the cytidine deaminase superfamily, and in particular, enzymes of the APOBEC family (APOBEC1, APOBEC2, APOBEC4, and APOBEC3 subgroups of enzymes), activation-induced cytidine deaminase (AID or AICDA) and CMP deaminases (see, e.g., Conticello et al., Mol. Biol. Evol. 22:367-77, 2005; Conticello, Genome Biol. 9:229, 2008; Muramatsu et al., J. Biol. Chem. 274: 18470-6, 1999); Carrington et al., Cells 9:1690 (2020)). In some embodiments, variants of any known cytidine deaminase or APOBEC protein are encompassed. Variants include proteins having a sequence that differs from wild-type protein by one or several mutations (i.e., substitutions, deletions, insertions), such as one or several single point substitutions. For instance, a shortened sequence could be used, *e.g.*, by deleting N-terminal, C-terminal, or internal amino acids, preferably one to four amino acids at the C-terminus of the sequence. As used herein, the term "variant" refers to allelic variants, splicing variants, and natural or artificial mutants, which are homologous to a reference sequence. The variant is "functional" in that it shows a catalytic activity of DNA editing.

[00105] As used herein, the term "APOBEC3A" refers to a cytidine deaminase such as the protein expressed by the human A3A gene. The APOBEC3A may have catalytic DNA editing activity. An amino acid sequence of APOBEC3A has been described (UniPROT accession ID: p31941) and is included herein as SEQ ID NO: 40. In some embodiments, the APOBEC3A protein is a human APOBEC3A protein and/or a wild-type protein. Variants include proteins having a sequence that differs from wild-type APOBEC3A protein by one or several mutations (i.e., substitutions, deletions, insertions), such as one or several single point substitutions. For instance, a shortened APOBEC3A sequence could be used, e.g. by deleting N-terminal, C-terminal, or internal amino acids, preferably one to four amino acids at the C-terminus of the sequence. As used herein, the term "variant" refers to allelic variants, splicing variants, and natural or artificial mutants, which are homologous to

an APOBEC3A reference sequence. The variant is "functional" in that it shows a catalytic activity of DNA editing. In some embodiments, an APOBEC3A (such as a human APOBEC3A) has a wild-type amino acid position 57 (as numbered in the wild-type sequence). In some embodiments, an APOBEC3A (such as a human APOBEC3A) has an asparagine at amino acid position 57 (as numbered in the wild-type sequence).

[00106] As used herein, a "nickase" is an enzyme that creates a single-strand break (also known as a "nick") in double strand DNA, *i.e.*, cuts one strand but not the other of the DNA double helix. As used herein, an "RNA-guided nickase" means a polypeptide or complex of polypeptides having DNA nickase activity, wherein the DNA nickase activity is sequence-specific and depends on the sequence of the RNA. Exemplary RNA-guided nickases include Cas nickases. Cas nickases include, but are not limited to, nickase forms of a Csm or Cmr complex of a type III CRISPR system, the Cas10, Csm1, or Cmr2 subunit thereof, a Cascade complex of a type I CRISPR system, the Cas3 subunit thereof, and Class 2 Cas nucleases. Class 2 Cas nickases include Class 2 Cas nuclease variants in which only one of the two catalytic domains is inactivated, which have RNA-guided DNA nickase activity. Class 2 Cas nickases include, for example, Cas9 (*e.g.*, H840A, D10A, or N863A variants of SpyCas9), Cpf1, C2c1, C2c2, C2c3, HF Cas9 (*e.g.*, N497A, R661A, Q695A, Q926A variants), HypaCas9 (*e.g.*, N692A, M694A, Q695A, H698A variants), eSPCas9(1.0) (*e.g.*, K810A, K1003A, R1060A variants), and eSPCas9(1.1) (*e.g.*, K848A, K1003A, R1060A variants) proteins and modifications thereof. Cpf1 protein, Zetsche et al., *Cell*, 163: 1-13 (2015), is homologous to Cas9, and contains a RuvC-like protein domain. Cpf1 sequences of Zetsche are incorporated by reference in their entirety. *See*, *e.g.*, Zetsche, Tables S1 and S3. "Cas9" encompasses *S. pyogenes* (Spy) Cas9, the variants of Cas9 listed herein, and equivalents thereof. *See*, *e.g.*, Makarova et al., *Nat Rev Microbiol*, 13(11): 722-36 (2015); Shmakov et al., *Molecular Cell,* 60:385-397 (2015).

[00107] Several Cas9 orthologs have been obtained from *N. meningitidis* (Esvelt et al., NAT. METHODS, vol. 10, 2013, 1116 - 1121; Hou et al., PNAS, vol. 110, 2013, pages 15644 – 15649; Edraki et al., Mol. Cell 73:714-726, 2019) (Nme1Cas9, Nme2Cas9, and Nme3Cas9). The Nme2Cas9 ortholog functions efficiently in mammalian cells, recognizes an N4CC PAM, and can be used for *in vivo* editing (Ran et al., NATURE, vol. 520, 2015, pages 186 - 191; Kim et al., NAT. COMMUN., vol. 8, 2017, pages 14500). Nme2Cas9 has been shown to be naturally resistant to off-target editing (Lee et al., MOL. THER., vol. 24, 2016, pages 645 - 654; Kim et al., 2017). See also *e.g.*, WO/2020081568 (*e.g.*, pages 28 and 42), describing an Nme2Cas9 D16A nickase, the contents of which are

hereby incorporated by reference in its entirety. Throughout, "NmeCas9" is generic and an encompasses any type of NmeCas9, including, Nme1Cas9, Nme2Cas9, and Nme3Cas9.

[00108] As used herein, the term "fusion protein" refers to a hybrid polypeptide which comprises polypeptides from at least two different proteins or sources. One polypeptide may be located at the amino-terminal (N-terminal) portion of the fusion protein or at the carboxy-terminal (C- terminal) protein thus forming an "amino-terminal fusion protein" or a "carboxy-terminal fusion protein," respectively. Any of the proteins provided herein may be produced by any method known in the art. For example, the proteins provided herein may be produced via recombinant protein expression and purification, which is especially suited for fusion proteins comprising a peptide linker. Methods for recombinant protein expression and purification are well known, and include those described by Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), the entire contents of which are incorporated herein by reference.

[00109] The term "linker," as used herein, refers to a chemical group or a molecule linking two adjacent molecules or moieties. Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond. In some embodiments, the linker is an amino acid or a plurality of amino acids (*e.g.*, a peptide or protein) such as a 16-amino acid residue "XTEN" linker, or a variant thereof (See, *e.g.*, the Examples; and Schellenberger et al. A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. Nat. Biotechnol. 27, 1186-1190 (2009)). In some embodiments, the XTEN linker comprises the sequence SGSETPGTSESATPES (SEQ ID NO: 46), SGSETPGTSESA (SEQ ID NO: 47), or SGSETPGTSESATPEGGSGGS (SEQ ID NO: 48). In some embodiments, the linker comprises one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272.

[00110] As used herein, the term "uracil glycosylase inhibitor", "uracil-DNA glycosylase inhibitor" or "UGI" refers to a protein that is capable of inhibiting a uracil-DNA glycosylase (UDG) base-excision repair enzyme (*e.g.*, UniPROT ID: P14739; SEQ ID NO: 27; SEQ ID NO:43).

[00111] As used herein, "open reading frame" or "ORF" of a gene refers to a sequence consisting of a series of codons that specify the amino acid sequence of the protein that the gene codes for. The ORF generally begins with a start codon (e.g., ATG in DNA or AUG in RNA) and ends with a stop codon, e.g., TAA, TAG or TGA in DNA or UAA, UAG, or UGA in RNA.

[00112] "Guide RNA", "gRNA", and "guide" are used herein interchangeably to refer to either a crRNA (also known as CRISPR RNA), or the combination of a crRNA and a trRNA (also known as tracrRNA). The crRNA and trRNA may be associated as a single RNA molecule (single guide RNA, sgRNA) or in two separate RNA molecules (dual guide RNA, dgRNA). "Guide RNA" or "gRNA" refers to each type. The trRNA may be a naturally-occurring sequence, or a trRNA sequence with modifications or variations compared to naturally-occurring sequences.

[00113] As used herein, a "guide sequence" or "guide region" or "spacer" or "spacer sequence" and the like refers to a sequence within a gRNA that is complementary to a target sequence and functions to direct a gRNA to a target sequence for binding or modification (*e.g.*, cleavage) by an RNA-guided nickase. A guide sequence can be 20 nucleotides in length, *e.g.*, in the case of *Streptococcus pyogenes* (*i.e.*, Spy Cas9 (also referred to as SpCas9)) and related Cas9 homologs/orthologs. Shorter or longer sequences can also be used as guides, e.g., 15-, 16-, 17-, 18-, 19-, 21-, 22-, 23-, 24-, or 25-nucleotides in length. A guide sequence can be 20-25 nucleotides in length, *e.g.*, in the case of Nme Cas9, *e.g.*, 20-, 21-, 22-, 23-, 24-or 25-nucleotides in length. For example, a guide sequence of 24 nucleotides in length can be used with Nme Cas9, *e.g.*, Nme2 Cas9.

[00114] In some embodiments, the target sequence is in a gene or on a chromosome, for example, and is complementary to the guide sequence. In some embodiments, the degree of complementarity or identity between a guide sequence and its corresponding target sequence may be about 75%, 80%, 85%, 90%, 95%, or 100%. In some embodiments, the guide sequence and the target region may be 100% complementary or identical. In other embodiments, the guide sequence and the target region may contain at least one mismatch. For example, the guide sequence and the target sequence may contain 1, 2, 3, or 4 mismatches, where the total length of the target sequence is at least 17, 18, 19, 20 or more base pairs. In some embodiments, the guide sequence and the target region may contain 1-4 mismatches where the guide sequence comprises at least 17, 18, 19, 20 or more nucleotides. In some embodiments, the guide sequence and the target region may contain 1, 2, 3, or 4 mismatches where the guide sequence comprises 20 nucleotides.

[00115] As used herein, a "target sequence" or "genomic target sequence" refers to a sequence of nucleic acid in a target gene that has complementarity to the guide sequence of the gRNA. The interaction of the target sequence and the guide sequence directs an RNA-guided DNA binding agent to bind, and potentially nick or cleave (depending on the activity of the agent), within the target sequence. Target sequences for Cas proteins include

both the positive and negative strands of genomic DNA *(i.e.*, the sequence given and the sequence's reverse compliment), as a nucleic acid substrate for a Cas protein is a double stranded nucleic acid. Accordingly, where a guide sequence is said to be "complementary to a target sequence," it is to be understood that the guide sequence may direct an RNA-guided DNA binding agent (*e.g.*, dCas9 or impaired Cas9) to bind to the reverse complement of a target sequence. Thus, in some embodiments, where the guide sequence binds the reverse complement of a target sequence, the guide sequence is identical to certain nucleotides of the target sequence (*e.g.*, the target sequence not including the PAM) except for the substitution of U for T in the guide sequence.

[00116] As used herein, a first sequence is considered to "comprise a sequence with at least X% identity to" a second sequence if an alignment of the first sequence to the second sequence shows that X% or more of the positions of the second sequence in its entirety are matched by the first sequence. For example, the sequence AAGA comprises a sequence with 100% identity to the sequence AAG because an alignment would give 100% identity in that there are matches to all three positions of the second sequence. The differences between RNA and DNA (generally the exchange of uridine for thymidine or vice versa) and the presence of nucleoside analogs such as modified uridines do not contribute to differences in identity or complementarity among polynucleotides as long as the relevant nucleotides (such as thymidine, uridine, or modified uridine) have the same complement (e.g., adenosine for all of thymidine, uridine, or modified uridine; another example is cytosine and 5-methylcytosine, both of which have guanosine as a complement). Thus, for example, the sequence 5'-AXG where X is any modified uridine, such as pseudouridine, N1-methyl pseudouridine, or 5-methoxyuridine, is considered 100% identical to AUG in that both are perfectly complementary to the same sequence (5'-CAU). Exemplary alignment algorithms are the Smith-Waterman and Needleman-Wunsch algorithms, which are well-known in the art. One skilled in the art will understand what choice of algorithm and parameter settings are appropriate for a given pair of sequences to be aligned; for sequences of generally similar length and expected identity >50% for amino acids or >75% for nucleotides, the Needleman-Wunsch algorithm with default settings of the Needleman-Wunsch algorithm interface provided by the EBI at the www.ebi.ac.uk web server are generally appropriate.

[00117] "mRNA" is used herein to refer to a polynucleotide that is not DNA and comprises an open reading frame that can be translated into a polypeptide (*i.e.*, can serve as a substrate for translation by a ribosome and amino-acylated tRNAs). mRNA can comprise a phosphate-sugar backbone including ribose residues or analogs thereof, *e.g.*, 2'-methoxy

ribose residues. In some embodiments, the sugars of an mRNA phosphate-sugar backbone consist essentially of ribose residues, 2'-methoxy ribose residues, or a combination thereof. In general, mRNAs do not contain a substantial quantity of thymidine residues (*e.g.*, 0 residues or fewer than 30, 20, 10, 5, 4, 3, or 2 thymidine residues; or less than 10%, 9%, 8%, 7%, 6%, 5%, 4%, 4%, 3%, 2%, 1%, 0.5%, 0.2%, or 0.1% thymidine content). An mRNA can contain modified uridines at some or all of its uridine positions.

[00118] "Modified uridine" is used herein to refer to a nucleoside other than thymidine with the same hydrogen bond acceptors as uridine and one or more structural differences from uridine. In some embodiments, a modified uridine is a substituted uridine, i.e., a uridine in which one or more non-proton substituents (*e.g.*, alkoxy, such as methoxy) takes the place of a proton. In some embodiments, a modified uridine is pseudouridine. In some embodiments, a modified uridine is a substituted pseudouridine, i.e., a pseudouridine in which one or more non-proton substituents (*e.g.*, alkyl, such as methyl) takes the place of a proton. In some embodiments, a modified uridine is any of a substituted uridine, pseudouridine, or a substituted pseudouridine.

[00119] "Uridine position" as used herein refers to a position in a polynucleotide occupied by a uridine or a modified uridine. Thus, for example, a polynucleotide in which "100% of the uridine positions are modified uridines" contains a modified uridine at every position that would be a uridine in a conventional RNA (where all bases are standard A, U, C, or G bases) of the same sequence. Unless otherwise indicated, a U in a polynucleotide sequence of a sequence table or sequence listing in or accompanying this disclosure can be a uridine or a modified uridine.

[00120] As used herein, the "minimal uridine codon(s)" for a given amino acid is the codon(s) with the fewest uridines (usually 0 or 1 except for a codon for phenylalanine, where the minimal uridine codon has 2 uridines). Modified uridine residues are considered equivalent to uridines for the purpose of evaluating uridine content.

[00121] As used herein, the "uridine dinucleotide (UU) content" of an ORF can be expressed in absolute terms as the enumeration of UU dinucleotides in an ORF or on a rate basis as the percentage of positions occupied by the uridines of uridine dinucleotides (for example, AUUAU would have a uridine dinucleotide content of 40% because 2 of 5 positions are occupied by the uridines of a uridine dinucleotide). Modified uridine residues are considered equivalent to uridines for the purpose of evaluating uridine dinucleotide content.

[00122] As used herein, the "minimal adenine codon(s)" for a given amino acid is the codon(s) with the fewest adenines (usually 0 or 1 except for a codon for lysine and

asparagine, where the minimal adenine codon has 2 adenines). Modified adenine residues are considered equivalent to adenines for the purpose of evaluating adenine content.

[00123] As used herein, the "adenine dinucleotide content" of an ORF can be expressed in absolute terms as the enumeration of AA dinucleotides in an ORF or on a rate basis as the percentage of positions occupied by the adenines of adenine dinucleotides (for example, UAAUA would have an adenine dinucleotide content of 40% because 2 of 5 positions are occupied by the adenines of an adenine dinucleotide). Modified adenine residues are considered equivalent to adenines for the purpose of evaluating adenine dinucleotide content.

[00124] As used herein, "indels" refer to insertion/deletion mutations consisting of a number of nucleotides that are either inserted or deleted, e.g., at the site of double-stranded breaks (DSBs) in a target nucleic acid.

[00125] As used herein, "knockdown" refers to a decrease in expression of a particular gene product (*e.g.*, protein, mRNA, or both). Knockdown of a protein can be measured either by detecting protein secreted by tissue or population of cells (*e.g.*, in serum or cell media) or by detecting total cellular amount of the protein from a tissue or cell population of interest. Methods for measuring knockdown of mRNA are known and include sequencing of mRNA isolated from a tissue or cell population of interest. In some embodiments, "knockdown" may refer to some loss of expression of a particular gene product, for example a decrease in the amount of mRNA transcribed or a decrease in the amount of protein expressed or secreted by a population of cells (including *in vivo* populations such as those found in tissues).

[00126] As used herein, "knockout" refers to a loss of expression of a particular protein in a cell. Knockout can be measured either by detecting the amount of protein secretion from a tissue or population of cells (*e.g.*, in serum or cell media) or by detecting total cellular amount of a protein a tissue or a population of cells. In some embodiments, the methods of the disclosure "knockout" a target protein one or more cells (e.g., in a population of cells including *in vivo* populations such as those found in tissues). In some embodiments, a knockout is not the formation of mutant of the target protein, for example, created by indels, but rather the complete loss of expression of the target protein in a cell, *i.e.*, decrease of expression to below the level of detection of the assay used.

[00127] As used herein, the terms "nuclear localization signal" (NLS) or "nuclear localization sequence" refers to an amino acid sequence which induces transport of molecules comprising such sequences or linked to such sequences into the nucleus of

eukaryotic cells. The nuclear localization signal may form part of the molecule to be transported. In some embodiments, the NLS may be fused to the molecule by a covalent bond, hydrogen bonds or ionic interactions. In some embodiments, the NLS may be fused to the molecule via a linker.

[00128] "β2M" or "B2M," as used herein, refers to nucleic acid sequence or protein sequence of "β-2 microglobulin;" the human gene has accession number NC_000015 (range 44711492..44718877), reference GRCh38.p13. The B2M protein is associated with MHC class I molecules as a heterodimer on the surface of nucleated cells and is required for MHC class I protein expression.

[00129] "*CIITA*" or "CIITA" or "C2TA," as used herein, refers to the nucleic acid sequence or protein sequence of "class II major histocompatibility complex transactivator;" the human gene has accession number NC_000016.10 (range 10866208..10941562), reference GRCh38.p13. The CIITA protein in the nucleus acts as a positive regulator of MHC class II gene transcription and is required for MHC class II protein expression.

[00130] As used herein, "MHC" or "MHC molecule(s)" or "MHC protein" or "MHC complex(es)," refers to a major histocompatibility complex molecule (or plural), and includes, *e.g.*, MHC class I and MHC class II molecules. In humans, MHC molecules are referred to as "human leukocyte antigen" complexes or "HLA molecules" or "HLA protein." The use of terms "MHC" and "HLA" are not meant to be limiting; as used herein, the term "MHC" may be used to refer to human MHC molecules, *i.e.*, HLA molecules. Therefore, the terms "MHC" and "HLA" are used interchangeably herein.

[00131] The term "HLA-A," as used herein in the context of HLA-A protein, refers to the MHC class I protein molecule, which is a heterodimer consisting of a heavy chain (encoded by the HLA-A gene) and a light chain (*i.e.*, beta-2 microglobulin). The term "HLA-A" or "HLA-A gene," as used herein in the context of nucleic acids refers to the gene encoding the heavy chain of the HLA-A protein molecule. The HLA-A gene is also referred to as "HLA class I histocompatibility, A alpha chain;" the human gene has accession number NC_000006.12 (29942532..29945870). The HLA-A gene is known to have thousands of different versions (also referred to as "alleles") across the population (and an individual may receive two different alleles of the HLA-A gene). A public database for HLA-A alleles, including sequence information, may be accessed at IPD-IMGT/HLA: https://www.ebi.ac.uk/ipd/imgt/hla/. All alleles of HLA-A are encompassed by the terms "HLA-A" and "HLA-A gene."

[00132] "HLA-B" as used herein in the context of nucleic acids refers to the gene encoding the heavy chain of the HLA-B protein molecule. The HLA-B is also referred to as "HLA class I histocompatibility, B alpha chain;" the human gene has accession number NC_000006.12 (31353875..31357179).

[00133] "HLA-C" as used herein in the context of nucleic acids refers to the gene encoding the heavy chain of the HLA-C protein molecule. The HLA-C is also referred to as "HLA class I histocompatibility, C alpha chain;" the human gene has accession number NC_000006.12 (31268749..31272092).

[00134] "TRBC1" and "TRBC2" as used herein in the context of nucleic acids refer to two homologous genes encoding the T-cell receptor β-chain. "TRBC" or "TRBC1/2" is used herein to refer to TRBC1 and TRBC2. The human wild-type TRBC1 sequence is available at NCBI Gene ID: 28639; Ensembl: ENSG00000211751. T-cell receptor Beta Constant, V_segment Translation Product, BV05S1J2.2, TCRBC1, and TCRB are gene synonyms for TRBC1. The human wild-type TRBC2 sequence is available at NCBI Gene ID: 28638; Ensembl: ENSG00000211772. T-cell receptor Beta Constant, V_segment Translation Product, and TCRBC2 are gene synonyms for TRBC2.

[00135] "TRAC" as used herein in the context of nucleic acids refers to the gene encoding the T-cell receptor α-chain. The human wild-type TRAC sequence is available at NCBI Gene ID: 28755; Ensembl: ENSG00000277734. T-cell receptor Alpha Constant, TCRA, IMD7, TRCA and TRA are gene synonyms for TRAC.

[00136] As used herein, the term "homozygous" refers to having two identical alleles of a particular gene.

[00137] As used herein, "treatment" refers to any administration or application of a therapeutic for disease or disorder in a subject, and includes inhibiting the disease, arresting its development, relieving one or more symptoms of the disease, curing the disease, or preventing one or more symptoms of the disease, including reoccurrence of the symptom.

[00138] As used herein, "delivering" and "administering" are used interchangeably, and include ex vivo and in vivo applications.

[00139] Co-administration, as used herein, means that a plurality of substances are administered sufficiently close together in time so that the agents act together. Co-administration encompasses administering substances together in a single formulation and administering substances in separate formulations close enough in time so that the agents act together.

[00140] As used herein, the phrase "pharmaceutically acceptable" means that which is useful in preparing a pharmaceutical composition that is generally non-toxic and is not biologically undesirable and that are not otherwise unacceptable for pharmaceutical use. Pharmaceutically acceptable generally refers to substances that are non-pyrogenic. Pharmaceutically acceptable can refer to substances that are sterile, especially for pharmaceutical substances that are for injection or infusion.

[00141] As used herein, a "subject" refers to any member of the animal kingdom. In some embodiments, "subject" refers to humans. In some embodiments, "subject" refers to non-human animals. In some embodiments, "subject" refers to primates. In some embodiments, subjects include, but are not limited to, mammals, birds, reptiles, amphibians, fish, insects, and/or worms. In certain embodiments, the non-human subject is a mammal (*e.g.*, a rodent, a mouse, a rat, a rabbit, a monkey, a dog, a cat, a sheep, cattle, a primate, and/or a pig). In some embodiments, a subject may be a transgenic animal, genetically-engineered animal, and/or a clone. In certain embodiments of the present invention the subject is an adult, an adolescent or an infant. In some embodiments, terms "individual" or "patient" are used and are intended to be interchangeable with "subject".

[00142] As used herein, "reduced or eliminated" expression of a protein on a cell refers to a partial or complete loss of expression of the protein relative to an unmodified cell. In some embodiments, the surface expression of a protein on a cell is measured by flow cytometry and has "reduced or eliminated" surface expression relative to an unmodified cell as evidenced by a reduction in fluorescence signal upon staining with the same antibody against the protein. A cell that has "reduced or eliminated" surface expression of a protein by flow cytometry relative to an unmodified cell may be referred to as "negative" for expression of that protein as evidenced by a fluorescence signal similar to a cell stained with an isotype control antibody. The "reduction or elimination" of protein expression can be measured by other known techniques in the field with appropriate controls known to those skilled in the art.

II. Exemplary compositions and methods

[00143] In some embodiments, a nucleic acid is provided, the nucleic acid comprising an open reading frame encoding a polypeptide comprising a cytidine deaminase (*e.g.*, A3A) and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI). In some embodiments, the nucleic acid is DNA or RNA. In

some embodiments, the nucleic acid is mRNA. In some embodiments, a polypeptide encoded by the mRNA is provided.

[00144] In some embodiments, a polypeptide or an mRNA encoding the polypeptide, are provided, the polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the polypeptide does not comprise a UGI. In some embodiments, the cytidine deaminase is A3A. In some embodiments, the RNA-guided nickase does not comprise a uracil glycosylase inhibitor (UGI). In some embodiments, a composition is provided comprising a first polypeptide, or an mRNA encoding a first polypeptide, comprising a cytidine deaminase (*e.g.*, A3A) and an RNA-guided nickase; and a second polypeptide, or an mRNA encoding a second polypeptide, comprising a uracil glycosylase inhibitor (UGI), wherein the second polypeptide is different from the first polypeptide.

[00145] In some embodiments, a composition is provided comprising a first nucleic acid comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (*e.g.*, A3A) and an RNA-guided nickase, and a second nucleic acid comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second nucleic acid is different from the first nucleic acid. In some embodiments, the first nucleic acid encodes a polypeptide that does not comprise a UGI.

[00146] In some embodiments, methods of modifying a target gene are provided comprising administering the compositions described herein. In some embodiments, the method comprises delivering to a cell a first nucleic acid comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase (*e.g.*, A3A) and an RNA-guided nickase, and a second nucleic acid comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second nucleic acid is different from the first nucleic acid.

[00147] In some embodiments, the methods comprise delivering to a cell a polypeptide comprising a cytidine deaminase (*e.g.*, A3A) and an RNA-guided nickase, or a nucleic acid encoding the polypeptide, and separately (e.g., not via the same nucleic acid construct) delivering to the cell a uracil glycosylase inhibitor (UGI), or a nucleic acid encoding the UGI.

[00148] In some embodiments, a molar ratio of the mRNA encoding UGI to the mRNA encoding the cytidine deaminase (*e.g.*, A3A) and the RNA-guided nickase is from about 1:35 to from about 30:1. In some embodiments, the molar ratio is from about 1:25 to about 25:1. In some embodiments, the molar ratio is from about 1:20 to about 25:1. In some embodiments, the molar ratio is from about 1:10 to about 22:1. In some embodiments, the

molar ratio is from about 1:5 to about 25:1. In some embodiments, the molar ratio is from about 1:1 to about 30:1. In some embodiments, the molar ratio is from about 2:1 to about 10:1. In some embodiments, the molar ratio is from about 5:1 to about 20:1. In some embodiments, the molar ratio is from about 1:1 to about 25:1. In some embodiments, the molar ratio may be about 1:35, 1:34, 1:33, 1:32, 1:31, 1:30, 1:32, 1:31, 1:30, 1:29, 1:28, 1:27, 1:26, 1:25, 1:24, 1:23, 1:22, 1:21, 1:20, 1:19, 1:18, 1:17, 1:16, 1:15, 1:14, 1:13, 1:12, 1:11, 1:10, 1:9, 1:8, 1:7, 1:6, 1:5, 1:4, 1:3, 1:2, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 21:1, 22:1, 23:1, 24:1, 25:1, 26:1, 27:1, 28:1, 29:1, or 30:1. In some embodiments, the molar ratio is equal to or larger than about 1:1. In some embodiments the molar ratio is about 1:1. In some embodiments the molar ratio is about 2:1. In some embodiments the molar ratio is about 3:1. In some embodiments the molar ratio is about 4:1. In some embodiments the molar ratio is about 5:1. In some embodiments the molar ratio is about 6:1. In some embodiments the molar ratio is about 7:1. In some embodiments the molar ratio is about 8:1. In some embodiments the molar ratio is about 9:1. In some embodiments the molar ratio is about 10:1. In some embodiments the molar ratio is about 11:1. In some embodiments the molar ratio is about 12:1. In some embodiments the molar ratio is about 13:1. In some embodiments the molar ratio is about 14:1. In some embodiments the molar ratio is about 15:1. In some embodiments the molar ratio is about 16:1. In some embodiments the molar ratio is about 17:1. In some embodiments the molar ratio is about 18:1. In some embodiments the molar ratio is about 19:1. In some embodiments the molar ratio is about 20:1. In some embodiments the molar ratio is about 21:1. In some embodiments the molar ratio is about 22:1. In some embodiments the molar ratio is about 23:1. In some embodiments the molar ratio is about 24:1. In some embodiments the molar ratio is about 25:1.

[00149] Similarly, in some embodiments, the molar ratio discussed above for the mRNA encoding the UGI protein to the mRNA encoding the cytidine deaminase (*e.g.*, A3A) and the RNA-guided nickase are similar if delivering protein.

[00150] For example, in some embodiments, a molar ratio of the UGI protein to be delivered to the cytidine deaminase (*e.g.*, A3A) and the RNA-guided nickase to be delivered is from about 1:35 to from about 30:1. In some embodiments, the molar ratio is from about 1:1 to about 30:1.

[00151] In some embodiments, the molar ratio of the UGI peptide and the cytidine deaminase (*e.g.*, A3A) and the RNA-guided nickase is from about 10:1 to about 50:1. In some embodiments, the molar ratio may be about 10:1, 11:1, 12:1, 13:1, 14:1, 15:1,

16:1, 17:1, 18:1, 19:1, 20:1, 21:1, 22:1, 23:1, 24:1, 25:1, 26:1, 27:1, 28:1, 29:1, 30:1, 31:1, 32:1, 33:1, 34:1, 35:1, 36:1, 37:1, 38:1, 39:1, 40:1, 41:1, 42:1, 43:1, 44:1, 45:1, 46:1, 47:1, 48:1, 49:1, or 50:1. In some embodiments, the molar ratio is from about 10:1- about 40:1. In some embodiments the molar ratio is from about 10:1- about 30:1. In some embodiments the molar ratio is about 2:1. In some embodiments the molar ratio is from about 10:1- about 20:1. In some embodiments the molar ratio is from about 10:1- about 15:1. In some embodiments the molar ratio is about 15:1- about 50:1. In some embodiments the molar ratio is about 6:1. In some embodiments the molar ratio is about 20:1- about 50:1. In some embodiments the molar ratio is about 8:1. In some embodiments the molar ratio is about 30:1- about 50:1. In some embodiments the molar ratio is about 30:1- about 40:1. In some embodiments the molar ratio is about 11:1. In some embodiments the molar ratio is about 20:1-about 30:1.

[00152] In some embodiments, the composition described herein further comprises at least one gRNA. In some embodiments, a composition is provided that comprises an mRNA described herein and at least one gRNA. In some embodiments, the gRNA is a single guide RNA (sgRNA). In some embodiments, the gRNA is a dual guide RNA (dgRNA).

[00153] In some embodiments, the composition is capable of effecting genome editing upon administration to the subject.

A. UGI

[00154] Without being bound by any theory, providing a UGI together with a polypeptide comprising a deaminase may be helpful in the methods described herein by inhibiting cellular DNA repair machinery (e.g., UDG and downstream repair effectors) that recognize a uracil in DNA as a form of DNA damage or otherwise would excise or modify the uracil and/or surrounding nucleotides. It should be understood that the use of a UGI may increase the editing efficiency of an enzyme that is capable of deaminating C residues.

[00155] Suitable UGI protein and nucleotide sequences are provided herein and additional suitable UGI sequences are known to those in the art, and include, for example, those published in Wang et al., Uracil-DNA glycosylase inhibitor gene of bacteriophage PBS2 encodes a binding protein specific for uracil-DNA glycosylase. J. Biol. Chem. 264: 1163-1171(1989); Lundquist et al., Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein. Role of specific carboxylic amino acids in complex formation with Escherichia coli uracil-DNA glycosylase. J. Biol. Chem. 272:21408-21419(1997); Ravishankar et al., X-ray analysis of a complex of Escherichia coli uracil DNA

glycosylase (EcUDG) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG. Nucleic Acids Res. 26:4880-4887(1998); and Putnam et al., Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with Escherichia coli uracil-DNA glycosylase. J. Mol. Biol. 287:331-346(1999), the entire contents of each are incorporated herein by reference. It should be appreciated that any proteins that are capable of inhibiting a uracil-DNA glycosylase base-excision repair enzyme are within the scope of the present disclosure. Additionally, any proteins that block or inhibit base-excision repair are also within the scope of this disclosure. In some embodiments, a uracil glycosylase inhibitor is a protein that binds uracil. In some embodiments, a uracil glycosylase inhibitor is a protein that binds uracil in DNA. In some embodiments, a uracil glycosylase inhibitor is a single-stranded binding protein. In some embodiments, a uracil glycosylase inhibitor is a catalytically inactive uracil DNA-glycosylase protein. In some embodiments, a uracil glycosylase inhibitor is a catalytically inactive uracil DNA-glycosylase protein that does not excise uracil from the DNA. In some embodiments, a uracil glycosylase inhibitor is a catalytically inactive UDG.

[00156] In some embodiments, a uracil glycosylase inhibitor (UGI) disclosed herein comprises an amino acid sequence with at least 80% to SEQ ID NO: 27 or 43. In some embodiments, any of the foregoing levels of identity is at least 90%, at least 95%, at least 98%, at least 99%, or 100%. In some embodiments, the UGI comprises an amino acid sequence with at least 90% identity to SEQ ID NO: 27 or 43. In some embodiments, the UGI comprises an amino acid sequence with at least 95% identity to SEQ ID NO: 27 or 43. In some embodiments, the UGI comprises an amino acid sequence with at least 98% identity to SEQ ID NO: 27 or 43. In some embodiments, the UGI comprises an amino acid sequence with at least 99% identity to SEQ ID NO: 27 or 43. In some embodiments, the UGI comprises the amino acid sequence of SEQ ID NO: 27 or 43.

B. Cytidine Deaminase

[00157] Cytidine deaminases encompass enzymes in the cytidine deaminase superfamily, and in particular, enzymes of the APOBEC family (APOBEC1, APOBEC2, APOBEC4, and APOBEC3 subgroups of enzymes), activation-induced cytidine deaminase (AID or AICDA) and CMP deaminases (see, *e.g.*, Conticello et al., Mol. Biol. Evol. 22:367-77, 2005; Conticello, Genome Biol. 9:229, 2008; Muramatsu et al., J. Biol. Chem. 274: 18470-6, 1999); and Carrington et al., Cells 9:1690 (2020)).

[00158] In some embodiments, the cytidine deaminase disclosed herein is an enzyme of APOBEC family. In some embodiments, the cytidine deaminase disclosed herein is an enzyme of APOBEC1, APOBEC2, APOBEC4, and APOBEC3 subgroups. In some embodiments, the cytidine deaminase disclosed herein is an enzyme of APOBEC3 subgroup. In some embodiments, the cytidine deaminase disclosed herein is an APOBEC3A deaminase (A3A).

[00159] In some embodiments, the cytidine deaminase is:

(i) an enzyme of the APOBEC family, optionally an enzyme of APOBEC3 subgroup;

(ii) a cytidine deaminase comprising an amino acid sequence that is at least 80 % identical to any one of SEQ ID NOs: 40, 41, and 960-1023;

(iii) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1013;

(iv) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009; or

(v) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 976, 981, 984, 986, and 1014-1023.

[00160] In some embodiments, the cytidine deaminase is a cytidine deaminase comprising an amino acid sequence having at least 80%, 85% 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 41, and 960-1023. In some embodiments, the cytidine deaminase is a cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identical to any one of SEQ ID NOs: 40, 41, and 960-1013. In some embodiments, the cytidine deaminase is a cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the cytidine deaminase is a cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identical to any one of SEQ ID NOs: 40, 976, 981, 984, 986, and 1014-1023. In some embodiments, the cytidine deaminase is a cytidine deaminase comprising an amino acid sequence that is at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identical to SEQ ID NO: 976, 977, 993-1006, and 1009.

1. APOBEC3A Deaminase

[00161] In some embodiments, an APOBEC3A deaminase (A3A) disclosed herein is a human A3A. In some embodiments, the A3A is a wild-type A3A.

[00162] In some embodiment, the A3A is an A3A variant. A3A variants share homology to wild-type A3A, or a fragment thereof. In some embodiments, a A3A variant has at least about 80% identity, at least about 85% identity, at least about 90% identity, at least about 95% identity, at least about 96% identity, at least about 97% identity, at least about 98% identity, at least about 99% identity, at least about 99.5% identity, or at least about 99.9% identity to a wild type A3A. In some embodiments, the A3A variant may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more amino acid changes compared to a wild type A3A. In some embodiments, the A3A variant comprises a fragment of an A3A, such that the fragment has at least about 80% identity, at least about 90% identity, at least about 95% identity, at least about 96% identity, at least about 97% identity, at least about 98% identity, at least about 99% identity, at least about 99.5% identity, or at least about 99.9% identity to the corresponding fragment of a wild-type A3A.

[00163] In some embodiments, an A3A variant is a protein having a sequence that differs from a wild-type A3A protein by one or several mutations, such as substitutions, deletions, insertions, one or several single point substitutions. In some embodiments, a shortened A3A sequence could be used, *e.g.*, by deleting N-terminal, C-terminal, or internal amino acids. In some embodiments, a shortened A3A sequence is used where one to four amino acids at the C-terminus of the sequence is deleted. In some embodiments, an APOBEC3A (such as a human APOBEC3A) has a wild-type amino acid position 57 (as numbered in the wild-type sequence). In some embodiments, an APOBEC3A (such as a human APOBEC3A) has an asparagine at amino acid position 57 (as numbered in the wild-type sequence).

[00164] In some embodiments, the wild-type A3A is a human A3A (UniPROT accession ID: p319411, SEQ ID NO: 40).

[00165] In some embodiments, the A3A disclosed herein comprises an amino acid sequence having at least 80% identity to SEQ ID NO: 40. In some embodiments, the level of identity is at least 85%, at least 87%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%. In some embodiments, the A3A comprises an amino acid sequence having at least 87% identity to SEQ ID NO: 40. In some embodiments, the A3A comprises an amino

acid sequence with at least 90% identity to SEQ ID NO: 40. In some embodiments, the A3A comprises an amino acid sequence with at least 95% identity to SEQ ID NO: 40. In some embodiments, the A3A comprises an amino acid sequence with at least 98% identity to SEQ ID NO: 40. In some embodiments, the A3A comprises an amino acid sequence with at least 99% identity to A3A ID NO: 40. In some embodiments, the A3A comprises the amino acid sequence of SEQ ID NO: 40.

C. Linkers

[00166] In some embodiments, the polypeptide comprising the A3A and the RNA-guided nickase described herein further comprises a linker that connects the A3A and the RNA-guided nickase. In some embodiments, the linker is an organic molecule, polymer, or chemical moiety. In some embodiments, the linker is a peptide linker. In some embodiments, the nucleic acid encoding the polypeptide comprising the A3A and the RNA-guided nickase further comprises a sequence encoding the peptide linker. mRNAs encoding the A3A-linker-RNA-guided nickase fusion protein are provided.

[00167] In some embodiments, the peptide linker is any stretch of amino acids having at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, or more amino acids.

[00168] In some embodiments, the peptide linker is the 16 residue "XTEN" linker, or a variant thereof (*See, e.g.*, the Examples; and Schellenberger et al. A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner. Nat. Biotechnol. 27, 1186-1190 (2009)). In some embodiments, the XTEN linker comprises a sequence that is any one of SGSETPGTSESATPES (SEQ ID NO: 46), SGSETPGTSESA (SEQ ID NO: 47), or SGSETPGTSESATPEGGSGGS (SEQ ID NO: 48). In some embodiments, the XTEN linker consists of the sequence SGSETPGTSESATPES (SEQ ID NO: 46), SGSETPGTSESA (SEQ ID NO: 47), or SGSETPGTSESATPEGGSGGS (SEQ ID NO: 48).

[00169] In some embodiments, the peptide linker comprises a $(GGGGS)_n$ (e.g., SEQ ID NOs: 212, 216, 221, 240), a $(G)_n$, an $(EAAAK)_n$ (e.g., SEQ ID NOs: 213, 219, 267), a $(GGS)_n$, an SGSETPGTSESATPES (SEQ ID NO: 46) motif (*see, e.g.*, Guilinger J P, Thompson D B, Liu D R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 2014; 32(6): 577-82; the entire contents are incorporated herein by reference), or an $(XP)_n$ motif, or a combination of any of these,

wherein n is independently an integer between 1 and 30. *See*, WO2015089406, *e.g.*, paragraph [0012], the entire content of which is incorporated herein by reference.

[00170] In some embodiments, the peptide linker comprises one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272. In some embodiments, the peptide linker comprises one or more sequences selected from SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 268, SEQ ID NO: 269, SEQ ID NO: 270. SEQ ID NO: 271 and SEQ ID NO: 272. In some embodiments, the peptide linker comprises a sequence of SEQ ID NO: 268.

D. RNA-guided nickase

[00171] In some embodiments, an RNA-guided nickase disclosed herein is a Cas nickase. In some embodiments, a RNA-guided nickase is from a specific Cas nuclease with its catalytic domain(s) being inactivated. In some embodiments, the RNA-guided nickase is a Class 2 Cas nickase, such as a Cas9 nickase or a Cpf1 nickase. In some embodiments, the RNA-guided nickase is an *S. pyogenes* Cas9 nickase. In some embodiments, the RNA-guided nickase is *Neisseria meningitidis* Cas9 nickase.

[00172] In some embodiments, the RNA-guided nickase is a modified Class 2 Cas protein or derived from a Class 2 Cas protein. In some embodiments, the RNA-guided nickase is modified or derived from a Cas protein, such as a Class 2 Cas nuclease (which may be, *e.g.*, a Cas nuclease of Type II, V, or VI). Class 2 Cas nuclease include, for example, Cas9, Cpf1, C2c1, C2c2, and C2c3 proteins and modifications thereof. Examples of Cas9 nucleases include those of the type II CRISPR systems of *S. pyogenes, S. aureus*, and other prokaryotes (*see, e.g.*, the list in the next paragraph), and modified (*e.g.*, engineered or mutant) versions thereof. *See, e.g.*, US2016/0312198 A1; US 2016/0312199 A1, which is incorporated by reference in its entirety. Other examples of Cas nucleases include a Csm or Cmr complex of a type III CRISPR system or the Cas10, Csm1, or Cmr2 subunit thereof; and a Cascade complex of a type I CRISPR system, or the Cas3 subunit thereof. In some embodiments, the Cas nuclease may be from a Type-IIA, Type-IIB, or Type-IIC system. For discussion of various CRISPR systems and Cas nucleases, *see, e.g.*, Makarova et al., NAT. REV. MICROBIOL. 9:467-477 (2011); Makarova et al., NAT. REV. MICROBIOL, 13: 722-36 (2015); Shmakov et al., MOLECULAR CELL, 60:385-397 (2015).

[00173] A Cas nickase described herein may be a nickase form of a Cas nuclease from the species including, but not limited to, Streptococcus pyogenes, Streptococcus thermophilus, Streptococcus sp., Staphylococcus aureus, Listeria innocua,

Lactobacillus gasseri, Francisella novicida, Wolinella succinogenes, Sutterella wadsworthensis, Gammaproteobacterium, Neisseria meningitidis, Campylobacter jejuni, Pasteurella multocida, Fibrobacter succinogene, Rhodospirillum rubrum, Nocardiopsis dassonvillei, Streptomyces pristinaespiralis, Streptomyces viridochromogenes, Streptomyces viridochromogenes, Streptosporangium roseum, Streptosporangium roseum, Alicyclobacillus acidocaldarius, Bacillus pseudomycoides, Bacillus selenitireducens, Exiguobacterium sibiricum, Lactobacillus delbrueckii, Lactobacillus salivarius, Lactobacillus buchneri, Treponema denticola, Microscilla marina, Burkholderiales bacterium, Polaromonas naphthalenivorans, Polaromonas sp., Crocosphaera watsonii, Cyanothece sp., Microcystis aeruginosa, Synechococcus sp., Acetohalobium arabaticum, Ammonifex degensii, Caldicelulosiruptor becscii, Candidatus Desulforudis, Clostridium botulinum, Clostridium difficile, Finegoldia magna, Natranaerobius thermophilus, Pelotomaculum thermopropionicum, Acidithiobacillus caldus, Acidithiobacillus ferrooxidans, Allochromatium vinosum, Marinobacter sp., Nitrosococcus halophilus, Nitrosococcus watsoni, Pseudoalteromonas haloplanktis, Ktedonobacter racemifer, Methanohalobium evestigatum, Anabaena variabilis, Nodularia spumigena, Nostoc sp., Arthrospira maxima, Arthrospira platensis, Arthrospira sp., Lyngbya sp., Microcoleus chthonoplastes, Oscillatoria sp., Petrotoga mobilis, Thermosipho africanus, Streptococcus pasteurianus, Neisseria cinerea, Campylobacter lari, Parvibaculum lavamentivorans, Corynebacterium diphtheria, Acidaminococcus sp., Lachnospiraceae bacterium ND2006, or Acaryochloris marina.

[00174] In some embodiments, the Cas nickase is a nickase form of the Cas9 nuclease from *Streptococcus pyogenes*. In some embodiments, the Cas nickase is a nickase form of the Cas9 nuclease from *Streptococcus thermophilus*. In some embodiments, the Cas nickase is a nickase form of the Cas9 nuclease from *Neisseria meningitidis*. See e.g., WO/2020081568, describing an Nme2Cas9 D16A nickase. In some embodiments, the Cas nickase is a nickase form of the Cas9 nuclease from *Staphylococcus aureus*. In some embodiments, the Cas nickase is a nickase form of the Cpf1 nuclease from *Francisella novicida*. In some embodiments, the Cas nickase is a nickase form of the Cpf1 nuclease from *Acidaminococcus sp.* In some embodiments, the Cas nickase is a nickase form of the Cpf1 nuclease from *Lachnospiraceae bacterium* ND2006. In further embodiments, the Cas nickase is a nickase form of the Cpf1 nuclease from *Francisella tularensis, Lachnospiraceae bacterium, Butyrivibrio proteoclasticus, Peregrinibacteria bacterium, Parcubacteria bacterium, Smithella, Acidaminococcus, Candidatus Methanoplasma termitum, Eubacterium eligens, Moraxella bovoculi, Leptospira inadai, Porphyromonas crevioricanis, Prevotella*

disiens, or Porphyromonas macacae. In certain embodiments, the Cas nickase is a nickase form of a Cpf1 nuclease from an *Acidaminococcus* or *Lachnospiraceae*. As discussed elsewhere, a nickase may be derived from (i.e. related to) a specific Cas nuclease in that the nickase is a form of the nuclease in which one of its two catalytic domains is inactivated, e.g., by mutating an active site residue essential for nucleolysis, such as D10, H840, or N863 in Spy Cas9. One skilled in the art will be familiar with techniques for easily identifying corresponding residues in other Cas proteins, such as sequence alignment and structural alignment, which is discussed in detail below.

[00175] In other embodiments, the Cas nickase may relate to a Type-I CRISPR/Cas system. In some embodiments, the Cas nickase may be a component of the Cascade complex of a Type-I CRISPR/Cas system. In some embodiments, the Cas nickase may be a Cas3 protein. In some embodiments, the Cas nickase may be from a Type-III CRISPR/Cas system.

[00176] In some embodiments, a Cas nickase is a nickase form of a Cas nuclease or a modified Cas nuclease in which an endonucleolytic active site is inactivated, e.g., by one or more alterations (e.g., point mutations) in a catalytic domain. See, e.g., US Pat. No. 8,889,356 for discussion of Cas nickases and exemplary catalytic domain alterations.

[00177] Wild type S. *pyogenes* Cas9 has two catalytic domains: RuvC and HNH. The RuvC domain cleaves the non-target DNA strand, and the HNH domain cleaves the target strand of DNA. In some embodiments, a Cas nuclease may comprise an amino acid substitution in the RuvC or RuvC-like nuclease domain. Exemplary amino acid substitutions in the RuvC or RuvC-like nuclease domain include D10A (based on the S. *pyogenes* Cas9 protein). *See, e.g.*, Zetsche et al. (2015) *Cell* Oct 22:163(3): 759-771. In some embodiments, the Cas nuclease may comprise an amino acid substitution in the HNH or HNH-like nuclease domain. Exemplary amino acid substitutions in the HNH or HNH-like nuclease domain include E762A, H840A, N863A, H983A, and D986A (based on the S. *pyogenes* Cas9 protein). *See, e.g.*, Zetsche et al. (2015). Further exemplary amino acid substitutions include D917A, E1006A, and D1255A (based on the *Francisella novicida* U112 Cpf1 (FnCpf1) sequence (UniProtKB - A0Q7Q2 (CPF1_FRATN)).

[00178] In some embodiments, a Cas nickase such as a Cas9 nickase has an inactivated RuvC or HNH domain. In some embodiments, a nickase is used having a RuvC domain with reduced activity. In some embodiments, a nickase is used having an inactive RuvC domain. In some embodiments, a nickase is used having an HNH domain with reduced activity. In some embodiments, a nickase is used having an inactive HNH domain.

[00179] In some embodiments, a Cas9 nickase has an active HNH nuclease domain and is able to cleave the non-targeted strand of DNA, i.e., the strand bound by the gRNA and has an inactive RuvC nuclease domain and is not able to cleave the targeted strand of the DNA, i.e., the strand where base editing by deaminase is desired.

[00180] An exemplary Cas9 nickase amino acid sequence is provided as SEQ ID NO: 70. An exemplary Cas9 nickase mRNA ORF sequence, which includes start and stop codons, is provided as SEQ ID NO: 71. An exemplary Cas9 nickase mRNA coding sequence, suitable for inclusion in a fusion protein, is provided as SEQ ID NO: 72.

[00181] In some embodiments, the RNA-guided nickase is a Class 2 Cas nickase described herein. In some embodiments, the RNA-guided nickase is a Cas9 nickase described herein.

[00182] In some embodiments, the RNA-guided nickase is an *S. pyogenes* Cas9 nickase described herein.

[00183] In some embodiments, the RNA-guided nickase is a D10A SpyCas9 nickase described herein. In some embodiments, the RNA-guided nickase comprises an amino acid sequence having at least 80%, 90%, 95%, 98%, or 99% identity to any one of SEQ ID NO: 70, 73, or 76. In some embodiments, the RNA-guided nickase comprises the amino acid sequence of SEQ ID NO: 70.

[00184] In some embodiments, the mRNA ORF sequence comprises encoding the RNA-guided nickase, which includes start and stop codons, comprises a nucleotide sequence having at least 80%, 90%, 95%, 98%, 99% or 100% identity to the nucleotide sequence of any one of SEQ ID NOs: 71, 74, or 77. In some embodiments, the mRNA sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 80%, 90%, 95%, 98%, 99% or 100% identity to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78. In some embodiments, the level of identity is at least 90%. In some embodiments, the level of identity is at least 95%. In some embodiments, the level of identity is at least 98%. In some embodiments, the level of identity is at least 99%. In some embodiments, the level of identity is at least 100%. In some embodiments, the sequence encoding the RNA-guided nickase comprises the nucleotide sequence of any one of SEQ ID NOs: 71, 72, 74, 75, 77, or 78.

[00185] In some embodiments, the RNA-guided nickase is *Neisseria meningitidis* (Nme) Cas9 nickase described herein.

[00186] In some embodiments, the RNA-guided nickase is a D16A NmeCas9 nickase described herein. In some embodiments, the D16A NmeCas9 nickase is a D16A

Nme2Cas9 nickase. In some embodiments, the D16A Nme2Cas9 nickase comprises an amino acid sequence at least 80%, 90%, 95%, 98%, 99% or 100% identical to SEQ ID NO: 387. In some embodiments, the sequence encoding the D16A Nme2Cas9 comprises a nucleotide sequence at least 80%, 90%, 95%, 98%, 99% or 100% identical to any one of SEQ ID NOs: 388-393.

E. Compositions comprising a cytidine deaminase and an RNA-guided nickase

[00187] In some embodiments, an mRNA encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase is provided, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

1. Exemplary Compositions

[00188] As described herein, compositions, methods, and uses are provided comprising an mRNA comprising an open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI). For each exemplary composition described below, the mRNA does not comprise a UGI.

[00189] In some embodiments, an mRNA encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase is provided. In some embodiments, an enzyme of APOBEC family and an RNA-guided nickase is provided. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and an RNA-guided nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC2 subgroup and an RNA-guided nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC4 subgroup and an RNA-guided nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and an RNA-guided nickase.

[00190] In some embodiments, an mRNA encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase is provided. In some embodiments, an enzyme of APOBEC family and a D10A SpyCas9 nickase is provided. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D10A SpyCas9 nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC2 subgroup and a D10A SpyCas9 nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC4 subgroup and a D10A SpyCas9 nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D10A SpyCas9 nickase.

[00191] In some embodiments, an mRNA encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase is provided. In some embodiments, an enzyme of APOBEC family and a D16A NmeCas9 nickase is provided. In some embodiments, an enzyme of APOBEC family and a D16A Nme2Cas9 nickase is provided. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D16A Nme2Cas9 nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC2 subgroup and a D16A Nme2Cas9 nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC4 subgroup and a D16A Nme2Cas9 nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D16A Nme2Cas9 nickase.

[00192] In some embodiments, the polypeptide lacks a UGI.

[00193] In some embodiments, the cytidine deaminase and the RNA-guided nickase are linked via a linker. In some embodiments, the cytidine deaminase and the RNA-guided nickase are linked via a peptide linker. In some embodiments, the peptide linker comprises one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272.

[00194] In some embodiments, the polypeptide further comprises one or more additional heterologous functional domains. In some embodiments, the polypeptide further comprises one or more nuclear localization sequences (NLSs) (described herein) at the C-terminal of the polypeptide or the N-terminal of the polypeptide.

[00195] In some embodiments, an mRNA encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase is provided. In some embodiments, an enzyme of APOBEC family and an RNA-guided nickase is provided. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and an RNA-guided nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC2 subgroup and an RNA-guided nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC4 subgroup and an RNA-guided nickase. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and an RNA-guided nickase.

[00196] In some embodiments, an mRNA encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase is provided. In some embodiments, an enzyme of APOBEC family and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC family and the D10A SpyCas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D10A SpyCas9 nickase, and a nuclear localization sequence (NLS) at the C-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D10A

SpyCas9 nickase, and a NLS at the N-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC family and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9 nickase, optionally via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC family and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9 nickase, optionally via a linker.

[00197] In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D16A NmeCas9 nickase, wherein the enzyme of APOBEC family and the D16A NmeCas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC family and the D16A Nme2Cas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D16A Nme2Cas9 nickase, and a nuclear localization sequence (NLS) at the C-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D16A Nme2Cas9 nickase, and a NLS at the N-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC family and the D16A Nme2Cas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D16A Nme2Cas9 nickase, optionally via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC family and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC family and the D16A Nme2Cas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D16A Nme2Cas9 nickase, optionally via a linker.

[00198] In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC1 subgroup and the D10A SpyCas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D10A SpyCas9 nickase, and a nuclear localization sequence (NLS) at the C-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D10A SpyCas9 nickase, and a NLS at the N-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC1 subgroup and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9

nickase, optionally via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC1 subgroup and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9 nickase, optionally via a linker.

[00199] In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC1 subgroup and the D16A Nme2Cas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC1 subgroup and the D16A Nme2Cas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D16A Nme2Cas9 nickase, and a nuclear localization sequence (NLS) at the C-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D16A Nme2Cas9 nickase, and a NLS at the N-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC1 subgroup and the D16A Nme2Cas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D16A Nme2Cas9 nickase, optionally via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC1 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC1 subgroup and the D16A Nme2Cas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D16A Nme2Cas9 nickase, optionally via a linker.

[00200] In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC3 subgroup and the D10A SpyCas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D10A SpyCas9 nickase, and a nuclear localization sequence (NLS) at the C-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D10A SpyCas9 nickase, and a NLS at the N-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC3 subgroup and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9 nickase, optionally via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D10A SpyCas9 nickase, wherein the enzyme of APOBEC3

subgroup and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9 nickase, optionally via a linker.

[00201] In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC3 subgroup and the D16A Nme2Cas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC3 subgroup and the D16A Nme2Cas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D16A Nme2Cas9 nickase, and a nuclear localization sequence (NLS) at the C-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D16A Nme2Cas9 nickase, and a NLS at the N-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC3 subgroup and the D16A Nme2Cas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D16A Nme2Cas9 nickase, optionally via a linker. In some embodiments, the polypeptide comprises an enzyme of APOBEC3 subgroup and a D16A Nme2Cas9 nickase, wherein the enzyme of APOBEC3 subgroup and the D16A Nme2Cas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D16A Nme2Cas9 nickase, optionally via a linker.

[00202] In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 268, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 269, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 270, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 271, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a

D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 272, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In any of the foregoing embodiments, the D10A SpyCas9 nickase may comprise an amino acid sequence that is at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to any one of SEQ ID NOs: 70, 73, or 76.

[00203] In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 268, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 269, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 270, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 271, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 272, and a cytidine deaminase comprising an amino acid sequence that is at least 85% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In any of the foregoing embodiments, the D16A Nme2Cas9 nickase may comprise an amino acid sequence that is at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to SEQ ID NO: 387.

[00204] In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 268, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 269, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some

embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 270, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 271, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D10A SpyCas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 272, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In any of the foregoing embodiments, the D10A SpyCas9 comprises an amino acid sequence that is at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to any one of SEQ ID NOs: 70, 73, or 76.

[00205] In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 268, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 269, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 270, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 271, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009. In some embodiments, the polypeptide comprises a D16A Nme2Cas9 nickase, a linker comprising the amino acid sequence of SEQ ID NO: 272, and a cytidine deaminase comprising an amino acid sequence selected from any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, *984-987, 993-1006, and 1009. In any of the foregoing embodiments, the D16A Nme2Cas9 nickase comprises an amino acid sequence that is at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identical to SEQ ID NO: 387.*

[00206] The polypeptide may be organized in any number of ways to form a single chain. The NLS can be N- or C-terminal, or both N- and C-terminals, and the cytidine deaminase can be N- or C-terminal as compared the RNA-guided nickase. In some embodiments, the polypeptide comprises, from N to C terminus, a cytidine deaminase, an optional linker, an RNA-guided nickase, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an RNA-guided nickase, an optional linker, a cytidine deaminase, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, and a cytidine deaminase. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, and a cytidine deaminase, and an optional NLS.

[00207] In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an enzyme of APOBEC family, an optional linker, an RNA-guided nickase, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, an enzyme of APOBEC family and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, an enzyme of APOBEC family, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an enzyme of APOBEC family, an optional linker, an RNA-guided nickase, and an optional NLS.

[00208] In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an enzyme of APOBEC3 subgroup, an optional linker, an RNA-guided nickase, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, an enzyme of APOBEC3 subgroup and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, an enzyme of APOBEC3 subgroup, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, an enzyme of APOBEC3 subgroup, and an optional NLS.

[00209] In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an enzyme of APOBEC family, an optional linker, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, an optional linker, an enzyme of APOBEC family and an

optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, an optional linker, an enzyme of APOBEC family, and an optional NLS,. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, an optional linker, and an enzyme of APOBEC family, and an optional NLS.

[00210] In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an enzyme of APOBEC3 subgroup, an optional linker, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, an optional linker, an enzyme of APOBEC3 subgroup and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, an optional linker, an enzyme of APOBEC3 subgroup, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, an optional linker, and an enzyme of APOBEC3 subgroup, and an optional NLS.

[00211] In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an enzyme of APOBEC3 subgroup, an optional linker, a D16A Nme2Cas9 nickase.

[00212] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS; (ii) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023; (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272, (iv) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, and (v) an optional NLS.

[00213] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS, (ii) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272, (iv) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023, and (v) an optional NLS.

[00214] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS, (ii) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and

211-272, (iv) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023, and (v) an optional NLS.

[00215] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS, (ii) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272, (iv) cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023, and (v) an optional NLS.

[00216] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS, (ii) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023; (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272, (iv) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, and (v) an optional NLS.

[00217] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS, (ii) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272, (iv) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023, and (v) an optional NLS.

[00218] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS, (ii) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272, (iv) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023, and (v) an optional NLS.

[00219] In some embodiments, the polypeptide comprises, from N to C terminus, (i) an optional NLS, (ii) a D10A SpyCas9 nickase or a D16A Nme2Cas9 nickase, (iii) a linker comprising one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272, and (iv) cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023, and (v) an optional NLS.

2. Compositions comprising an APOBEC3A deaminase and an RNA-guided nickase

[00220] In some embodiments, an mRNA encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase is provided. In some embodiments, the polypeptide comprises a human A3A and an RNA-guided nickase. In some embodiments, the polypeptide comprises a wild-type A3A and an RNA-guided nickase. In

some embodiments, the polypeptide comprises an A3A variant and an RNA-guided nickase. In some embodiments, the polypeptide comprises an A3A and a Cas9 nickase. In some embodiments, the polypeptide comprises an A3A and a D10A SpyCas9 nickase. In some embodiments, the polypeptide comprises a human A3A and a D10A SpyCas9 nickase. In some embodiments, the polypeptide comprises an A3A variant and a D10A SpyCas9 nickase. In some embodiments, the polypeptide lacks a UGI. In some embodiments, the A3A and the RNA-guided nickase are linked via a linker. In some embodiments, the polypeptide further comprises one or more additional heterologous functional domains. In some embodiments, the polypeptide further comprises a nuclear localization sequence (NLS) (described herein) at the C-terminal of the polypeptide or the N-terminal of the polypeptide.

[00221] In some embodiments, the polypeptide comprises a human A3A and a D10A SpyCas9 nickase, wherein the human A3A and the D10A SpyCas9 nickase are fused via a linker. In some embodiments, the polypeptide comprises a human A3A and a D10A SpyCas9 nickase, and a nuclear localization sequence (NLS) at the C-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises a human A3A and a D10A SpyCas9 nickase, and a NLS at the N-terminus of the fused polypeptide. In some embodiments, the polypeptide comprises a human A3A and a D10A SpyCas9 nickase, wherein the human A3A and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9 nickase, optionally via a linker. In some embodiments, the polypeptide comprises a human A3A and a D10A SpyCas9 nickase, wherein the human A3A and the D10A SpyCas9 nickase are fused via a linker, and a NLS fused to the C-terminus of the D10A SpyCas9 nickase, optionally via a linker.

[00222] The polypeptide may be organized in any number of ways to form a single chain. The NLS can be N- or C-terminal, or both N- and C-terminals. and the A3A can be N- or C-terminal as compared the RNA-guided nickase. In some embodiments, the polypeptide comprises, from N to C terminus, an A3A, an optional linker, an RNA-guided nickase, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an RNA-guided nickase, an optional linker, an A3A, and an optional NLS. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, and an A3A. In some embodiments, the polypeptide comprises, from N to C terminus, an optional NLS, an RNA-guided nickase, an optional linker, and an A3A, and an optional NLS.

[00223] In any of the foregoing embodiments, the polypeptide may comprise an amino acid sequence having at least 80% identity to SEQ ID NOs: 3 or 6. In some

embodiments, any of the foregoing levels of identity is at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%. In some embodiments, the polypeptide disclosed herein may comprise an amino acid sequence with at least 90% identity to SEQ ID NOs: 3 or 6. In some embodiments, the polypeptide disclosed herein may comprise an amino acid sequence with at least 95% identity to SEQ ID NOs: 3 or 6. In some embodiments, the polypeptide disclosed herein may comprise an amino acid sequence with at least 98% identity to SEQ ID NOs: 3 or 6. In some embodiments, the polypeptide disclosed herein may comprise an amino acid sequence with at least 99% identity to SEQ ID NOs: 3 or 6. In some embodiments, the polypeptide disclosed herein may comprise an amino acid sequence of SEQ ID NOs: 3 or 6.

[00224] In any of the foregoing embodiments, a nucleic acid sequence comprising an open reading frame encoding the polypeptide disclosed herein may comprise a nucleic acid sequence having at least 80% identity to SEQ ID NOs: 2 or 5. In some embodiments, any of the foregoing levels of identity is at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%.

[00225] In any of the foregoing embodiments, an mRNA sequence encoding the polypeptide disclosed herein may comprise a nucleic acid sequence having at least 80% identity to SEQ ID NOs: 1 or 4. In some embodiments, any of the foregoing levels of identity is at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%.

[00226] In any of the foregoing embodiments, the polypeptide may comprise an amino acid sequence having at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or 100% identity to SEQ ID NOs: 303, 306, 309, or 312. In some embodiments, the polypeptide disclosed herein may comprise an amino acid sequence of SEQ ID NOs: 303, 306, 309, or 312. In any of the foregoing embodiments, a nucleic acid sequence comprising an open reading frame encoding the polypeptide disclosed herein may comprise a nucleic acid sequence having at least 80%, at least 90%, at least 95%, at least 98% , at least 99%, or 100% identity to SEQ ID NOs: SEQ ID NOs: 302, 305, 308, or 311. In some embodiments, a nucleic acid sequence comprising an open reading frame encoding the polypeptide disclosed herein comprises a nucleic acid sequence of SEQ ID NOs: SEQ ID NOs: 302, 305, 308, or 311. In any of the foregoing embodiments, an mRNA sequence encoding the polypeptide disclosed herein may comprise a nucleic acid sequence having at least 80%, at least 90%, at least 95%, at least 98% , at least 99%, or 100% identity to SEQ ID NOs: 301, 304, 307, or 310. In any of the foregoing embodiments, an mRNA sequence encoding the polypeptide

disclosed herein may comprise a nucleic acid sequence of SEQ ID NOs: 301, 304, 307, or 310.

[00227] In any of the foregoing embodiments, the A3A may comprise an amino acid sequence having at least 80% identity to SEQ ID NO: 40. In some embodiments, the level of identity is at least 85%, at least 87%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%. In some embodiments, the A3A comprises an amino acid sequence of SEQ ID NO: 40.

[00228] In any of the foregoing embodiments, the RNA-guided nickase may comprise an amino acid sequence having at least 80%, 90%, 95%, 98%, or 99% identity to any one of SEQ ID NOs: 70, 73, or 76. In some embodiments, the level of identity is at least 85%, at least 87%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%. In some embodiments, the RNA-guided nickase comprises the amino acid sequence of SEQ ID NO: 70. In some embodiments, the RNA-guided nickase comprises the amino acid sequence of SEQ ID NO: 73. In some embodiments, the RNA-guided nickase comprises the amino acid sequence of SEQ ID NO: 76.

[00229] In any of the foregoing embodiments, the A3A may comprise an amino acid sequence having at least 80% identity to SEQ ID NO: 40 and the RNA-guided nickase may comprise an amino acid sequence having at least 80%, 90%, 95%, 98%, or 99% identity to any one of SEQ ID NOs: 70, 73, or 76. In some embodiments, the A3A comprises an amino acid sequence of SEQ ID NO: 40 and the RNA-guided nickase comprises an amino acid sequence of SEQ ID NO: 70.

F. Additional Features

1. Codon-optimization

[00230] In some embodiments, the UGI or polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase) and an RNA-guided nickase is encoded by an open reading frame (ORF) comprising a codon optimized nucleic acid sequence. In some embodiment, the codon optimized nucleic acid sequence comprises minimal adenine codons and/or minimal uridine codons.

[00231] A given ORF can be reduced in uridine content or uridine dinucleotide content, for example, by using minimal uridine codons in a sufficient fraction of the ORF. For example, an amino acid sequence for the polypeptide described herein can be back-

translated into an ORF sequence by converting amino acids to codons, wherein some or all of the ORF uses the exemplary minimal uridine codons shown below. In some embodiments, at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% of the codons in the ORF are codons listed in Table 1.

Table 1. Exemplary minimal uridine codons

	Amino Acid	Minimal uridine codon
A	Alanine	GCA or GCC or GCG
G	Glycine	GGA or GGC or GGG
V	Valine	GUC or GUA or GUG
D	Aspartic acid	GAC
E	Glutamic acid	GAA or GAG
I	Isoleucine	AUC or AUA
T	Threonine	ACA or ACC or ACG
N	Asparagine	AAC
K	Lysine	AAG or AAA
S	Serine	AGC
R	Arginine	AGA or AGG
L	Leucine	CUG or CUA or CUC
P	Proline	CCG or CCA or CCC
H	Histidine	CAC
Q	Glutamine	CAG or CAA
F	Phenylalanine	UUC
Y	Tyrosine	UAC
C	Cysteine	UGC
W	Tryptophan	UGG
M	Methionine	AUG

[00232] In some embodiments, the ORF may consist of a set of codons of which at least about 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% of the codons are codons listed in Table 1.

[00233] A given ORF can be reduced in adenine content or adenine dinucleotide content, for example, by using minimal adenine codons in a sufficient fraction of the ORF. For example, an amino acid sequence for the polypeptide described herein can be back-translated into an ORF sequence by converting amino acids to codons, wherein some or all of the ORF uses the exemplary minimal adenine codons shown below. In some embodiments, at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% of the codons in the ORF are codons listed in Table 2.

Table 2. Exemplary minimal adenine codons

	Amino Acid	Minimal adenine codon
A	Alanine	GCU or GCC or GCG
G	Glycine	GGU or GGC or GGG
V	Valine	GUC or GUU or GUG
D	Aspartic acid	GAC or GAU
E	Glutamic acid	GAG
I	Isoleucine	AUC or AUU
T	Threonine	ACU or ACC or ACG
N	Asparagine	AAC or AAU
K	Lysine	AAG
S	Serine	UCU or UCC or UCG
R	Arginine	CGU or CGC or CGG
L	Leucine	CUG or CUC or CUU
P	Proline	CCG or CCU or CCC
H	Histidine	CAC or CAU
Q	Glutamine	CAG
F	Phenylalanine	UUC or UUU
Y	Tyrosine	UAC or UAU
C	Cysteine	UGC or UGU
W	Tryptophan	UGG
M	Methionine	AUG

[00234] In some embodiments, the ORF may consist of a set of codons of which at least about 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% of the codons are codons listed in Table 2.

[00235] To the extent feasible, any of the features described above with respect to low adenine content can be combined with any of the features described above with respect to low uridine content. So too for uridine and adenine dinucleotides. Similarly, the content of uridine nucleotides and adenine dinucleotides in the ORF may be as set forth above. Similarly, the content of uridine dinucleotides and adenine nucleotides in the ORF may be as set forth above.

[00236] A given ORF can be reduced in uridine and adenine nucleotide and/or dinucleotide content, for example, by using minimal uridine and adenine codons in a sufficient fraction of the ORF. For example, an amino acid sequence for the polypeptide described herein can be back-translated into an ORF sequence by converting amino acids to codons, wherein some or all of the ORF uses the exemplary minimal uridine and adenine codons shown below. In some embodiments, at least about 50%, 55%, 60%, 65%, 70%, 75%,

80%, 85%, 90%, 95%, 98%, 99%, or 100% of the codons in the ORF are codons listed in Table 3.

Table 3. Exemplary minimal uridine and adenine codons

	Amino Acid	Minimal uridine and adenine codon
A	Alanine	GCC or GCG
G	Glycine	GGC or GGG
V	Valine	GUC or GUG
D	Aspartic acid	GAC
E	Glutamic acid	GAG
I	Isoleucine	AUC
T	Threonine	ACC or ACG
N	Asparagine	AAC
K	Lysine	AAG
S	Serine	AGC or UCC or UCG
R	Arginine	CGC or CGG
L	Leucine	CUG or CUC
P	Proline	CCG or CCC
H	Histidine	CAC
Q	Glutamine	CAG
F	Phenylalanine	UUC
Y	Tyrosine	UAC
C	Cysteine	UGC
W	Tryptophan	UGG
M	Methionine	AUG

[00237] In some embodiments, the ORF may consist of a set of codons of which at least about 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% of the codons are codons listed in Table 3. As can be seen in Table 3, each of the three listed serine codons contains either one A or one U. In some embodiments, uridine minimization is prioritized by using AGC codons for serine. In some embodiments, adenine minimization is prioritized by using UCC and/or UCG codons for serine.

[00238] In some embodiments, the ORF may have codons that increase translation in a mammal, such as a human. In further embodiments, the mRNA comprises an ORF having codons that increase translation in an organ, such as the liver, of the mammal, e.g., a human. In further embodiments, the ORF may have codons that increase translation in a cell type, such as a hepatocyte, of the mammal, e.g., a human. An increase in translation in a mammal, cell type, organ of a mammal, human, organ of a human, etc., can be determined relative to the extent of translation wild-type sequence of the ORF, or relative to an ORF

having a codon distribution matching the codon distribution of the organism from which the ORF was derived or the organism that contains the most similar ORF at the amino acid level. Alternatively, in some embodiments, an increase in translation for a Cas9 sequence in a mammal, cell type, organ of a mammal, human, organ of a human, etc., is determined relative to translation of an ORF with the sequence of SEQ ID NO: 2 or 5 with all else equal, including any applicable point mutations, heterologous domains, and the like. In some embodiments, at least about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% of the codons in an ORF are codons corresponding to highly expressed tRNAs (e.g., the highest-expressed tRNA for each amino acid) in a mammal, such as a human. In some embodiments, at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% of the codons in an ORF are codons corresponding to highly expressed tRNAs (e.g., the highest-expressed tRNA for each amino acid) in a mammalian organ, such as a human organ.

[00239] Alternatively, codons corresponding to highly expressed tRNAs in an organism (e.g., human) in general may be used.

[00240] Any of the foregoing approaches to codon selection can be combined with the minimal uridine and/or adenine codons shown above, e.g., by starting with the codons of Table 1, 2, or 3, and then where more than one option is available, using the codon that corresponds to a more highly-expressed tRNA, either in the organism (e.g., human) in general, or in an organ or cell type of interest(e.g., human liver or human hepatocytes).

[00241] In some embodiments, at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% of the codons in an ORF are codons from a codon set shown in Table 4 (e.g., the low U 1, low A, or low A/U codon set). The codons in the low U 1, low G, low A, and low A/U sets use codons that minimize the indicated nucleotides while also using codons corresponding to highly expressed tRNAs where more than one option is available. In some embodiments, at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% of the codons in an ORF are codons from the low U 1 codon set shown in Table 4. In some embodiments, at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% of the codons in an ORF are codons from the low A codon set shown in Table 4. In some embodiments, at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% of the codons in an ORF are codons from the low A/U codon set shown in Table 4.

Table 4. Exemplary Codon Sets.

Amino Acid	Low U 1	Low U 2	Low A	Low A/U
Gly	GGC	GGG	GGC	GGC
Glu	GAG	GAA	GAG	GAG
Asp	GAC	GAC	GAC	GAC
Val	GTG	GTA	GTG	GTG
Ala	GCC	GCG	GCC	GCC
Arg	AGA	CGA	CGG	CGG
Ser	AGC	AGC	TCC	AGC
Lys	AAG	AAA	AAG	AAG
Asn	AAC	AAC	AAC	AAC
Met	ATG	ATG	ATG	ATG
Ile	ATC	ATA	ATC	ATC
Thr	ACC	ACG	ACC	ACC
Trp	TGG	TGG	TGG	TGG
Cys	TGC	TGC	TGC	TGC
Tyr	TAC	TAC	TAC	TAC
Leu	CTG	CTA	CTG	CTG
Phe	TTC	TTC	TTC	TTC
Gln	CAG	CAA	CAG	CAG
His	CAC	CAC	CAC	CAC

2. Heterologous functional domains; nuclear localization signals (NLS)

[00242] In some embodiments, the polypeptide comprising a cytidine deaminase (e.g., A3A) and an RNA-guided nickase further comprises one or more additional heterologous functional domains (e.g., is or comprises a ternary or higher-order fusion polypeptide).

[00243] In some embodiments, the heterologous functional domain may facilitate transport of the polypeptide into the nucleus of a cell. For example, the heterologous functional domain may be a nuclear localization signal (NLS). In some embodiments, the polypeptide may be fused with 1-10 NLS(s). In some embodiments, the polypeptide may be fused with 1-5 NLS(s). In some embodiments, the polypeptide may be fused with one NLS. Where one NLS is used, the NLS may be fused at the N-terminus or the C-terminus of the polypeptide sequence. In some embodiments, the polypeptide may be fused C-terminally to at least one NLS. An NLS may also be inserted within the polypeptide sequence. In other embodiments, the polypeptide may be fused with more than one NLS. In some embodiments, the polypeptide may be fused with 2, 3, 4, or 5 NLSs. In some embodiments, the polypeptide may be fused with two NLSs. In certain circumstances, the two NLSs may be the same (*e.g.*, two SV40 NLSs) or different. In some embodiments, the polypeptide is fused to two SV40 NLS sequences at the carboxy terminus. In some embodiments, the polypeptide may be fused with two NLSs, one at the N-terminus and one at the C-terminus. In some embodiments, the polypeptide may be fused with 3 NLSs. In some embodiments, the polypeptide may be fused with no NLS. In some embodiments, the NLS may be a monopartite sequence, such as, *e.g.*, the SV40 NLS, PKKKRKV (SEQ ID NO: 63) or PKKKRRV (SEQ ID NO: 121). In some embodiments, the NLS may be a bipartite sequence, such as the NLS of nucleoplasmin, KRPAATKKAGQAKKKK (SEQ ID NO: 122). In a specific embodiment, a single PKKKRKV (SEQ ID NO: 63) NLS may be fused at the C-terminus of the polypeptide. One or more linkers are optionally included at the fusion site (e.g., between the polypeptide and NLS). In some embodiments, one or more NLS(s) according to any of the foregoing embodiments are present in the polypeptide in combination with one or more additional heterologous functional domains, such as any of the heterologous functional domains described below.

[00244] In some embodiments of the mRNA disclosed herein, the cytidine deaminase (e.g., A3A) is located N-terminal to the RNA-guided nickase in the polypeptide.

In some embodiments of the mRNA disclosed herein, the encoded RNA-guided nickase comprises a nuclear localization signal (NLS). In some embodiments, the NLS is fused to the C-terminus of the RNA-guided nickase. In some embodiments, the NLS is fused to the C-terminus of the RNA-guided nickase via a linker. In some embodiments, the NLS is fused to the N-terminus of the RNA-guided nickase. In some embodiments, the NLS is fused to the N-terminus of the RNA-guided nickase via a linker (e.g., SEQ ID NO: 61). In some embodiments, the NLS comprises a sequence having at least 80%, 85%, 90%, or 95% identity to any one of SEQ ID NOs: 63 and 110-122. In some embodiments, the NLS comprises the sequence of any one of SEQ ID NOs: 63 and 110-122. In some embodiments, the NLS is encoded by a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of any one of SEQ ID NOs: 63 and 110-122.

[00245] In some embodiments, the heterologous functional domain may be capable of modifying the intracellular half-life of the A3A and/or the RNA-guided nickase in the polypeptide. In some embodiments, the half-life of the A3A and/or the RNA-guided nickase in the polypeptide may be increased. In some embodiments, the half-life of the A3A and/or the RNA-guided nickase in the polypeptide may be reduced. In some embodiments, the heterologous functional domain may be capable of increasing the stability of the A3A and/or the RNA-guided nickase in the polypeptide. In some embodiments, the heterologous functional domain may be capable of reducing the stability of the A3A and/or the RNA-guided nickase in the polypeptide . In some embodiments, the heterologous functional domain may act as a signal peptide for protein degradation. In some embodiments, the protein degradation may be mediated by proteolytic enzymes, such as, for example, proteasomes, lysosomal proteases, or calpain proteases. In some embodiments, the heterologous functional domain may comprise a PEST sequence. In some embodiments, the polypeptide may be modified by addition of ubiquitin or a polyubiquitin chain. In some embodiments, the ubiquitin may be a ubiquitin-like protein (UBL). Non-limiting examples of ubiquitin-like proteins include small ubiquitin-like modifier (SUMO), ubiquitin cross-reactive protein (UCRP, also known as interferon-stimulated gene-15 (ISG15)), ubiquitin-related modifier-1 (URM1), neuronal-precursor-cell-expressed developmentally downregulated protein-8 (NEDD8, also called Rub1 in *S. cerevisiae*), human leukocyte antigen F-associated (FAT10), autophagy-8 (ATG8) and -12 (ATG12), Fau ubiquitin-like protein (FUB1), membrane-anchored UBL (MUB), ubiquitin fold-modifier-1 (UFM1), and ubiquitin-like protein-5 (UBL5).

[00246] In some embodiments, the heterologous functional domain may be a marker domain. Non-limiting examples of marker domains include fluorescent proteins, purification tags, epitope tags, and reporter gene sequences. In some embodiments, the marker domain may be a fluorescent protein. Any known fluorescent proteins may be used as the marker domain such as GFP, YFP, EBFP, ECFP, DsRed or any other suitable fluorescent protein. In some embodiments, the marker domain may be a purification tag and/or an epitope tag. Non-limiting exemplary tags include glutathione-S-transferase (GST), chitin binding protein (CBP), maltose binding protein (MBP), thioredoxin (TRX), poly(NANP), tandem affinity purification (TAP) tag, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, 8xHis, biotin carboxyl carrier protein (BCCP), poly-His, and calmodulin. In some embodiments, the marker domain may be a reporter gene. Non-limiting exemplary reporter genes include glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT), beta-galactosidase, beta-glucuronidase, luciferase, or fluorescent proteins.

[00247] In additional embodiments, the heterologous functional domain may target the polypeptide to a specific organelle, cell type, tissue, or organ. In some embodiments, the heterologous functional domain may target the polypeptide to mitochondria.

3. UTRs; Kozak sequences

[00248] In some embodiments, the nucleic acid (e.g., mRNA) disclosed herein comprises a 5' UTR, 3' UTR, or 5' and 3' UTRs from Hydroxysteroid 17-Beta Dehydrogenase 4 (HSD17B4 or HSD) or globin such as human alpha globin (HBA), human beta globin (HBB), *Xenopus laevis* beta globin (XBG), bovine growth hormone, cytomegalovirus (CMV), mouse Hba-al, heat shock protein 90 (Hsp90), glyceraldehyde 3-phosphate dehydrogenase (GAPDH), beta-actin, alpha-tubulin, tumor protein (p53), or epidermal growth factor receptor (EGFR).

[00249] In some embodiments, the nucleic acid disclosed herein comprises a 5' UTR from HSD and a 3' UTR from a human albumin gene. In some embodiments, an mRNA disclosed herein comprises a 5' UTR with at least 90% identity to any one of SEQ ID NO: 93 and a 3' UTR with at least 90% identity to any one of SEQ ID NO: 69.

[00250] In some embodiments, the nucleic acid disclosed herein comprises a 5' UTR with at least 90% identity to any one of SEQ ID NOs: 91-98. In some embodiments, an

mRNA disclosed herein comprises a 3' UTR with at least 90% identity to any one of SEQ ID NOs: 69, 99-106. In some embodiments, any of the foregoing levels of identity is at least 95%, at least 98%, at least 99%, or 100%. In some embodiments, an mRNA disclosed herein comprises a 5' UTR having the sequence of any one of SEQ ID NOs: 91-98. In some embodiments, an mRNA disclosed herein comprises a 3' UTR having the sequence of any one of SEQ ID NOs: 69, 99-106. In some embodiments, the mRNA comprises a 5' UTR and a 3' UTR from the same source.

[00251] In some embodiments, the nucleic acid described herein does not comprise a 5' UTR, e.g., there are no additional nucleotides between the 5' cap and the start codon. In some embodiments, the mRNA comprises a Kozak sequence (described below) between the 5' cap and the start codon, but does not have any additional 5' UTR. In some embodiments, the mRNA does not comprise a 3' UTR, e.g., there are no additional nucleotides between the stop codon and the poly-A tail.

[00252] In some embodiments, the nucleic acid herein comprises a Kozak sequence. The Kozak sequence can affect translation initiation and the overall yield of a polypeptide translated from an mRNA. A Kozak sequence includes a methionine codon that can function as the start codon. A minimal Kozak sequence is NNNRUGN wherein at least one of the following is true: the first N is A or G and the second N is G. In the context of a nucleotide sequence, R means a purine (A or G). In some embodiments, the Kozak sequence is RNNRUGN, NNNRUGG, RNNRUGG, RNNAUGN, NNNAUGG, RNNAUGG, or GCCACCAUG. In some embodiments, the Kozak sequence is rccRUGg, rccAUGg, gccAccAUG, gccRccAUGG (SEQ ID NO: 107) or gccgccRccAUGG (SEQ ID NO: 108), with zero mismatches or with up to one or two mismatches to positions in lowercase.

4. Poly-A tail

[00253] In some embodiments, the nucleic acid disclosed herein further comprises a poly-adenylated (poly-A) tail. The poly-A tails may comprise at least 8 consecutive adenine nucleotides, but also comprise one or more non-adenine nucleotide. As used herein, "non-adenine nucleotides" refer to any natural or non-natural nucleotides that do not comprise adenine. Guanine, thymine, and cytosine nucleotides are exemplary non-adenine nucleotides. Thus, the poly-A tails on the nucleic acid described herein may comprise consecutive adenine nucleotides located 3' to nucleotides encoding a polypeptide of interest. In some instances, the poly-A tails on mRNA comprise non-consecutive adenine nucleotides located 3' to nucleotides encoding a polypeptide comprising a cytidine deaminase

(*e.g.*, A3A) and an RNA-guided nickase or a sequence of interest, wherein non-adenine nucleotides interrupt the adenine nucleotides at regular or irregularly spaced intervals.

[00254] In some embodiments, the poly-A tail is encoded in the plasmid used for in vitro transcription of mRNA and becomes part of the transcript. The poly-A sequence encoded in the plasmid, *i.e.*, the number of consecutive adenine nucleotides in the poly-A sequence, may not be exact, *e.g.*, a 100 poly-A sequence in the plasmid may not result in a precisely 100 poly-A sequence in the transcribed mRNA. In some embodiments, the poly-A tail is not encoded in the plasmid, and is added by PCR tailing or enzymatic tailing, e.g., using *E. coli* poly(A) polymerase.

[00255] In some embodiments, the one or more non-adenine nucleotides are positioned to interrupt the consecutive adenine nucleotides so that a poly(A) binding protein can bind to a stretch of consecutive adenine nucleotides. In some embodiments, one or more non-adenine nucleotide(s) is located after at least 8, 9, 10, 11, or 12 consecutive adenine nucleotides. In some embodiments, the one or more non-adenine nucleotide is located after 8-50 consecutive adenine nucleotides. In some embodiments, the one or more non-adenine nucleotide is located after 8-100 consecutive adenine nucleotides.

[00256] In some embodiments, the poly-A tail comprises or contains one non-adenine nucleotide or one consecutive stretch of 2-10 non-adenine nucleotides.

[00257] In some embodiments, the non-adenine nucleotide is guanine, cytosine, or thymine. In some instances, where more than one non-adenine nucleotide is present, the non-adenine nucleotide may be selected from: a) guanine and thymine nucleotides; b) guanine and cytosine nucleotides; c) thymine and cytosine nucleotides; or d) guanine, thymine and cytosine nucleotides. An exemplary poly-A tail comprising non-adenine nucleotides is provided as SEQ ID NO: 109.

5. Modified nucleotides

[00258] In some embodiments, the nucleic acid disclosed herein comprises a modified uridine at some or all uridine positions. In some embodiments, the modified uridine is a uridine modified at the 5 position, *e.g.*, with a halogen or C1-C3 alkoxy. In some embodiments, the modified uridine is a pseudouridine modified at the 1 position, e.g., with a C1-C3 alkyl. The modified uridine can be, for example, pseudouridine, N1-methyl-pseudouridine, 5-methoxyuridine, 5-iodouridine, or a combination thereof.

[00259] In some embodiments, at least 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 100% of the

uridine positions in the nucleic acid disclosed herein are modified uridines. In some embodiments, 10%-25%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-75%, 75-85%, 85-95%, or 90-100% of the uridine positions in an mRNA disclosed herein are modified uridines, e.g., 5-methoxyuridine, 5-iodouridine, N1-methyl pseudouridine, pseudouridine, or a combination thereof.

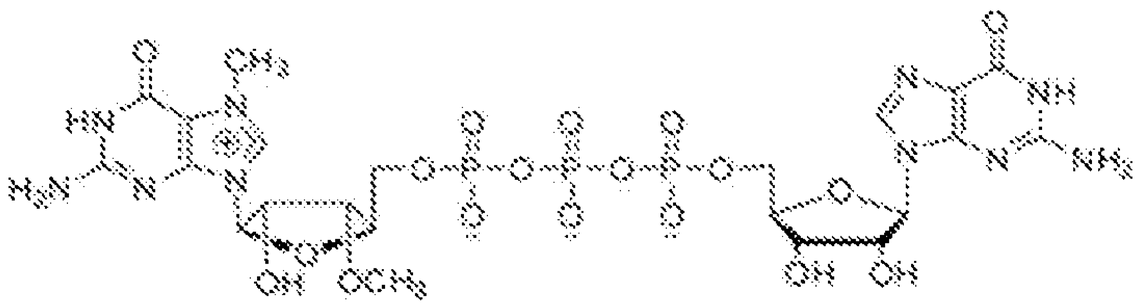
[00260] In some embodiments, at least 10% of the uridine is substituted with a modified uridine. In some embodiments, 15% to 45% of the uridine is substituted with the modified uridine. In some embodiments, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or 100%of the uridine is substituted with the modified uridine.

6. 5' Cap

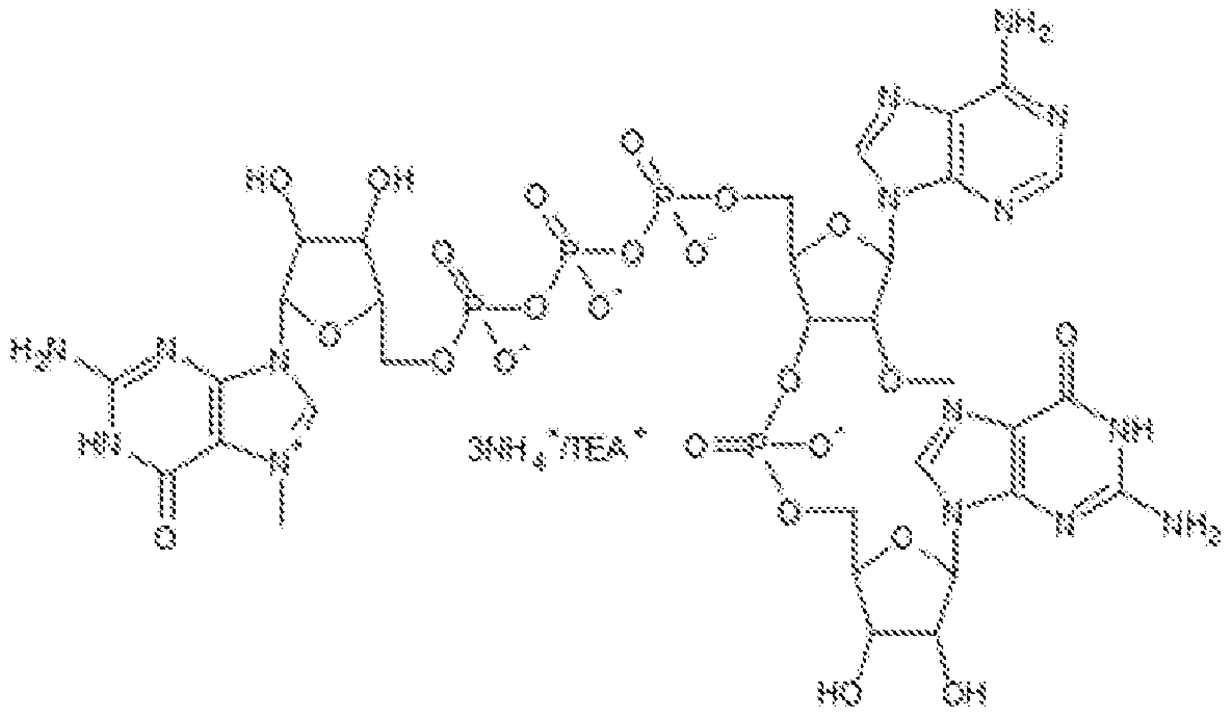
[00261] In some embodiments, the nucleic acid disclosed herein comprises a 5' cap, such as a Cap0, Cap1, or Cap2. A 5' cap is generally a 7-methylguanine ribonucleotide (which may be further modified, as discussed below *e.g.,* with respect to ARCA) linked through a 5'-triphosphate to the 5' position of the first nucleotide of the 5'-to-3' chain of the nucleic acid, i.e., the first cap-proximal nucleotide. In Cap0, the riboses of the first and second cap-proximal nucleotides of the mRNA both comprise a 2'-hydroxyl. In Cap1, the riboses of the first and second transcribed nucleotides of the mRNA comprise a 2'-methoxy and a 2'-hydroxyl, respectively. In Cap2, the riboses of the first and second cap-proximal nucleotides of the mRNA both comprise a 2'-methoxy. See, e.g., Katibah et al. (2014) *Proc Natl Acad Sci USA* 111(33):12025-30; Abbas et al. (2017) *Proc Natl Acad Sci USA* 114(11):E2106-E2115. Most endogenous higher eukaryotic nucleic acids, including mammalian nucleic acids such as human nucleic acids, comprise Cap1 or Cap2. Cap0 and other cap structures differing from Cap1 and Cap2 may be immunogenic in mammals, such as humans, due to recognition as "non-self" by components of the innate immune system such as IFIT-1 and IFIT-5, which can result in elevated cytokine levels including type I interferon. Components of the innate immune system such as IFIT-1 and IFIT-5 may also compete with eIF4E for binding of a nucleic acids with a cap other than Cap1 or Cap2, potentially inhibiting translation of the nucleic acid.

[00262] A cap can be included co-transcriptionally. For example, ARCA (anti-reverse cap analog; Thermo Fisher Scientific Cat. No. AM8045) is a cap analog comprising a 7-methylguanine 3'-methoxy-5'-triphosphate linked to the 5' position of a guanine ribonucleotide which can be incorporated in vitro into a transcript at initiation. ARCA results

in a Cap0 cap or a Cap0-like cap in which the 2' position of the first cap-proximal nucleotide is hydroxyl. See, e.g., Stepinski et al., (2001) "Synthesis and properties of mRNAs containing the novel 'anti-reverse' cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl(3'deoxy)GpppG," *RNA* 7: 1486–1495. The ARCA structure is shown below.



[00263] CleanCap™ AG (m7G(5')ppp(5')(2'OMeA)pG; TriLink Biotechnologies Cat. No. N-7113) or CleanCap™ GG (m7G(5')ppp(5')(2'OMeG)pG; TriLink Biotechnologies Cat. No. N-7133) can be used to provide a Cap1 structure co-transcriptionally. 3'-O-methylated versions of CleanCap™ AG and CleanCap™ GG are also available from TriLink Biotechnologies as Cat. Nos. N-7413 and N-7433, respectively. The CleanCap™ AG structure is shown below. CleanCap™ structures are sometimes referred to herein using the last three digits of the catalog numbers listed above (e.g., "CleanCap™ 113" for TriLink Biotechnologies Cat. No. N-7113).



[00264] Alternatively, a cap can be added to an RNA post-transcriptionally. For example, Vaccinia capping enzyme is commercially available (New England Biolabs Cat. No. M2080S) and has RNA triphosphatase and guanylyltransferase activities, provided by its D1 subunit, and guanine methyltransferase, provided by its D12 subunit. As such, it can add a 7-methylguanine to an RNA, so as to give Cap0, in the presence of S-adenosyl methionine and GTP. See, e.g., Guo, P. and Moss, B. (1990) *Proc. Natl. Acad. Sci. USA* 87, 4023-4027; Mao, X. and Shuman, S. (1994) *J. Biol. Chem.* 269, 24472-24479. For additional discussion of caps and capping approaches, see, e.g., WO2017/053297 and Ishikawa et al., *Nucl. Acids. Symp. Ser*. (2009) No. 53, 129-130.

G. Guide RNA (gRNA)

[00265] In some embodiments, the compositions comprise at least one guide RNA (gRNA), and the methods comprise delivering at least one gRNA, wherein the gRNA directs the editor to a desired genomic location. In some embodiments, a composition comprises an mRNA described herein and at least one gRNA. In some embodiments, a composition comprises a polypeptide described herein and at least one gRNA. In some embodiments, the gRNA is a single guide RNA (sgRNA). In some embodiments, the gRNA is a dual guide RNA (dgRNA).

[00266] A gRNA disclosed herein may comprise a guide sequence that directs a polypeptide comprising a cytidine deaminase (*e.g.*, an APOBEC3A deaminase) and an RNA-guided nickase to a cytosine (C) located in any region of a gene (*e.g.*, within the coding region of a gene) for cytosine (C) to thymine (T) conversion ("C-to-T conversion").

[00267] In some embodiments, the C-to-T conversion alters a DNA sequence, such as a human genetic sequence. In some embodiments, the C-to-T conversion alters the coding sequence of a gene. In some embodiments, the C-to-T conversion generates a stop codon, for example, a premature stop codon within the coding region of a gene. In some embodiments, the C-to-T conversion eliminates a stop codon. In some embodiments, the C-to-T conversion alters the regulatory sequence of a gene (*e.g.*, a gene promotor or gene repressor). In some embodiments, the C-to-T conversion alters the splicing of a gene. In some embodiments, the C-to-T conversion corrects a genetic defect associated with a disease or disorder.

[00268] In some embodiments, a guide RNA (gRNA) comprises a guide sequence that directs a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase) and an RNA-guided nickase to a splice donor or acceptor site in a gene. In some embodiments, the splice donor or acceptor is a splice donor site. In some embodiments, the splice donor or acceptor site is a splice acceptor site.

[00269] In some embodiments, a guide RNA (gRNA) comprises a guide sequence that directs a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase) and an RNA-guided nickase to an acceptor splice site boundary. In some embodiments, a guide RNA (gRNA) comprises a guide sequence that directs a polypeptide comprising a cytidine deaminase (e.g., A3A) and an RNA-guided nickase to a donor splice site boundary.

[00270] In some embodiments, a guide RNA (gRNA) comprises a guide sequence that directs a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase) and an RNA-guided nickase to make a single-strand cut in a gene at a cut site 3' of an acceptor splice site boundary or 5' of an acceptor splice site boundary. In this and the following discussion, 3' and 5' indicate directions in the sense of the strand being cut.

[00271] In some embodiments, a guide RNA (gRNA) disclosed herein comprises a guide sequence that directs a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase) and an RNA-guided nickase for making a single-strand cut in a gene at a cut site that is 3' of a donor splice site boundary or 5' of a donor splice site boundary.

[00272] A "splice site," as used herein, refers to the three nucleotides that make up an acceptor splice site or a donor splice site (defined below), or any other nucleotides known in the art that are part of a splice site. *See e.g.*, Burset et al., Nucleic Acids Research 28(21):4364-4375 (2000) (describing canonical and non-canonical splice sites in mammalian genomes). The three nucleotides that make up an "acceptor splice site" are two conserved residues (*e.g.*, AG in humans) at the 3' of an intron and a boundary nucleotide (*i.e.*, the first nucleotide of the exon 3' of the AG). The three nucleotides that make up a "donor splice site" are two conserved residues (*e.g.*, GT (gene) or GU (in RNA such as pre-mRNA) in human) at the 5' end of an intron and a boundary nucleotide (*i.e.*, the first nucleotide of the exon 5' of the GT).

[00273] In some embodiments, a composition comprising at least one gRNA is provided in combination with a nucleic acid *(e.g.*, an mRNA) disclosed herein. In some embodiments, one or more gRNA is provided as a separate molecule from the nucleic acid (*e.g.*, an mRNA) disclosed herein. In some embodiments, a gRNA is provided as a part, such as a part of a UTR, of the nucleic acid disclosed herein.

[00274] In some embodiments, a composition is provided comprising a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a gRNA. In some embodiments, a ribonucleoprotein complex (RNP) is provided, the RNP comprising a polypeptide comprising a cytidine deaminase and an RNA-guided nickase and a gRNA. In some embodiments, the polypeptide does not comprise a UGI.

[00275] The gRNA comprises a guide sequence targeting a particular gene or genetic sequence. In some embodiments, the gRNA is a Cas nickase guide. In some embodiments, the gRNA is a Class 2 Cas nickase guide. In further embodiments, the gRNA is a Cpf1 or Cas9 guide. In some embodiments, the gRNA is a Nme nickase guide. In some

embodiments, the Nme nickase is a Nme1, Nme2, or Nme3 nickase. In some embodiments, the gRNA comprises a guide sequence 5' of an RNA that forms two or more hairpin or stem-loop structures. CRISPR/Cas gRNA structures are known in the art and vary with their cognate Cas nuclease. In general, the gRNA used together with any particular Cas9 or Nme nickase described herein must function with that nickase. For example, when the polypeptide disclosed herein comprises a SpyCas9 nickase, the gRNA provided is a SpyCas9 guide RNA (as described herein). When the polypeptide disclosed herein comprises a NmeCas9 nickase, the guide RNA is a NmeCas9 guide RNA (as described herein).

[00276] In some embodiments, the gRNA comprises a guide sequence that direct an RNA-guided nickase (*e.g.*, Cas9 nickase), to a target DNA sequence in a target locus, such as a target gene. Targets and exemplary target sequences targeting each gene are exemplified herein and include, but are not limited to, targets and guide sequences disclosed in e.g., WO2017185054 (for trinucleotide repeats in transcription factor four (*TCF4*)); WO 2018119182 A1 (targeting *SERPINA1*); WO 2019/067872 (targeting transthyretin (*TTR*); WO 2020/028327 A1 (targeting hydroxyacid oxidase 1 (*HAO1*), the contents of each of which are hereby incorporated by reference in their entirety. One skilled in the art will be familiar with suitable guide sequences for targeting other genes or loci of interest.

[00277] The gRNA may comprise a crRNA comprising 17, 18, 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides of a guide sequence. The gRNA may further comprise a trRNA. In each composition and method embodiment described herein, the crRNA and trRNA may be associated as a single RNA (sgRNA), or may be on separate RNAs (dgRNA). In the context of sgRNAs, the crRNA and trRNA components may be covalently linked, e.g., via a phosphodiester bond or other covalent bond.

[00278] In each of the composition, use, and method embodiments described herein, the gRNA may comprise two RNA molecules as a "dual guide RNA" or "dgRNA". The dgRNA comprises a first RNA molecule comprising a crRNA comprising a guide sequence, and a second RNA molecule comprising a trRNA. The first and second RNA molecules may not be covalently linked, but may form a RNA duplex via the base pairing between portions of the crRNA and the trRNA.

[00279] In each of the composition, use, and method embodiments described herein, the gRNA may comprise a single RNA molecule as a "single guide RNA" or "sgRNA". The sgRNA may comprise a crRNA (or a portion thereof) comprising a guide sequence covalently linked to, e.g., a trRNA. The sgRNA may comprise 17, 18, 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides of a guide sequence. In some embodiments, the

crRNA and the trRNA are covalently linked via a linker. In some embodiments, the sgRNA forms a stem-loop structure via the base pairing between portions of the crRNA and the trRNA. In some embodiments, the crRNA and the trRNA are covalently linked via one or more bonds that are not a phosphodiester bond.

[00280] In some embodiments, the trRNA may comprise all or a portion of a trRNA sequence derived from a naturally-occurring CRISPR/Cas system. In some embodiments, the trRNA comprises a truncated or modified wild type trRNA. The length of the trRNA depends on the CRISPR/Cas system used. In some embodiments, the trRNA comprises or consists of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or more than 100 nucleotides. In some embodiments, the trRNA may comprise certain secondary structures, such as, for example, one or more hairpin or stem-loop structures, or one or more bulge structures.

[00281] The gRNAs provided herein can be useful for recognizing (e.g., hybridizing to) a target sequence in the gene. In some embodiments, the selection of the one or more gRNAs is determined based on target sequences within the gene. In some embodiments, a gRNA complementary or having complementarity to a target sequence within the target locus is used to direct a polypeptide comprising a cytidine deaminase (e.g., A3A) and an RNA-guided nickase to a particular location in the locus. The target locus may be recognized and nicked by a Cas nickase comprising a gRNA.

[00282] In some embodiments, the guide sequence is at least 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, or 90% identical to a target sequence. In some embodiments, the target sequence may be complementary to the guide sequence of the gRNA. In some embodiments, the degree of complementarity or identity between a guide sequence of a gRNA and its corresponding target sequence may be about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%. In some embodiments, the target sequence and the guide sequence of the gRNA may be 100% complementary or identical. In other embodiments, the target sequence and the guide sequence of the gRNA may contain at least one mismatch. For example, the target sequence and the guide sequence of the gRNA may contain 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mismatches, where the total length of the target sequence is at least about 17, 18, 19, 20 or more base pairs. In some embodiments, the target sequence and the guide sequence of the gRNA may contain 1, 2, 3, 4, 5, or 6 mismatches where the guide sequence is 20 nucleotides.

[00283] The gRNA may comprise a guide sequence linked to additional nucleotides to form a crRNA, e.g., with the following exemplary nucleotide sequence following the guide sequence at its 3' end: GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 139).

[00284] In the case of an sgRNA, the guide sequence may be linked to additional nucleotides to form a sgRNA, e.g., with the following exemplary nucleotide sequence following the 3' end of the guide sequence:
GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 140) in 5' to 3' orientation.

[00285] In some embodiments, the sgRNA comprises the modification pattern shown below in SEQ ID NO: 141, where N is any natural or non-natural nucleotide, and where the totality of the N's comprise a guide sequence as described herein and the modified sgRNA comprises the following sequence:
mN*mN*mN*NNNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU (SEQ ID NO: 141), where "N" may be any natural or non-natural nucleotide. For example, encompassed herein is SEQ ID NO: 141, where the N's are replaced with any of the guide sequences disclosed herein. The modifications remain as shown in SEQ ID NO: 141 despite the substitution of N's for the nucleotides of a guide. That is, although the nucleotides of the guide replace the "N's", the first three nucleotides are 2'OMe modified and there are phosphorothioate linkages between the first and second nucleotides, the second and third nucleotides and the third and fourth nucleotides.

[00286] Fig. 23A shows an exemplary sgRNA (SEQ ID NO: 141, methylation not shown) in a possible secondary structure with labels designating individual nucleotides of the conserved region of the sgRNA, including the lower stem, bulge, upper stem, nexus (the nucleotides of which can be referred to as N1 through N18, respectively, in the 5' to 3' direction), and the hairpin region which includes hairpin 1 and hairpin 2 regions. A nucleotide between hairpin 1 and hairpin 2 is labeled n. A guide region may be present on an sgRNA and is indicated in this figure as "(N)x" preceding the conserved region of the sgRNA. In some embodiments, the sgRNA may further comprise one or more nucleotides between the lower stem and bulge regions, between the bulge and the upper stem region, between the upper stem and the nexus, or the between the nexus and the hairpin 1 region or between the hairpin 1 and hairpin 2 regions.

[00287] In some embodiments, a conserved portion of the sgRNA is a conserved region of a spyCas9 or a spyCas9 equivalent. In some embodiments, a conserved portion of the sgRNA is not from *S. pyogenes* Cas9, such as *Staphylococcus aureus* Cas9 ("saCas9"). Further description of regions of exemplary sgRNAs are provided in WO2019/237069 published December 12, 2019, the entire contents of which are incorporated herein by reference.

[00288] The SpyCas9 gRNA may comprise internal linkers. In some embodiments, the internal linker may have a bridging length of about 3-30, optionally 12-21 atoms, and the linker substitutes for at least 2 nucleotides of the gRNA. In some embodiments, the internal linker has a bridging length of about 6-18 atoms, optionally about 6-12 atoms, and the linker substitutes for at least 2 nucleotides of the gRNA.
In some embodiments the internal linker comprises at least two ethylene glycol subunits covalently linked to each other. In some embodiments, the internal linker comprises a PEG-linker.

[00289] In some embodiments, the internal linker comprises a PEG-linker having from 1 to 10 ethylene glycol units. In some embodiments, the internal linker comprises a PEG-linker having from 3 to 6 ethylene glycol units. In some embodiments, the internal linker comprises a PEG-linker having 3 ethylene glycol units. In some embodiments, the internal linker comprises a PEG-linker having 6 ethylene glycol units.

[00290] In some embodiments, the conserved portion of a spyCas9 guide RNA comprises a repeat-anti-repeat region, a hairpin 1 region, and a hairpin 2 region, and further comprises at least one of:

1) a first internal linker substituting for at least 2 nucleotides of an upper stem region of the repeat-anti-repeat region of the sgRNA;
2) a second internal linker substituting for 1 or 2 nucleotides of the hairpin 1 of the sgRNA; or
3) a third internal linker substituting for at least 2 nucleotides of the hairpin 2 of the sgRNA.

[00291]

Exemplary locations of the linkers in the spyCas9 guide RNA are as shown in the following: NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUA**(L1)**UAGCAAGUUAAAAUAAG GC**UA**GUCCGUUAUCAACUU**(L1)**AAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO: 524) where N are nucleotides encoding a guide sequence.

[00292] As used herein, "Linker 1" or "L1" refers to an internal linker having a bridging length of about 15-21 atoms. As used herein, "Linker 2" or "L2" refers to an internal linker having a bridging length of about 6-12 atoms.

[00293] In some embodiments, the spyCas9 guide RNA comprising internal linkers may be chemically modified. Exemplary modifications include a modification pattern of the following sequence:

mA*mC*mG*CAAAUAUCAGUCCAGCGGUUUUAGAmGmCmUmA(L1)mUmAmGmC AAGUUAAAAUAAGGC(L2)GUCCGUUAUCAC(L1)GGGCACCGAGUCGG*mU*mG* mC (SEQ ID NO: 523).

[00294] In some embodiments, the gRNA comprises a 3' tail. In some embodiments, the 3' tail consists of a nucleotide comprising a uracil or modified uracil. In some embodiments, the 3' terminal nucleotide is a modified nucleotide. In some embodiments, wherein the 3' tail comprises a modification of any one or more of the nucleotides present in the 3' tail. In further embodiments, wherein the modification of the 3' tail is one or more of 2'-O-methyl (2'-OMe) modified nucleotide and a phosphorothioate (PS) linkage between nucleotides. penultimate nucleotide.

1. Short-single guide RNA (short-sgRNA)

[00295] In some embodiments, an sgRNA provided herein is a short-single guide RNAs (short-sgRNAs), *e.g.*, comprising a conserved portion of an sgRNA comprising a hairpin region, wherein the hairpin region lacks at least 5-10 nucleotides or 6-10 nucleotides. In some embodiments, the 5-10 nucleotides or 6-10 nucleotides are consecutive.

[00296] In some embodiments, a short-sgRNA lacks at least nucleotides 54-58 (AAAAA) of the conserved portion of a spyCas9 sgRNA. In some embodiments, a short-sgRNA is a non-spyCas9 sgRNA that lacks nucleotides corresponding to nucleotides 54-58 (AAAAA) of the conserved portion of a spyCas9 as determined, for example, by pairwise or structural alignment.

[00297] Structural alignment is useful where molecules share similar structures despite considerable sequence variation. Structural alignment involves identifying corresponding residues across two (or more) sequences by (i) modeling the structure of a first sequence using the known structure of the second sequence or (ii) comparing the structures of the first and second sequences where both are known, and identifying the residue in the first sequence most similarly positioned to a residue of interest in the second sequence. Corresponding residues are identified in some algorithms based on distance minimization given position

(e.g., nucleobase position 1 or the 1' carbon of the pentose ring for polynucleotides, or alpha carbons for polypeptides) in the overlaid structures (e.g., what set of paired positions provides a minimized root-mean-square deviation for the alignment). When identifying positions in a non-spyCas9 gRNA corresponding to positions described with respect to spyCas9 gRNA, spyCas9 gRNA can be the "second" sequence. Where a non-spyCas9 gRNA of interest does not have an available known structure, but is more closely related to another non-spyCas9 gRNA that does have a known structure, it may be most effective to model the non-spyCas9 gRNA of interest using the known structure of the closely related non-spyCas9 gRNA, and then compare that model to the spyCas9 gRNA structure to identify the desired corresponding residue in the non-spyCas9 gRNA of interest. There is an extensive literature on structural modeling and alignment for proteins; representative disclosures include US 6859736; US 8738343; and those cited in Aslam et al., Electronic Journal of Biotechnology 20 (2016) 9–13. For discussion of modeling a structure based on a known related structure or structures, *see, e.g.*, Bordoli et al., *Nature Protocols* 4 (2009) 1–13, and references cited therein. *See also* Figure 2(F) from Nishimasu et al., *Cell* 162(5): 1113-1126 (2015) for alignment of nucleic acid.

[00298] In some embodiments, the short-sgRNA described herein comprises a conserved portion comprising a hairpin region, wherein the hairpin region lacks 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides. In some embodiments, the lacking nucleotides are 5-10 lacking nucleotides or 6-10 lacking nucleotides. In some embodiments, the lacking nucleotides are consecutive. In some embodiments, the lacking nucleotides span at least a portion of hairpin 1 and a portion of hairpin 2. In some embodiments, the 5-10 lacking nucleotides comprise or consist of nucleotides 54-58, 54-61, or 53-60 of SEQ ID NO: 140.

[00299] In some embodiments, the short-sgRNA described herein further comprises a nexus region, wherein the nexus region lacks at least one nucleotide (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides in the nexus region). In some embodiments, the short-sgRNA lacks each nucleotide in the nexus region.

[00300] In some embodiments, the SpyCas9 short-sgRNA described herein comprises a sequence of NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGUGCU (SEQ ID NO: 521). In some embodiments, the short-sgRNA described herein comprises a modification pattern as shown in SEQ ID NO: 520:

mN*mN*mN*GUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGmUmGmC*mU (SEQ ID NO: 520), where A, C, G, U, and N are adenine, cytosine, guanine, uracil, and any ribonucleotide, respectively, unless otherwise indicated. An m is indicative of a 2'O-methyl modification, and an * is indicative of a phosphorothioate linkage between the nucleotides.

[00301] In some embodiments, a gRNA described herein is an *N. meningitidis Cas9* (NmeCas9) gRNA comprising a conserved portion comprising a repeat/anti-repeat region, a hairpin 1 region, and a hairpin 2 region, wherein one or more of the repeat/anti-repeat region, the hairpin 1 region, and the hairpin 2 region are shortened. Exemplary wild-type NmeCas9 guide RNA comprises a sequence of $(N)_{20-25}$ GUUGUAGCUCCCUUUCUCAUUUCGGAAACGAAAUGAGAACCGUUGCUACAAU AAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAAGCUUCUGCUU UAAGGGGCAUCGUUUA (SEQ ID NO: 512). $(N)_{20-25}$ as used herein represent 20-25, i.e., 20, 21, 22, 23, 24, or 25 consecutive N. A, C, G, and U represent nucleotides having adenine, cytosine, guanine, and uracil bases, respectively. In some embodiments, $(N)_{20-25}$ has 24 nucleotides in length. N is any natural or non-natural nucleotide, and where the totality of the N's comprises a guide sequence.

[00302] In some embodiments, the conserved portion of the NmeCas9 short-gRNA comprises:

(a) a shortened repeat/anti-repeat region, wherein the shortened repeat/anti-repeat region lacks 2-24 nucleotides, wherein

(i) one or more of nucleotides 37-48 and 53-64 is deleted and optionally one or more of nucleotides 37-64 is substituted relative to SEQ ID NO: 512; and

(ii) nucleotide 36 is linked to nucleotide 65 by at least 2 nucleotides; or

b) a shortened hairpin 1 region, wherein the shortened hairpin 1 lacks 2-10, optionally 2-8 nucleotides, wherein

(i) one or more of nucleotides 82-86 and 91-95 is deleted and optionally one or more of positions 82-96 is substituted relative to SEQ ID NO: 512 and

(ii) nucleotide 81 is linked to nucleotide 96 by at least 4 nucleotides; or

(c) a shortened hairpin 2 region, wherein the shortened hairpin 2 lacks 2-18, optionally 2-16 nucleotides, wherein

(i) one or more of nucleotides 113-121 and 126-134 is deleted and optionally one or more of nucleotides 113-134 is substituted relative to SEQ ID NO: 512; and

(ii) nucleotide 112 is linked to nucleotide 135 by at least 4 nucleotides;

wherein one or both nucleotides 144-145 are optionally deleted relative to SEQ ID NO: 512; and wherein at least 10 nucleotides are modified nucleotides.

[00303] In some embodiments, the NmeCas9 short-gRNA comprises one of the following sequences in 5' to 3' orientation:

$(N)_{20-25}$ GUUGUAGCUCCCUGAAACCGUUGCUACAAUAAGGCCGUCGAAAGAUGU GCCGCAACGCUCUGCCUUCUGGCAUCGUU (SEQ ID NO: 513);

$(N)_{20-25}$ GUUGUAGCUCCCUGAAACCGUUGCUACAAUAAGGCCGUCGAAAGAUGU GCCGCAACGCUCUGCCUUCUGGCAUCGUUUAUU (SEQ ID NO: 514);

$(N)_{20-25}$ GUUGUAGCUCCCUGGAAACCCGUUGCUACAAUAAGGCCGUCGAAAGA UGUGCCGCAACGCUCUGCCUUCUGGCAUCGUUUAUU (SEQ ID NO: 515).

[00304] In some embodiments, at least 10 nucleotides of the conserved portion of the NmeCas9 short-sgRNA are modified nucleotides.

[00305] In some embodiments, the NmeCas9 short-sgRNA comprises a conserved region comprising one of the following sequences in 5' to 3' orientation:

GUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGm CCmGmUmCmGmAmAmAmGmAmUGUGCmCGCmAmAmCmGCUCUmGmCCmUmU mCmUGmGCmAmUC*mG*mU*mU (SEQ ID NO: 516); or

GUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGm CCmGmUmCmGmAmAmAmGmAmUGUGCmCGmCAAmCGCUCUmGmCCmUmUmC mUGGCAUCG*mU*mU (SEQ ID NO: 517).

[00306] The shortened NmeCas9 gRNA may comprise internal linkers disclosed herein.

[00307] "Internal linker" as used herein describes a non-nucleotide segment joining two nucleotides within a guide RNA. If the gRNA contains a spacer region, the internal linker is located outside of the spacer region (e.g., in the scaffold or conserved region of the gRNA). For Type V guides, it is understood that the last hairpin is the only hairpin in the structure, i.e., the repeat-anti-repeat region. In some embodiments, the internal linker comprises a PEG-linker disclosed herein.

[00308] Exemplary locations of the linkers are as shown in the following:

$(N)_{20-25}$ GUUGUAGCUCCCUUC**(L1)**GACCGUUGCUACAAUAAGGCCGUC**(L1)**GAUGU GCCGCAACGCUCUGCC**(L1)**GGCAUCGUU (SEQ ID NO: 518). As used herein, (L1) refers to an internal linker having a bridging length of about 15-21 atoms.

[00309] In some embodiments, the shortened NmeCas9 guide RNA comprising internal linkers may be chemically modified. Exemplary modifications include a modification pattern of the following sequence: mN*mN*mN*mNmNNNmNmNNmNNmNNNNNmNNNNmNNNmGUUGmUmAmGmC UCCCmUmUmC(L1)mGmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmC(L1) mGmAmUGUGCmCGmCAAmCGCUCUmGmCC(L1)GGCAUCG*mU*mU (SEQ ID NO: 519).

2. Modifications

[00310] In some embodiments, the gRNA (*e.g.*, sgRNA, short-sgRNA, dgRNA, or crRNA) is modified. The term "modified" or "modification" in the context of a gRNA described herein includes, the modifications described above, including, for example, (a) end modifications, e.g., 5′ end modifications or 3′ end modifications, including 5' or 3' protective end modifications, (b) nucleobase (or "base") modifications, including replacement or removal of bases, (c) sugar modifications, including modifications at the 2′, 3′, and/or 4′ positions, (d) internucleoside linkage modifications, and (e) backbone modifications, which can include modification or replacement of the phosphodiester linkages and/or the ribose sugar. A modification of a nucleotide at a given position includes a modification or replacement of the phosphodiester linkage immediately 3' of the sugar of the nucleotide. Thus, for example, a nucleic acid comprising a phosphorothioate between the first and second sugars from the 5' end is considered to comprise a modification at position 1. The term "modified gRNA" generally refers to a gRNA having a modification to the chemical structure of one or more of the base, the sugar, and the phosphodiester linkage or backbone portions, including nucleotide phosphates, all as detailed and exemplified herein (see the modification patterns shown in e.g., SEQ ID NOs: 142-145, 181-185 and 191-203).

[00311] Further description and exemplary patterns of modifications are provided in in Table 1 of WO2019/237069 published December 12, 2019, the entire contents of which are incorporated herein by reference.

[00312] In some embodiments, a gRNA comprises modifications at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or more YA sites. In some embodiments, the pyrimidine of the YA site comprises a modification (which includes a modification altering the internucleoside linkage immediately 3' of the sugar of the pyrimidine). In some embodiments, the adenine of the YA site comprises a modification (which includes a

modification altering the internucleoside linkage immediately 3' of the sugar of the adenine). In some embodiments, the pyrimidine and the adenine of the YA site comprise modifications, such as sugar, base, or internucleoside linkage modifications. The YA modifications can be any of the types of modifications set forth herein. In some embodiments, the YA modifications comprise one or more of phosphorothioate, 2'-OMe, or 2'-fluoro. In some embodiments, the YA modifications comprise pyrimidine modifications comprising one or more of phosphorothioate, 2'-OMe, 2'-H, inosine, or 2'-fluoro. In some embodiments, the YA modification comprises a bicyclic ribose analog (e.g., an LNA, BNA, or ENA) within an RNA duplex region that contains one or more YA sites. In some embodiments, the YA modification comprises a bicyclic ribose analog (e.g., an LNA, BNA, or ENA) within an RNA duplex region that contains a YA site, wherein the YA modification is distal to the YA site.

[00313] In some embodiments, the guide sequence (or guide region) of a gRNA comprises 1, 2, 3, 4, 5, or more YA sites ("guide region YA sites") that may comprise YA modifications. In some embodiments, one or more YA sites located at 5-end, 6-end, 7-end, 8-end, 9-end, or 10-end from the 5' end of the 5' terminus (where "5-end", etc., refers to position 5 to the 3' end of the guide region, i.e., the most 3' nucleotide in the guide region) comprise YA modifications. A modified guide region YA site comprises a YA modification.

[00314] In some embodiments, a modified guide region YA site is within 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, or 9 nucleotides of the 3' terminal nucleotide of the guide region. For example, if a modified guide region YA site is within 10 nucleotides of the 3' terminal nucleotide of the guide region and the guide region is 20 nucleotides long, then the modified nucleotide of the modified guide region YA site is located at any of positions 11-20. In some embodiments, a modified guide region YA site is at or after nucleotide 4, 5, 6, 7, 8, 9, 10, or 11 from the 5' end of the 5' terminus.

[00315] In some embodiments, a modified guide region YA site is other than a 5' end modification. For example, a sgRNA can comprise a 5' end modification as described herein and further comprise a modified guide region YA site. Alternatively, a sgRNA can comprise an unmodified 5' end and a modified guide region YA site. Alternatively, a short-sgRNA can comprise a modified 5' end and an unmodified guide region YA site.

[00316] In some embodiments, a modified guide region YA site comprises a modification that at least one nucleotide located 5' of the guide region YA site does not comprise. For example, if nucleotides 1-3 comprise phosphorothioates, nucleotide 4 comprises only a 2'-OMe modification, and nucleotide 5 is the pyrimidine of a YA site and

comprises a phosphorothioate, then the modified guide region YA site comprises a modification (phosphorothioate) that at least one nucleotide located 5' of the guide region YA site (nucleotide 4) does not comprise. In another example, if nucleotides 1-3 comprise phosphorothioates, and nucleotide 4 is the pyrimidine of a YA site and comprises a 2'-OMe, then the modified guide region YA site comprises a modification (2'-OMe) that at least one nucleotide located 5' of the guide region YA site (any of nucleotides 1-3) does not comprise. This condition is also always satisfied if an unmodified nucleotide is located 5' of the modified guide region YA site.

[00317] In some embodiments, the modified guide region YA sites comprise modifications as described for YA sites above. The guide region of a gRNA may be modified according to any embodiment comprising a modified guide region set forth herein.

[00318] Conserved region YA sites 1-10 are illustrated in Fig. 23B. In some embodiments, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 conserved region YA sites comprise modifications. In some embodiments, conserved region YA sites 1, 8, or 1 and 8 comprise YA modifications. In some embodiments, conserved region YA sites 1, 2, 3, 4, and 10 comprise YA modifications. In some embodiments, YA sites 2, 3, 4, 8, and 10 comprise YA modifications. In some embodiments, conserved region YA sites 1, 2, 3, and 10 comprise YA modifications. In some embodiments, YA sites 2, 3, 8, and 10 comprise YA modifications. In some embodiments, YA sites 1, 2, 3, 4, 8, and 10 comprise YA modifications. In some embodiments, 1, 2, 3, 4, 5, 6, 7, or 8 additional conserved region YA sites comprise YA modifications.

[00319] In some embodiments, the modified conserved region YA sites comprise modifications as described for YA sites above. Any embodiments set forth elsewhere in this disclosure may be combined to the extent feasible with any of the foregoing embodiments.

[00320] In some embodiments, the 5' and/or 3' terminus regions of a gRNA are modified.

[00321] In some embodiments, the terminal (i.e., last) 1, 2, 3, 4, 5, 6, or 7 nucleotides in the 3' terminus region are modified. Throughout, this modification may be referred to as a "3' end modification". In some embodiments, the terminal (i.e., last) 1, 2, 3, 4, 5, 6, or 7 nucleotides in the 3' terminus region comprise more than one modification. In some embodiments, the 3' end modification comprises or further comprises any one or more of the following: a modified nucleotide selected from 2'-O-methyl (2'-O-Me) modified nucleotide, 2'-O-(2-methoxyethyl) (2'-O-moe) modified nucleotide, a 2'-fluoro (2'-F)

modified nucleotide, a phosphorothioate (PS) linkage between nucleotides, an inverted abasic modified nucleotide, or combinations thereof. In some embodiments, the 3' end modification comprises or further comprises modifications of 1, 2, 3, 4, 5, 6, or 7 nucleotides at the 3' end of the gRNA. In some embodiments, the 3' end modification comprises or further comprises one PS linkage, wherein the linkage is between the last and second to last nucleotide. In some embodiments, the 3' end modification comprises or further comprises two PS linkages between the last three nucleotides. In some embodiments, the 3' end modification comprises or further comprises four PS linkages between the last four nucleotides. In some embodiments, the 3' end modification comprises or further comprises PS linkages between any one or more of the last 2, 3, 4, 5, 6, or 7 nucleotides. In some embodiments, the gRNA comprising a 3' end modification comprises or further comprises a 3' tail, wherein the 3' tail comprises a modification of any one or more of the nucleotides present in the 3' tail. In some embodiments, the 3' tail is fully modified. In some embodiments, the 3' tail comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, or 1-10 nucleotides, optionally where any one or more of these nucleotides are modified. In some embodiments, a gRNA is provided comprising a 3' end modification, wherein the 3' end modification comprises the 3' end modification as shown in any one of SEQ ID Nos: 141-145. In some embodiments, a gRNA is provided comprising a 3' protective end modification. In some embodiments, the 3' tail comprises between 1 and about 20 nucleotides, between 1 and about 15 nucleotides, between 1 and about 10 nucleotides, between 1 and about 5 nucleotides, between 1 and about 4 nucleotides, between 1 and about 3 nucleotides, and between 1 and about 2 nucleotides. In some embodiments, the gRNA does not comprise a 3' tail.

[00322] In some embodiments, the 5' terminus region is modified, for example, the first 1, 2, 3, 4, 5, 6, or 7 nucleotides of the gRNA are modified. Throughout, this modification may be referred to as a "5' end modification". In some embodiments, the first 1, 2, 3, 4, 5, 6, or 7 nucleotides of the 5' terminus region comprise more than one modification. In some embodiments, at least one of the terminal (i.e., first) 1, 2, 3, 4, 5, 6, or 7 nucleotides at the 5' end are modified. In some embodiments, both the 5' and 3' terminus regions (e.g., ends) of the gRNA are modified. In some embodiments, only the 5' terminus region of the gRNA is modified. In some embodiments, only the 3' terminus region (plus or minus a 3' tail) of the conserved portion of a gRNA is modified. In some embodiments, the gRNA comprises modifications at 1, 2, 3, 4, 5, 6, or 7 of the first 7 nucleotides at a 5' terminus region of the gRNA. In some embodiments, the gRNA comprises modifications at 1, 2, 3, 4, 5, 6, or 7 of the 7 terminal nucleotides at a 3' terminus region. In some embodiments, 2, 3, or

4 of the first 4 nucleotides at the 5' terminus region, and/or 2, 3, or 4 of the terminal 4 nucleotides at the 3' terminus region are modified. In some embodiments, 2, 3, or 4 of the first 4 nucleotides at the 5' terminus region are linked with phosphorothioate (PS) bonds. In some embodiments, the modification to the 5' terminus and/or 3' terminus comprises a 2'-O-methyl (2'-O-Me) or 2'-O-(2-methoxyethyl) (2'-O-moe) modification. In some embodiments, the modification comprises a 2'-fluoro (2'-F) modification to a nucleotide. In some embodiments, the modification comprises a phosphorothioate (PS) linkage between nucleotides. In some embodiments, the modification comprises an inverted abasic nucleotide. In some embodiments, the modification comprises a protective end modification. In some embodiments, the modification comprises a more than one modification selected from protective end modification, 2'-O-Me, 2'-O-moe, 2'-fluoro (2'-F), a phosphorothioate (PS) linkage between nucleotides, and an inverted abasic nucleotide. In some embodiments, an equivalent modification is encompassed. In some embodiments, a gRNA is provided comprising a 5' end modification, wherein the 5' end modification comprises a 5' end modification as shown in any one of SEQ ID Nos: 141-145.

[00323] In some embodiments, a gRNA is provided comprising a 5' end modification and a 3' end modification. In some embodiments, the gRNA comprises modified nucleotides that are not at the 5' or 3' ends.

[00324] In some embodiments, a sgRNA is provided comprising an upper stem modification, wherein the upper stem modification comprises a modification to any one or more of US1-US12 in the upper stem region. In some embodiments, a sgRNA is provided comprising an upper stem modification, wherein the upper stem modification comprises a modification of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all 12 nucleotides in the upper stem region. In some embodiments, an sgRNA is provided comprising an upper stem modification, wherein the upper stem modification comprises 1, 2, 3, 4, or 5 YA modifications in a YA site. In some embodiments, the upper stem modification comprises a 2'-OMe modified nucleotide, a 2'-O-moe modified nucleotide, a 2'-F modified nucleotide, and/or combinations thereof. Other modifications described herein, such as a 5' end modification and/or a 3' end modification may be combined with an upper stem modification.

[00325] In some embodiments, the sgRNA comprises a modification in the hairpin region. In some embodiments, the hairpin region modification comprises at least one modified nucleotide selected from a 2'-O-methyl (2'-OMe) modified nucleotide, a 2'-fluoro (2'-F) modified nucleotide, and/or combinations thereof. In some embodiments, the hairpin region modification is in the hairpin 1 region. In some embodiments, the hairpin region

modification is in the hairpin 2 region. In some embodiments, the hairpin modification comprises 1, 2, or 3 YA modifications in a YA site. In some embodiments, the hairpin modification comprises at least 1, 2, 3, 4, 5, or 6 YA modifications. Other modifications described herein, such as an upper stem modification, a 5' end modification, and/or a 3' end modification may be combined with a modification in the hairpin region.

[00326] In some embodiments, a gRNA comprises a substituted and optionally shortened hairpin 1 region, wherein at least one of the following pairs of nucleotides are substituted in the substituted and optionally shortened hairpin 1 with Watson-Crick pairing nucleotides: H1-1 and H1-12, H1-2 and H1-11, H1-3 and H1-10, and/or H1-4 and H1-9. "Watson-Crick pairing nucleotides" include any pair capable of forming a Watson-Crick base pair, including A-T, A-U, T-A, U-A, C-G, and G-C pairs, and pairs including modified versions of any of the foregoing nucleotides that have the same base pairing preference. In some embodiments, the hairpin 1 region lacks any one or two of H1-5 through H1-8. In some embodiments, the hairpin 1 region lacks one, two, or three of the following pairs of nucleotides: H1-1 and H1-12, H1-2 and H1-11, H1-3 and H1-10 and/or H1-4 and H1-9. In some embodiments, the hairpin 1 region lacks 1-8 nucleotides of the hairpin 1 region. In any of the foregoing embodiments, the lacking nucleotides may be such that the one or more nucleotide pairs substituted with Watson-Crick pairing nucleotides (H1-1 and H1-12, H1-2 and H1-11, H1-3 and H1-10, and/or H1-4 and H1-9) form a base pair in the gRNA.

[00327] In some embodiments, the gRNA further comprises an upper stem region lacking at least 1 nucleotide, e.g., any of the shortened upper stem regions indicated in Table 7 of U.S. Application No. 62/946,905, the contents of which are hereby incorporated by reference in its entirety, or described elsewhere herein, which may be combined with any of the shortened or substituted hairpin 1 regions described herein.

[00328] In some embodiments, the gRNA described herein further comprises a nexus region, wherein the nexus region lacks at least one nucleotide.

3. Chemical Modifications of gRNAs

[00329] In some embodiments, the gRNA is chemically modified. A gRNA comprising one or more modified nucleosides or nucleotides is called a "modified" gRNA or "chemically modified" gRNA, to describe the presence of one or more non-naturally and/or naturally occurring components or configurations that are used instead of or in addition to the canonical A, G, C, and U residues. Modified nucleosides and nucleotides can include one or more of: (i) alteration, *e.g.*, replacement, of one or both of the non-linking phosphate oxygens

and/or of one or more of the linking phosphate oxygens in the phosphodiester backbone linkage (an exemplary backbone modification); (ii) alteration, *e.g.*, replacement, of a constituent of the ribose sugar, *e.g.*, of the 2' hydroxyl on the ribose sugar (an exemplary sugar modification); (iii) wholesale replacement of the phosphate moiety with "dephospho" linkers (an exemplary backbone modification); (iv) modification or replacement of a naturally occurring nucleobase, including with a non-canonical nucleobase (an exemplary base modification); (v) replacement or modification of the ribose-phosphate backbone (an exemplary backbone modification); (vi) modification of the 3' end or 5' end of the oligonucleotide, *e.g.*, removal, modification or replacement of a terminal phosphate group or conjugation of a moiety, cap or linker (such 3' or 5' cap modifications may comprise a sugar and/or backbone modification); and (vii) modification or replacement of the sugar (an exemplary sugar modification).

[00330] Chemical modifications such as those listed above can be combined to provide modified gRNAs comprising nucleosides and nucleotides (collectively "residues") that can have two, three, four, or more modifications. For example, a modified residue can have a modified sugar and a modified nucleobase. In some embodiments, every base of a gRNA is modified, *e.g.*, all bases have a modified phosphate group, such as a phosphorothioate group. In certain embodiments, all, or substantially all, of the phosphate groups of an gRNA molecule are replaced with phosphorothioate groups. In some embodiments, modified gRNAs comprise at least one modified residue at or near the 5' end of the RNA. In some embodiments, modified gRNAs comprise at least one modified residue at or near the 3' end of the RNA.

[00331] In some embodiments, the gRNA comprises one, two, three or more modified residues. In some embodiments, at least 5% (*e.g.*, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or 100%) of the positions in a modified gRNA are modified nucleosides or nucleotides.

[00332] In some embodiments of a backbone modification, the phosphate group of a modified residue can be modified by replacing one or more of the oxygens with a different substituent. Further, the modified residue, *e.g.*, modified residue present in a modified nucleic acid, can include the wholesale replacement of an unmodified phosphate moiety with a modified phosphate group as described herein. In some embodiments, the

backbone modification of the phosphate backbone can include alterations that result in either an uncharged linker or a charged linker with unsymmetrical charge distribution.

[00333] Examples of modified phosphate groups include phosphorothioate, phosphoroselenates, borano phosphates, borano phosphate esters, hydrogen phosphonates, phosphoroamidates, alkyl or aryl phosphonates and phosphotriesters.

[00334] Scaffolds that can mimic nucleic acids can also be constructed wherein the phosphate linker and ribose sugar are replaced by nuclease resistant nucleoside or nucleotide surrogates. Such modifications may comprise backbone and sugar modifications. In some embodiments, the nucleobases can be tethered by a surrogate backbone. Examples can include, without limitation, the morpholino, cyclobutyl, pyrrolidine and peptide nucleic acid (PNA) nucleoside surrogates.

[00335] The modified nucleosides and modified nucleotides can include one or more modifications to the sugar group, *i.e.* at sugar modification. For example, the 2' hydroxyl group (OH) can be modified, *e.g.* replaced with a number of different "oxy" or "deoxy" substituents. In some embodiments, modifications to the 2' hydroxyl group can enhance the stability of the nucleic acid since the hydroxyl can no longer be deprotonated to form a 2'-alkoxide ion. Examples of 2' hydroxyl group modifications can include alkoxy or aryloxy (OR, wherein "R" can be, *e.g.*, alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or a sugar); polyethyleneglycols (PEG), $O(CH_2CH_2O)_nCH_2CH_2OR$ wherein R can be, *e.g.*, H or optionally substituted alkyl, and n can be an integer from 0 to 20. In some embodiments, the 2' hydroxyl group modification can be 2'-O-Me. In some embodiments, the 2' hydroxyl group modification can be a 2'-fluoro modification, which replaces the 2' hydroxyl group with a fluoride. In some embodiments, the 2' hydroxyl group modification can include "locked" nucleic acids (LNA) in which the 2' hydroxyl can be connected, *e.g.*, by a C_{1-6} alkylene or C_{1-6} heteroalkylene bridge, to the 4' carbon of the same ribose sugar, where exemplary bridges can include methylene, propylene, ether, or amino bridges. In some embodiments, the 2' hydroxyl group modification can included "unlocked" nucleic acids (UNA) in which the ribose ring lacks the C2'-C3' bond. In some embodiments, the 2' hydroxyl group modification can include the methoxyethyl group (MOE), ($OCH_2CH_2OCH_3$, *e.g.*, a PEG derivative).

[00336] "Deoxy" 2' modifications can include hydrogen (*i.e.* deoxyribose sugars, *e.g.*, at the overhang portions of partially dsRNA); halo (*e.g.*, bromo, chloro, fluoro, or iodo); amino (wherein amino can be, *e.g.*, NH_2; alkylamino, dialkylamino, heterocyclyl, arylamino, diarylamino, heteroarylamino, diheteroarylamino, or amino acid);

$NH(CH_2CH_2NH)_nCH2CH_2$- amino (wherein amino can be, *e.g.*, as described herein), -NHC(O)R (wherein R can be, *e.g.*, alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar), cyano; mercapto; alkyl-thio-alkyl; thioalkoxy; and alkyl, cycloalkyl, aryl, alkenyl and alkynyl, which may be optionally substituted with *e.g.*, an amino as described herein.

[00337] The sugar modification can comprise a sugar group which may also contain one or more carbons that possess the opposite stereochemical configuration than that of the corresponding carbon in ribose. Thus, a modified nucleic acid can include nucleotides containing *e.g.*, arabinose, as the sugar. The modified nucleic acids can also include abasic sugars. These abasic sugars can also be further modified at one or more of the constituent sugar atoms. The modified nucleic acids can also include one or more sugars that are in the L form, *e.g.* L- nucleosides.

[00338] The modified nucleosides and modified nucleotides described herein, which can be incorporated into a modified nucleic acid, can include a modified base, also called a nucleobase. Examples of nucleobases include, but are not limited to, adenine (A), guanine (G), cytosine (C), and uracil (U). These nucleobases can be modified or wholly replaced to provide modified residues that can be incorporated into modified nucleic acids. The nucleobase of the nucleotide can be independently selected from a purine, a pyrimidine, a purine analog, or pyrimidine analog. In some embodiments, the nucleobase can include, for example, naturally-occurring and synthetic derivatives of a base.

[00339] In embodiments employing a dual guide RNA, each of the crRNA and the tracr RNA can contain modifications. Such modifications may be at one or both ends of the crRNA and/or tracr RNA. In embodiments comprising an sgRNA, one or more residues at one or both ends of the sgRNA may be chemically modified, or the entire sgRNA may be chemically modified. Certain embodiments comprise a 5' end modification. Certain embodiments comprise a 3' end modification. In certain embodiments, one or more or all of the nucleotides in single stranded overhang of a gRNA molecule are deoxynucleotides.

[00340] In some embodiments, the gRNAs disclosed herein comprise one of the modification patterns disclosed in WO2018/107028 A1, published June 14, 2018 the contents of which are hereby incorporated by reference in their entirety.

[00341] The terms “mA,” “mC,” “mU,” or “mG” may be used to denote a nucleotide that has been modified with 2’-O-Me. The terms “fA,” “fC,” “fU,” or “fG” may be used to denote a nucleotide that has been substituted with 2’-F. A “*” may be used to depict a PS modification. The terms A*, C*, U*, or G* may be used to denote a nucleotide that is linked to the next (e.g., 3’) nucleotide with a PS bond. The terms “mA*,” “mC*,” “mU*,” or

"mG*" may be used to denote a nucleotide that has been substituted with 2'-O-Me and that is linked to the next (e.g., 3') nucleotide with a PS bond.

H. Lipids; formulation; delivery

[00342] Disclosed herein are various embodiments using lipid nucleic acid assembly compositions comprising nucleic acids(s), or composition(s) described herein. In some embodiments, the lipid nucleic acid assembly composition comprises a nucleic acid (e.g., mRNA) comprising an open reading frame encoding a first polypeptide comprising a cytidine deaminase (e.g., A3A) and an RNA-guided nickase. In some embodiments, the lipid nucleic acid assembly composition comprises a first nucleic acid comprising an open reading frame encoding a first polypeptide comprising a cytidine deaminase (e.g., A3A) and an RNA-guided nickase and a second nucleic acid encoding a UGI.

[00343] As used herein, a "lipid nucleic acid assembly composition" refers to lipid-based delivery compositions, including lipid nanoparticles (LNPs) and lipoplexes. LNP refers to lipid nanoparticles <100nM. LNPs are formed by precise mixing a lipid component (*e.g.*, in ethanol) with an aqueous nucleic acid component and LNPs are uniform in size. Lipoplexes are particles formed by bulk mixing the lipid and nucleic acid components and are between about 100nm and 1 micron in size. In certain embodiments the lipid nucleic acid assemblies are LNPs. As used herein, a "lipid nucleic acid assembly" comprises a plurality of (*i.e.* more than one) lipid molecules physically associated with each other by intermolecular forces. A lipid nucleic acid assembly may comprise a bioavailable lipid having a pKa value of <7.5 or <7. The lipid nucleic acid assemblies are formed by mixing an aqueous nucleic acid-containing solution with an organic solvent-based lipid solution, *e.g.*, 100% ethanol. Suitable solutions or solvents include or may contain: water, PBS, Tris buffer, NaCl, citrate buffer, ethanol, chloroform, diethylether, cyclohexane, tetrahydrofuran, methanol, isopropanol. A pharmaceutically acceptable buffer may optionally be comprised in a pharmaceutical formulation comprising the lipid nucleic acid assemblies, *e.g.*, for an *ex vivo* therapy. In some embodiments, the aqueous solution comprises an RNA, such as an mRNA or a gRNA. In some embodiments, the aqueous solution comprises an mRNA encoding an RNA-guided DNA binding agent, such as Cas9.

[00344] As used herein, lipid nanoparticle (LNP) refers to a particle that comprises a plurality of (i.e., more than one) lipid molecules physically associated with each other by intermolecular forces. The LNPs may be, e.g., microspheres (including unilamellar and multilamellar vesicles, e.g., "liposomes"—lamellar phase lipid bilayers that, in some

embodiments, are substantially spherical—and, in more particular embodiments, can comprise an aqueous core, e.g., comprising a substantial portion of RNA molecules), a dispersed phase in an emulsion, micelles, or an internal phase in a suspension. Emulsions, micelles, and suspensions may be suitable compositions for local and/or topical delivery. See also, e.g., WO2017173054A1, the contents of which are hereby incorporated by reference in their entirety. Any LNP known to those of skill in the art to be capable of delivering nucleotides to subjects may be utilized with the guide RNAs and the nucleic acid encoding an RNA-guided nickase and the nucleic acid encoding a cytidine deaminase described herein.

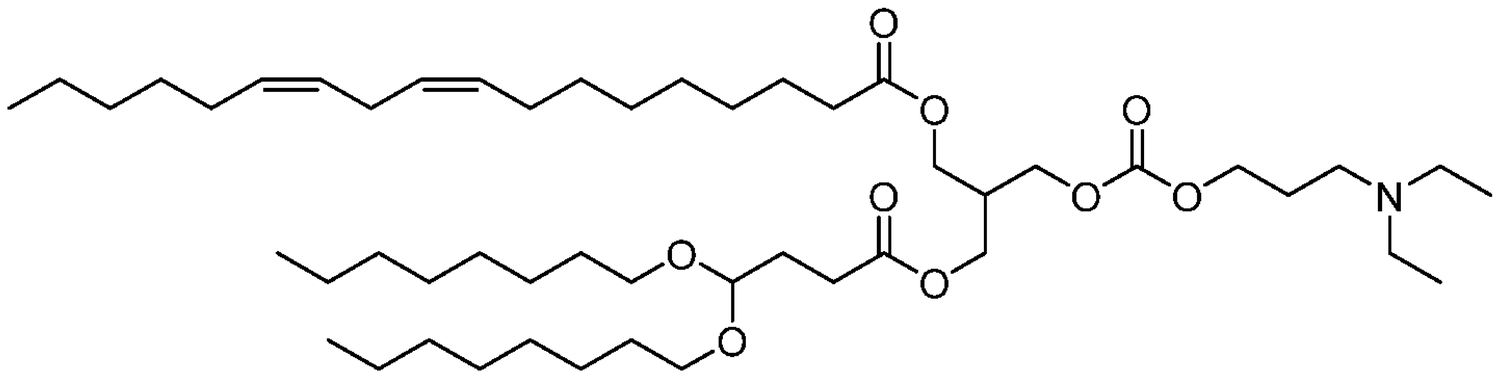
[00345] In some embodiments, the aqueous solution comprises a nucleic acid encoding a polypeptide comprising an A3A and an RNA-guided nickase. A pharmaceutical formulation comprising the lipid nucleic acid assembly composition may optionally comprise a pharmaceutically acceptable buffer.

[00346] In some embodiments, the lipid nucleic acid assembly compositions include an "amine lipid" (sometimes herein or elsewhere described as an "ionizable lipid" or a "biodegradable lipid"), together with an optional "helper lipid", a "neutral lipid", and a stealth lipid such as a PEG lipid. In some embodiments, the amine lipids or ionizable lipids are cationic depending on the pH.

1. Amine Lipids

[00347] In some embodiments, lipid nucleic acid assembly compositions comprise an "amine lipid", which is, for example an ionizable lipid such as Lipid A or its equivalents, including acetal analogs of Lipid A.

[00348] In some embodiments, the amine lipid is Lipid A, which is (9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-((((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-((((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl (9Z,12Z)-octadeca-9,12-dienoate. Lipid A can be depicted as:

[00349] 

[00350] Lipid A may be synthesized according to WO2015/095340 (*e.g.*, pp. 84-86). In some embodiments, the amine lipid is an equivalent to Lipid A.

[00351] In some embodiments, an amine lipid is an analog of Lipid A. In some embodiments, a Lipid A analog is an acetal analog of Lipid A. In particular lipid nucleic acid assembly compositions, the acetal analog is a C4-C12 acetal analog. In some embodiments, the acetal analog is a C5-C12 acetal analog. In additional embodiments, the acetal analog is a C5-C10 acetal analog. In further embodiments, the acetal analog is chosen from a C4, C5, C6, C7, C9, C10, C11, and C12 acetal analog.

[00352] Amine lipids and other "biodegradable lipids" suitable for use in the lipid nucleic acid assemblies described herein are biodegradable *in vivo* or *ex vivo*. The amine lipids have low toxicity (*e.g.*, are tolerated in animal models without adverse effect in amounts of greater than or equal to 10 mg/kg). In some embodiments, lipid nucleic acid assemblies comprising an amine lipid include those where at least 75% of the amine lipid is cleared from the plasma or the engineered cell within 8, 10, 12, 24, or 48 hours, or 3, 4, 5, 6, 7, or 10 days. In some embodiments, lipid nucleic acid assemblies comprising an amine lipid include those where at least 50% of the nucleic acid, *e.g.*, mRNA or gRNA, is cleared from the plasma within 8, 10, 12, 24, or 48 hours, or 3, 4, 5, 6, 7, or 10 days. In some embodiments, lipid nucleic acid assemblies comprising an amine lipid include those where at least 50% of the lipid nucleic acid assembly is cleared from the plasma within 8, 10, 12, 24, or 48 hours, or 3, 4, 5, 6, 7, or 10 days, for example by measuring a lipid (*e.g.* an amine lipid), nucleic acid, *e.g.*, RNA/mRNA, or other component. In some embodiments, lipid-encapsulated versus free lipid, RNA, or nucleic acid component of the lipid nucleic acid assembly is measured.

[00353] Biodegradable lipids include, for example the biodegradable lipids of WO/2020/219876, WO/2020/118041, WO/2020/072605, WO/2019/067992, WO/2017/173054, WO2015/095340, and WO2014/136086, and LNPs include LNP compositions described therein, the lipids and compositions of which are hereby incorporated by reference.

[00354] Lipid clearance may be measured as described in literature. *See* Maier, M.A., *et al.* Biodegradable Lipids Enabling Rapidly Eliminated Lipid Nanoparticles for Systemic Delivery of RNAi Therapeutics. *Mol. Ther.* 2013, 21(8), 1570-78 ("*Maier*"). For example, in *Maier*, LNP-siRNA systems containing luciferases-targeting siRNA were administered to six- to eight-week old male C57Bl/6 mice at 0.3 mg/kg by intravenous bolus injection *via* the lateral tail vein. Blood, liver, and spleen samples were collected at 0.083, 0.25, 0.5, 1, 2, 4, 8, 24, 48, 96, and 168 hours post-dose. Mice were perfused with saline before tissue collection and blood samples were processed to obtain plasma. All samples were processed and analyzed by LC-MS. Further, *Maier* describes a procedure for assessing toxicity after

administration of LNP-siRNA formulations. For example, a luciferase-targeting siRNA was administered at 0, 1, 3, 5, and 10 mg/kg (5 animals/group) via single intravenous bolus injection at a dose volume of 5 mL/kg to male Sprague-Dawley rats. After 24 hours, about 1 mL of blood was obtained from the jugular vein of conscious animals and the serum was isolated. At 72 hours post-dose, all animals were euthanized for necropsy. Assessments of clinical signs, body weight, serum chemistry, organ weights and histopathology were performed. Although *Maier* describes methods for assessing siRNA-LNP formulations, these methods may be applied to assess clearance, pharmacokinetics, and toxicity of administration of lipid nucleic acid assembly compositions of the present disclosure.

[00355] Ionizable and bioavailable lipids for LNP delivery of nucleic acids known in the art are suitable. Lipids may be ionizable depending upon the pH of the medium they are in. For example, in a slightly acidic medium, the lipid, such as an amine lipid, may be protonated and thus bear a positive charge. Conversely, in a slightly basic medium, such as, for example, blood where pH is approximately 7.35, the lipid, such as an amine lipid, may not be protonated and thus bear no charge.

[00356] The ability of a lipid to bear a charge is related to its intrinsic pKa. In some embodiments, the amine lipids of the present disclosure may each, independently, have a pKa in the range of from about 5.1 to about 7.4. In some embodiments, the bioavailable lipids of the present disclosure may each, independently, have a pKa in the range of from about 5.1 to about 7.4, such as from about 5.5 to about 6.6, from about 5.6 to about 6.4, from about 5.8 to about 6.2, or from about 5.8 to about 6.5. For example, the amine lipids of the present disclosure may each, independently, have a pKa in the range of from about 5.8 to about 6.5. Lipids with a pKa ranging from about 5.1 to about 7.4 are effective for delivery of cargo in vivo, *e.g.* to the liver. Further, it has been found that lipids with a pKa ranging from about 5.3 to about 6.4 are effective for delivery in vivo, *e.g.* to tumors. *See*, *e.g.*, WO2014/136086.

2. Additional Lipids

[00357] "Neutral lipids" suitable for use in a lipid nucleic acid assembly composition of the disclosure include, for example, a variety of neutral, uncharged or zwitterionic lipids. Examples of neutral phospholipids suitable for use in the present disclosure include, but are not limited to, 5-heptadecylbenzene-1,3-diol (resorcinol), dipalmitoylphosphatidylcholine (DPPC), distearoylphosphatidylcholine (DSPC), pohsphocholine (DOPC), dimyristoylphosphatidylcholine (DMPC), phosphatidylcholine (PLPC), 1,2-distearoyl-sn-glycero-3-phosphocholine (DAPC), phosphatidylethanolamine (PE), egg phosphatidylcholine

(EPC), dilauryloylphosphatidylcholine (DLPC), dimyristoylphosphatidylcholine (DMPC), 1-myristoyl-2-palmitoyl phosphatidylcholine (MPPC), 1-palmitoyl-2-myristoyl phosphatidylcholine (PMPC), 1-palmitoyl-2-stearoyl phosphatidylcholine (PSPC), 1,2-diarachidoyl-sn-glycero-3-phosphocholine (DBPC), 1-stearoyl-2-palmitoyl phosphatidylcholine (SPPC), 1,2-dieicosenoyl-sn-glycero-3-phosphocholine (DEPC), palmitoyloleoyl phosphatidylcholine (POPC), lysophosphatidyl choline, dioleoyl phosphatidylethanolamine (DOPE), dilinoleoylphosphatidylcholine distearoylphosphatidylethanolamine (DSPE), dimyristoyl phosphatidylethanolamine (DMPE), dipalmitoyl phosphatidylethanolamine (DPPE), palmitoyloleoyl phosphatidylethanolamine (POPE), lysophosphatidylethanolamine and combinations thereof. In one embodiment, the neutral phospholipid may be selected from the group consisting of distearoylphosphatidylcholine (DSPC) and dimyristoyl phosphatidyl ethanolamine (DMPE). In another embodiment, the neutral phospholipid may be distearoylphosphatidylcholine (DSPC).

[00358] "Helper lipids" include steroids, sterols, and alkyl resorcinols. Helper lipids suitable for use in the present disclosure include, but are not limited to, cholesterol, 5-heptadecylresorcinol, and cholesterol hemisuccinate. In one embodiment, the helper lipid may be cholesterol. In one embodiment, the helper lipid may be cholesterol hemisuccinate.

[00359] "Stealth lipids" are lipids that alter the length of time the nanoparticles can exist *in vivo* (*e.g.*, in the blood). Stealth lipids may assist in the formulation process by, for example, reducing particle aggregation and controlling particle size. Stealth lipids used herein may modulate pharmacokinetic properties of the lipid nucleic acid assembly or aid in stability of the nanoparticle *ex vivo*. Stealth lipids suitable for use in a lipid nucleic acid assembly composition of the disclosure include, but are not limited to, stealth lipids having a hydrophilic head group linked to a lipid moiety. Stealth lipids suitable for use in a lipid nucleic acid assembly composition of the present disclosure and information about the biochemistry of such lipids can be found in Romberg *et al.*, Pharmaceutical Research, Vol. 25, No. 1, 2008, pg. 55-71 and Hoekstra *et al.*, Biochimica et Biophysica Acta 1660 (2004) 41-52. Additional suitable PEG lipids are disclosed, *e.g.*, in WO 2006/007712.

[00360] In one embodiment, the hydrophilic head group of stealth lipid comprises a polymer moiety selected from polymers based on PEG. Stealth lipids may comprise a lipid moiety. In some embodiments, the stealth lipid is a PEG lipid.

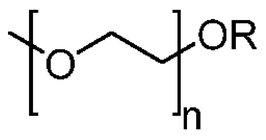
[00361] In one embodiment, a stealth lipid comprises a polymer moiety selected from polymers based on PEG (sometimes referred to as poly(ethylene oxide)), poly(oxazoline),

poly(vinyl alcohol), poly(glycerol), poly(N-vinylpyrrolidone), polyaminoacids and poly[N-(2-hydroxypropyl)methacrylamide].

[00362] In one embodiment, the PEG lipid comprises a polymer moiety based on PEG (sometimes referred to as poly(ethylene oxide)).

[00363] The PEG lipid further comprises a lipid moiety. In some embodiments, the lipid moiety may be derived from diacylglycerol or diacylglycamide, including those comprising a dialkylglycerol or dialkylglycamide group having alkyl chain length independently comprising from about C4 to about C40 saturated or unsaturated carbon atoms, wherein the chain may comprise one or more functional groups such as, for example, an amide or ester. In some embodiments, the alkyl chain length comprises about C10 to C20. The dialkylglycerol or dialkylglycamide group can further comprise one or more substituted alkyl groups. The chain lengths may be symmetrical or asymmetrical.

[00364] Unless otherwise indicated, the term "PEG" as used herein means any polyethylene glycol or other polyalkylene ether polymer. In one embodiment, PEG is an optionally substituted linear or branched polymer of ethylene glycol or ethylene oxide. In one embodiment, PEG is unsubstituted. In one embodiment, the PEG is substituted, *e.g.*, by one or more alkyl, alkoxy, acyl, hydroxy, or aryl groups. In one embodiment, the term includes PEG copolymers such as PEG-polyurethane or PEG-polypropylene (see, *e.g.*, J. Milton Harris, Poly(ethylene glycol) chemistry: biotechnical and biomedical applications (1992)); in another embodiment, the term does not include PEG copolymers. In one embodiment, the PEG has a molecular weight of from about 130 to about 50,000, in a sub-embodiment, about 150 to about 30,000, in a sub-embodiment, about 150 to about 20,000, in a sub-embodiment about 150 to about 15,000, in a sub-embodiment, about 150 to about 10,000, in a sub-embodiment, about 150 to about 6,000, in a sub-embodiment, about 150 to about 5,000, in a sub-embodiment, about 150 to about 4,000, in a sub-embodiment, about 150 to about 3,000, in a sub-embodiment, about 300 to about 3,000, in a sub-embodiment, about 1,000 to about 3,000, and in a sub-embodiment, about 1,500 to about 2,500.

[00365] In some embodiments, the PEG (*e.g.*, conjugated to a lipid moiety or lipid, such as a stealth lipid), is a "PEG-2K," also termed "PEG 2000," which has an average molecular weight of about 2,000 daltons. PEG-2K is represented herein by the following formula (I), wherein n is 45, meaning that the number averaged degree of polymerization comprises about 45 subunits 
(I)
. However, other PEG embodiments known in the art may

be used, including, *e.g.*, those where the number-averaged degree of polymerization comprises about 23 subunits (n=23), and/or 68 subunits (n=68). In some embodiments, n may range from about 30 to about 60. In some embodiments, n may range from about 35 to about 55. In some embodiments, n may range from about 40 to about 50. In some embodiments, n may range from about 42 to about 48. In some embodiments, n may be 45. In some embodiments, R may be selected from H, substituted alkyl, and unsubstituted alkyl. In some embodiments, R may be unsubstituted alkyl. In some embodiments, R may be methyl.

[00366] In any of the embodiments described herein, the PEG lipid may be selected from PEG-dilauroylglycerol, PEG-dimyristoylglycerol (PEG-DMG) (catalog # GM-020 from NOF, Tokyo, Japan), PEG-dipalmitoylglycerol, PEG-distearoylglycerol (PEG-DSPE) (catalog # DSPE-020CN, NOF, Tokyo, Japan), PEG-dilaurylglycamide, PEG-dimyristylglycamide, PEG-dipalmitoylglycamide, and PEG-distearoylglycamide, PEG-cholesterol (1-[8'-(Cholest-5-en-3[beta]-oxy)carboxamido-3',6'-dioxaoctanyl]carbamoyl-[omega]-methyl-poly(ethylene glycol), PEG-DMB (3,4-ditetradecoxylbenzyl-[omega]-methyl-poly(ethylene glycol)ether), 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (PEG2k-DMG) (cat. #880150P from Avanti Polar Lipids, Alabaster, Alabama, USA), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (PEG2k-DSPE) (cat. #880120C from Avanti Polar Lipids, Alabaster, Alabama, USA), 1,2-distearoyl-sn-glycerol, methoxypolyethylene glycol (PEG2k-DSG; GS-020, NOF Tokyo, Japan), poly(ethylene glycol)-2000-dimethacrylate (PEG2k-DMA), and 1,2-distearyloxypropyl-3-amine-N-[methoxy(polyethylene glycol)-2000] (PEG2k-DSA). In one embodiment, the PEG lipid may be PEG2k-DMG. In some embodiments, the PEG lipid may be PEG2k-DSG. In one embodiment, the PEG lipid may be PEG2k-DSPE. In one embodiment, the PEG lipid may be PEG2k-DMA. In one embodiment, the PEG lipid may be PEG2k-C-DMA. In one embodiment, the PEG lipid may be compound S027, disclosed in WO2016/010840 (paragraphs [00240] to [00244]). In one embodiment, the PEG lipid may be PEG2k-DSA. In one embodiment, the PEG lipid may be PEG2k-C11. In some embodiments, the PEG lipid may be PEG2k-C14. In some embodiments, the PEG lipid may be PEG2k-C16. In some embodiments, the PEG lipid may be PEG2k-C18.

3. Formulations

[00367] The lipid nucleic acid assembly may contain (i) a biodegradable lipid, (ii) an optional neutral lipid, (iii) a helper lipid, and (iv) a stealth lipid, such as a PEG lipid. The lipid nucleic acid assembly may contain a biodegradable lipid and one or more of a neutral lipid, a helper lipid, and a stealth lipid, such as a PEG lipid.

[00368] The lipid nucleic acid assembly may contain (i) an amine lipid for encapsulation and for endosomal escape, (ii) a neutral lipid for stabilization, (iii) a helper lipid, also for stabilization, and (iv) a stealth lipid, such as a PEG lipid. The lipid nucleic acid assembly may contain an amine lipid and one or more of a neutral lipid, a helper lipid, also for stabilization, and a stealth lipid, such as a PEG lipid.

[00369] The mRNAs required to achieve the described functional effects described herein may be delivered to a cell in one or more lipid nucleic acid assembly composition(s). For example, one lipid nucleic acid assembly composition may be formulated for delivery comprising mRNA encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase and additional mRNAs encoding, for example, one or more UGIs and one or more gRNAs. Alternatively, the mRNA encoding the polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase, mRNA encoding one or more UGI, and mRNA encoding one or more gRNAs may be formulated in separate lipid nucleic acid assembly compositions. As such, one or multiple lipid nucleic acid assembly composition(s) may be delivered to a cell in vitro or in vivo.

[00370] In some embodiments, a method of modifying a target gene in a cell is provided, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, , optionally lipid nanoparticles, comprising:

(a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase;

(b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs.

[00371] In some embodiments, parts (a) and (b) are in separate lipid nucleic acid assembly compositions. In some embodiments, parts (a) and (b) are in the same lipid nucleic acid assembly composition. In some embodiments, parts (a) and (c) are in separate lipid nucleic

acid assembly compositions. In some embodiments, parts (a) and (c) are in the same lipid nucleic acid assembly composition. In some embodiments, parts (b) and (c) are in separate lipid nucleic acid assembly compositions. In some embodiments, parts (a) and (c) are in the same lipid nucleic acid assembly composition, and part (b) is in a separate lipid nucleic acid assembly composition. In some embodiments, parts (a), (b), and (c) are each in separate lipid nucleic acid assembly compositions. In some embodiments, parts (a), (b), and (c) are in the same lipid nucleic acid assembly composition. In some embodiments, the one or more guide RNAs are each in separate lipid nucleic acid assembly compositions.

[00372] In some embodiments, the method further comprise delivering one or more guide RNAs in one or more lipid nucleic acid assembly compositions that are separate from the lipid nucleic acid assembly compositions comprising the A3A and UGI.

[00373] In some embodiments, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 lipid nucleic acid assembly compositions are delivered to the cell. In some embodiments, the at least one lipid nucleic acid assembly composition comprises lipid nanoparticle (LNPs). In some embodiments, all lipid nucleic acid assembly compositions comprise LNPs. In some embodiments, at least one lipid nucleic acid assembly composition is a lipoplex composition.

[00374] In some embodiments, the lipid nucleic acid assembly composition, e.g., LNP composition, comprises an mRNA that encodes a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase, as described herein. In some embodiments, the lipid nucleic acid assembly composition, e.g. LNP composition, comprises an mRNA that encodes a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase; and a gRNA.

[00375] In some embodiments, the lipid nucleic acid assembly composition comprises a first lipid nucleic acid assembly composition comprising an mRNA encoding a first polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase. In some embodiments, the lipid nucleic acid assembly composition further comprises a second lipid nucleic acid assembly composition comprising a gRNA.

[00376] In some embodiments, the lipid nucleic acid assembly composition comprises a first composition comprising a mRNA encoding a first polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; and one or more second mRNA encoding a uracil glycosylase inhibitor (UGI). In some embodiments, the lipid nucleic acid assembly composition further comprises one or more gRNA.

[00377] In some embodiments, the lipid nucleic acid assembly composition further comprises a second lipid nucleic acid assembly composition comprising a gRNA. In some embodiments, the lipid nucleic acid assembly composition comprises first and second lipid nucleic acid assembly compositions, wherein the first composition comprises an mRNA encoding polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; and the second composition comprises one or more mRNA encoding a uracil glycosylase inhibitor (UGI). In some embodiments, the first lipid nucleic acid assembly composition or the second lipid nucleic acid assembly composition further comprises one or more gRNA. In some embodiments, the lipid nucleic acid assembly composition further comprises a third lipid nucleic acid assembly composition comprising one or more gRNA.

[00378] In some embodiments, the lipid nucleic acid assembly composition comprises a first lipid nucleic acid assembly composition comprising an mRNA comprising an open reading frame encoding a first polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase and a second lipid nucleic acid assembly composition comprising one or more guide RNA (gRNA).

[00379] In some embodiments, the lipid nucleic acid assembly composition comprises a first composition comprising a first mRNA comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; and a second mRNA comprising one or more second open reading frame encoding a uracil glycosylase inhibitor (UGI). In some embodiments, the lipid nucleic acid assembly composition further comprises one or more gRNA. In some embodiments, the lipid nucleic acid assembly composition further comprises a second lipid nucleic acid assembly composition comprising one or more gRNA.

[00380] In some embodiments, the lipid nucleic acid assembly composition comprises first and second lipid nucleic acid assembly compositions, wherein the first composition comprises a first mRNA comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; the second composition comprises a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI). In some embodiments, the first lipid nucleic acid assembly composition or the second lipid nucleic acid assembly composition further comprises a gRNA. In some embodiments, the lipid nucleic acid assembly composition further comprises a third lipid nucleic acid assembly composition comprising a gRNA.

[00381] In some embodiments, the lipid nucleic acid assembly composition comprises a first composition comprising a first mRNA comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; and the second composition comprises a uracil glycosylase inhibitor (UGI). In some embodiments, the first lipid nucleic acid assembly composition or the second lipid nucleic acid assembly composition further comprises a gRNA. In some embodiments, the lipid nucleic acid assembly composition further comprises a third lipid nucleic acid assembly composition comprising a gRNA.

[00382] In some embodiments, the lipid nucleic acid assembly composition comprises first and second lipid nucleic acid assembly compositions, wherein the first composition comprises a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; the second composition comprises a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI). In some embodiments, the first lipid nucleic acid assembly composition or the second lipid nucleic acid assembly composition further comprises a gRNA. In some embodiments, the lipid nucleic acid assembly composition further comprises a third lipid nucleic acid assembly composition comprising a gRNA.

[00383] In certain embodiments, a lipid nucleic acid assembly composition may comprise mRNA, optionally a gRNA, an amine lipid, a helper lipid, a neutral lipid, and a stealth lipid. In certain lipid nucleic acid assembly compositions, the helper lipid is cholesterol. In other lipid nucleic acid assembly compositions, the neutral lipid is DSPC. In additional embodiments, the stealth lipid is PEG2k-DMG or PEG2k-C11. In certain embodiments, the lipid nucleic acid assembly composition comprises Lipid A or an equivalent of Lipid A; a helper lipid; a neutral lipid; a stealth lipid. In certain compositions, the amine lipid is Lipid A. In certain compositions, the amine lipid is Lipid A or an acetal analog thereof; the helper lipid is cholesterol; the neutral lipid is DSPC; and the stealth lipid is PEG2k-DMG.

[00384] Embodiments of the present disclosure also provide lipid nucleic acid assembly compositions described according to the molar ratio between the positively charged amine groups of the amine lipid (N) and the negatively charged phosphate groups (P) of the nucleic acid to be encapsulated. This may be mathematically represented by the equation N/P. In some embodiments, a lipid nucleic acid assembly composition may comprise a lipid component that comprises an amine lipid, a helper lipid, a neutral lipid, and a PEG lipid; and a nucleic acid component, wherein the N/P ratio is about 3 to 10. In some embodiments, a

lipid nucleic acid assembly composition may comprise a lipid component that comprises an amine lipid, a helper lipid, and a PEG lipid; and a nucleic acid component, wherein the N/P ratio is about 3 to 10. In some embodiments, a lipid nucleic acid assembly composition may comprise a lipid component that comprises an amine lipid, a helper lipid, a neutral lipid, and a helper lipid; and an RNA component, wherein the N/P ratio is about 3 to 10. In some embodiments, a lipid nucleic acid assembly composition may comprise a lipid component that comprises an amine lipid, a helper lipid, and a PEG lipid; and an RNA component, such as an mRNA or gRNA, wherein the N/P ratio is about 3 to 10. In one embodiment, the N/P ratio may be about 5 to 7. In one embodiment, the N/P ratio may be about 3 to 7. In one embodiment, the N/P ratio may be about 4.5 to 8. In one embodiment, the N/P ratio may be about 6. In one embodiment, the N/P ratio may be 6 ± 1. In one embodiment, the N/P ratio may be 6 ± 0.5. In some embodiments, the N/P ratio will be ±30%, ±25%, ±20%, ±15%, ±10%, ±5%, or ±2.5% of the target N/P ratio. In certain embodiments, LNP inter-lot variability will be less than 15%, less than 10% or less than 5%.

[00385] In some embodiments, lipid nucleic acid assembly compositions are formed by mixing an aqueous RNA solution with an organic solvent-based lipid solution, e.g., 100% ethanol. Suitable solutions or solvents include or may contain: water, PBS, Tris buffer, NaCl, citrate buffer, ethanol, chloroform, diethylether, cyclohexane, tetrahydrofuran, methanol, isopropanol. A pharmaceutically acceptable buffer, e.g., for in vivo administration of lipid nucleic acid assembly compositions, may be used. In certain embodiments, a buffer is used to maintain the pH of the composition comprising lipid nucleic acid assembly compositions at or above pH 6.5. In certain embodiments, a buffer is used to maintain the pH of the composition comprising LNPs at or above pH 7.0. In certain embodiments, the composition has a pH ranging from about 7.2 to about 7.7. In additional embodiments, the composition has a pH ranging from about 7.3 to about 7.7 or ranging from about 7.4 to about 7.6. In further embodiments, the composition has a pH of about 7.2, 7.3, 7.4, 7.5, 7.6, or 7.7. The pH of a composition may be measured with a micro pH probe. In certain embodiments, a cryoprotectant is included in the composition. Non-limiting examples of cryoprotectants include sucrose, trehalose, glycerol, DMSO, and ethylene glycol. Exemplary compositions may include up to 10% cryoprotectant, such as, for example, sucrose. In certain embodiments, the lipid nucleic acid assembly composition may include about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10% cryoprotectant. In certain embodiments, the lipid nucleic acid assembly composition may include about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10% sucrose. In some embodiments, the lipid nucleic acid assembly composition may include a buffer. In some

embodiments, the buffer may comprise a phosphate buffer (PBS), a Tris buffer, a citrate buffer, and mixtures thereof. In certain exemplary embodiments, the buffer comprises NaCl. In certain embodiments, NaCl is omitted. Exemplary amounts of NaCl may range from about 20 mM to about 45 mM. Exemplary amounts of NaCl may range from about 40 mM to about 50 mM. In some embodiments, the amount of NaCl is about 45 mM. In some embodiments, the buffer is a Tris buffer. Exemplary amounts of Tris may range from about 20 mM to about 60 mM. Exemplary amounts of Tris may range from about 40 mM to about 60 mM. In some embodiments, the amount of Tris is about 50 mM. In some embodiments, the buffer comprises NaCl and Tris. Certain exemplary embodiments of the lipid nucleic acid assembly compositions contain 5% sucrose and 45 mM NaCl in Tris buffer. In other exemplary embodiments, compositions contain sucrose in an amount of about 5% w/v, about 45 mM NaCl, and about 50 mM Tris at pH 7.5. The salt, buffer, and cryoprotectant amounts may be varied such that the osmolality of the overall formulation is maintained. For example, the final osmolality may be maintained at less than 450 mOsm/L. In further embodiments, the osmolality is between 350 and 250 mOsm/L. Certain embodiments have a final osmolality of 300 +/- 20 mOsm/L.

[00386] In some embodiments, microfluidic mixing, T-mixing, or cross-mixing is used. In certain aspects, flow rates, junction size, junction geometry, junction shape, tube diameter, solutions, and/or nucleic acids and lipid concentrations may be varied. Lipid nucleic acid assembly compositions may be concentrated or purified, *e.g.*, via dialysis, tangential flow filtration, or chromatography. The lipid nucleic acid assembly compositions may be stored as a suspension, an emulsion, or a lyophilized powder, for example. In some embodiments, an lipid nucleic acid assembly composition is stored at 2-8° C, in certain aspects, the LNP compositions are stored at room temperature. In additional embodiments, a lipid nucleic acid assembly composition is stored frozen, for example at -20° C or -80° C. In other embodiments, a lipid nucleic acid assembly composition is stored at a temperature ranging from about 0° C to about -80° C. Frozen lipid nucleic acid assembly compositions may be thawed before use, for example on ice, at room temperature, or at 25° C.

[00387] The lipid nucleic acid assembly compositions may be, *e.g.*, microspheres, a dispersed phase in an emulsion, micelles, or an internal phase in a suspension.

[00388] Moreover, in some embodiments, the lipid nucleic acid assembly compositions are biodegradable, in that they do not accumulate to cytotoxic levels *in vivo* at a therapeutically effective dose. In some embodiments, the lipid nucleic acid assembly

compositions do not cause an innate immune response that leads to substantial adverse effects at a therapeutic dose level. In some embodiments, the lipid nucleic acid assembly compositions provided herein do not cause toxicity at a therapeutic dose level.

[00389] The LNPs disclosed herein may have a size (e.g., Z-average diameter) of about 1 to about 150 nm. In some embodiments, the LNPs have a size of about 10 to about 200 nm. In some embodiments, the LNPs have a size of about 50 to about 100 nm. In some embodiments, the LNPs have a size of about 60 to about 100 nm. In some embodiments, the LNPs have a size of about 75 to about 100 nm. In some embodiments, the LNP composition comprises a population of the LNP with an average diameter of about 20-100 nm. In some embodiments, the LNP composition comprises a population of the LNP with an average diameter of about 50-100 nm. In some embodiments, the LNP composition comprises a population of the LNP with an average diameter of about 60-100 nm. In some embodiments, the LNP composition comprises a population of the LNP with an average diameter of or about 75-100 nm. Unless indicated otherwise, all sizes referred to herein are the average sizes (diameters) of the fully formed nanoparticles, as measured by dynamic light scattering on a Malvern Zetasizer. The nanoparticle sample is diluted in phosphate buffered saline (PBS) so that the count rate is approximately 200-400 kcps. The data is presented as a weighted-average of the intensity measure (Z-average diameter).

[00390] In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from about 50% to about 100%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from about 50% to about 70%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from about 70% to about 90%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from about 90% to about 100%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from about 75% to about 95%.

[00391] In some embodiments, the LNPs are formed with an average molecular weight ranging from about 1.00E+05 g/mol to about 1.00E+10 g/mol. In some embodiments, the LNPs are formed with an average molecular weight ranging from about 5.00E+05 g/mol to about 7.00E+07g/mol. In some embodiments, the LNPs are formed with an average molecular weight ranging from about 1.00E+06 g/mol to about 1.00E+10 g/mol. In some embodiments, the LNPs are formed with an average molecular weight ranging from about 1.00E+07 g/mol to about 1.00E+09 g/mol. In some embodiments, the LNPs are formed with an average molecular weight ranging from about 5.00E+06 g/mol to about 5.00E+09 g/mol.

[00392] In some embodiments, the polydispersity (Mw/Mn; the ratio of the weight averaged molar mass (Mw) to the number averaged molar mass (Mn)) may range from about 1.000 to about 2.000. In some embodiments, the Mw/Mn may range from about 1.00 to about 1.500. In some embodiments, the Mw/Mn may range from about 1.020 to about 1.400. In some embodiments, the Mw/Mn may range from about 1.010 to about 1.100. In some embodiments, the Mw/Mn may range from about 1.100 to about 1.350.

[00393] Dynamic Light Scattering ("DLS") can be used to characterize the polydispersity index ("pdi") and size of the LNPs of the present disclosure. DLS measures the scattering of light that results from subjecting a sample to a light source. PDI, as determined from DLS measurements, represents the distribution of particle size (around the mean particle size) in a population, with a perfectly uniform population having a PDI of zero. In some embodiments, the pdi may range from 0.005 to 0.75. In some embodiments, the pdi may range from 0.01 to 0.5. In some embodiments, the pdi may range from 0.02 to 0.4. In some embodiments, the pdi may range from 0.03 to 0.35. In some embodiments, the pdi may range from 0.1 to 0.35. In some embodiments, the pdi may range about zero to about 0.4, such as about zero to about 0.35. In some embodiments, the pdi may range from about zero to about 0.35, about zero to about 0.3, about zero to about 0.25, or about zero to about 0.2. In some embodiments, the pdi is less than about 0.08, 0.1, 0.15, 0.2, or 0.4.

[00394] In some embodiments, LNPs disclosed herein have a size of 1 to 250 nm. In some embodiments, the LNPs have a size of 10 to 200 nm. In further embodiments, the LNPs have a size of 20 to 150 nm. In some embodiments, the LNPs have a size of 50 to 150 nm. In some embodiments, the LNPs have a size of 50 to 100 nm. In some embodiments, the LNPs have a size of 50 to 120 nm. In some embodiments, the LNPs have a size of 75 to 150 nm. In some embodiments, the LNPs have a size of 30 to 200 nm. Unless indicated otherwise, all sizes referred to herein are the average sizes (diameters) of the fully formed nanoparticles, as measured by dynamic light scattering on a Malvern Zetasizer. The nanoparticle sample is diluted in phosphate buffered saline (PBS) so that the count rate is approximately 200-400 kcts. The data is presented as a weighted-average of the intensity measure. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from 50% to 100%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from 50% to 70%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from 70% to 90%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from

90% to 100%. In some embodiments, the LNPs are formed with an average encapsulation efficiency ranging from 75% to 95%.

[00395] Electroporation is also a well-known means for delivery of cargo, and any electroporation methodology may be used for delivery of any one of the RNAs disclosed herein.

[00396] In some embodiments, the methods comprises a method for delivering a composition comprising the mRNA disclosed herein to an *ex vivo* cell, wherein the mRNA is encapsulated in an LNP. In some embodiments, the composition comprises the mRNA and one or more additional RNAs disclosed herein encapsulated in the LNP.

[00397] In some embodiments, a lipid nucleic acid assembly composition comprises a lipid component, wherein the lipid component comprises an amine lipid, a neutral lipid, a helper lipid, and a stealth lipid; and wherein the N/P ratio is about 1-10.

[00398] In some instances, the lipid component comprises Lipid A or its acetal analog, cholesterol, DSPC, and PEG-DMG; and wherein the N/P ratio is about 1-10. In some embodiments, the lipid component comprises: about 40-60 mol-% amine lipid; about 5-15 mol-% neutral lipid; and about 1.5-10 mol-% PEG lipid, wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 3-10. In some embodiments, the lipid component comprises about 50-60 mol-% amine lipid; about 8-10 mol-% neutral lipid; and about 2.5-4 mol-% PEG lipid, wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 3-8. In some instances, the lipid component comprises: about 50-60 mol-% amine lipid; about 5-15 mol-% DSPC; and about 2.5-4 mol-% PEG lipid, wherein the remainder of the lipid component is cholesterol, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 3-8. In some instances, the lipid component comprises: 48-53 mol-% Lipid A; about 8-10 mol-% DSPC; and 1.5-10 mol-% PEG lipid, wherein the remainder of the lipid component is cholesterol, and wherein the N/P ratio of the lipid nucleic acid assembly composition is 3-8 ±0.2.

[00399] In some embodiments, the lipid component comprises about 50-60 mol-% amine lipid such as Lipid A, about 8-10 mol-% neutral lipid; and about 2.5-4 mol-% stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6. In some embodiments, the lipid component comprises about 50-60 mol-% amine lipid such as Lipid A; about 27-39.5 mol-% helper lipid; about 8-10 mol-% neutral lipid; and about 2.5-4 mol-% stealth lipid (e.g., a PEG lipid), wherein the N/P ratio of the lipid nucleic acid assembly

composition is about 5-7 (e.g., about 6). In some embodiments, the lipid component comprises about 50-60 mol-% amine lipid such as Lipid A; about 5-15 mol-% neutral lipid; and about 2.5-4 mol-% Stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 3-10. In some embodiments, the lipid component comprises about 40-60 mol-% amine lipid such as Lipid A; about 5-15 mol-% neutral lipid; and about 2.5-4 mol-% Stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6. In some embodiments, the lipid component comprises about 50-60 mol-% amine lipid such as Lipid A; about 5-15 mol-% neutral lipid; and about 1.5-10 mol-% Stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6. In some embodiments, the lipid component comprises about 40-60 mol-% amine lipid such as Lipid A; about 0-10 mol-% neutral lipid; and about 1.5-10 mol-% Stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 3-10. In some embodiments, the lipid component comprises about 40-60 mol-% amine lipid such as Lipid A; less than about 1 mol-% neutral lipid; and about 1.5-10 mol-% Stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the lipid nucleic acid assembly composition is about 3-10. In some embodiments, the lipid component comprises about 40-60 mol-% amine lipid such as Lipid A; and about 1.5-10 mol-% Stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, wherein the N/P ratio of the lipid nucleic acid assembly composition is about 3-10, and wherein the lipid nucleic acid assembly composition is essentially free of or free of neutral phospholipid. In some embodiments, the lipid component comprises about 50-60 mol-% amine lipid such as Lipid A; about 8-10 mol-% neutral lipid; and about 2.5-4 mol-% Stealth lipid (e.g., a PEG lipid), wherein the remainder of the lipid component is helper lipid, and wherein the N/P ratio of the Lipid nucleic acid assembly composition is about 3-7.

[00400] In some embodiments, the amine lipid is present at about 50 mol-%. In some embodiments, the neutral lipid is present at about 9 mol-%. In some embodiments, the stealth lipid is present at about 3 mol-%. In some embodiments, the helper lipid is present at about 38 mol-%.

[00401] In some embodiments, the lipid component comprises, consists essentially of, or consists of: about 50 mol-% amine lipid such as Lipid A; about 9 mol-%

neutral lipid such as DSPC; about 3 mol-% of a stealth lipid such as a PEG lipid, such as PEG2k-DMG, and the remainder of the lipid component is helper lipid such as cholesterol, wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6. In some embodiments, the amine lipid is Lipid A. In some embodiments, the neutral lipid is DSPC. In some embodiments, the stealth lipid is a PEG lipid. In some embodiments, the stealth lipid is a PEG2k-DMG. In some embodiments, the helper lipid is cholesterol. In some embodiments, the lipid comprises a lipid component and the lipid component comprises: about 50 mol-% Lipid A; about 9 mol-% DSPC; about 3 mol-% of PEG2k-DMG, and the remainder of the lipid component is cholesterol wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6.

[00402] In some embodiments, the lipid component comprises, consists essentially of, or consists of: about 25 to 45 mol-% amine lipid such as Lipid A; about 10 to 30 mol-% neutral lipid such as DSPC; about 1.5 to 3.5 mol-% of a stealth lipid such as a PEG lipid, such as PEG2k-DMG, and about 25 to 65 mol% helper lipid such as cholesterol, wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6. In some embodiments, the lipid component comprises, consists essentially of, or consists of: about 35 mol-% amine lipid such as Lipid A; about 15 mol-% neutral lipid such as DSPC; about 2.5 mol-% of a stealth lipid such as a PEG lipid, such as PEG2k-DMG, and the remainder of the lipid component is helper lipid such as cholesterol, wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6. In some embodiments, the amine lipid is Lipid A. In some embodiments, the neutral lipid is DSPC. In some embodiments, the stealth lipid is a PEG lipid. In some embodiments, the stealth lipid is a PEG2k-DMG. In some embodiments, the helper lipid is cholesterol. In some embodiments, the lipid comprises a lipid component and the lipid component comprises: about 35 mol-% Lipid A; about 15 mol-% DSPC; about 2.5 mol-% of PEG2k-DMG, and the remainder of the lipid component is cholesterol wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6.

I. Exemplary Uses, Methods, And Treatments

[00403] In some embodiments, a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is for use in genome editing, e.g., editing a target gene, or modifying a target gene. In some embodiments, a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is for use in modifying a target gene, e.g., altering its sequence or epigenetic status. In some embodiments, the nucleic acid (e.g., mRNA), polypeptide,

composition, or lipid nucleic acid assembly composition disclosed herein is for use in the manufacture of a medicament for genome editing or modifying a target gene.

[00404] In some embodiments, the use of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is provided for the preparation of a medicament for genome editing, e.g., editing a target gene. In some embodiments, the use of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition is provided for the preparation of a medicament for modifying a target gene, e.g., altering its sequence or epigenetic status. In some embodiments, the use of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is provided for the preparation of a medicament for causing C-to-T conversion within a target gene.

[00405] In some embodiments, a method of genome editing or modifying a target gene is provided, the method comprising delivering to a cell the mRNA, composition, or lipid nanoparticle(s) described herein.

[00406] In some embodiments, the method generates a cytosine (C) to thymine (T) conversion within a target gene.

[00407] In some embodiments, the method causes at least 50% C-to-T conversion relative to the total edits in the target sequence. As used herein, the "total edits in the target sequence" is the sum of each read with an indel or at least one conversion, wherein an indel can comprise more than one nucleotide. Indel is calculated as the total number of sequencing reads with one or more base inserted or deleted within the 20 bp scoring region divided by the total number of sequencing reads, including wild type. C-to-T conversions or C-to-A/G conversions were scored in a 40 bp region including 10 bp upstream and 10 bp downstream of the 20 bp sgRNA target sequence. Any sequencing methods (e.g., NGS) that allow reading of sequences diverged from the wild-type alignment may be used. In some embodiments, the method causes at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% C-to-T conversion relative to the total edits in the target sequence.

[00408] In some embodiments, the ratio of C-to-T conversion to unintended edits is larger than 1:1. As used herein, an "unintended edit" is any edit in the target region that is not a C-to-T conversion. In some embodiments, the ratio of C-to-T conversion to unintended edits is larger than 2:1, larger than 3:1, larger than 4:1, larger than 5:1, larger than 6:1, larger than 7:1, or larger than 8:1. In some embodiments, the ratio of C-to-T conversion

to unintended edits is from 2:1 to 99:1. In some embodiments, the ratio of C-to-T conversion to unintended edits is 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, or 8:1.

[00409] In some embodiments, the method causes the A3A to make a base edit corresponding to any one of positions -1 to 10 relative to the 5' end of the guide sequence.

[00410] In some embodiments, the method causes the A3A to make a base edit at a position 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the guide sequence.

[00411] In some embodiments, the nickase is a SpyCas9 nickase, and the method causes the cytidine deaminase to make a base edit at a cytidine present at position 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 nucleotides from the 5' end of the guide sequence.

[00412] In some embodiments, the nickase is a NmeCas9 nickase, and the method causes the cytidine deaminase to make a base edit at a cytidine present at position 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides from the 5' end of the guide sequence.

[00413] In some embodiments, the composition comprises a first mRNA comprising a first open reading frame encoding a polypeptide comprising an A3A and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), and a gRNA; and the first mRNA, the second mRNA, and the gRNA if present, delivered at a ratio of about 6:2:3 (w:w:w).In some embodiments, the target gene is in a subject, such as a mammal, such as a human.

[00414] In some embodiments, methods are provided for modifying a target gene comprising delivering to a cell a first mRNA comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase, a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, and at least one guide RNA (gRNA).

[00415] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs. In some embodiments, the one or more guide RNAs are each in separate lipid nucleic acid assembly compositions.

[00416] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, and /or one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and/or one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

[00417] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises at least two gRNAs selected from: one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

[00418] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

[00419] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first

open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one gRNA that targets a gene that reduces or eliminates expression of HLA-A on the surface of a cell, and /or one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and/or one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

[00420] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises at least two gRNAs selected from: one gRNA that targets a gene that reduces or eliminates expression of HLA-A on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

[00421] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one gRNA that targets a gene that reduces or eliminates expression of HLA-A on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

[00422] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs, wherein the method comprises one gRNA selected from a gRNA that targets TRAC, TRBC, B2M, HLA-A, or CIITA. In some embodiments, one gRNA targets TRAC. In some embodiments, one gRNA targets TRBC. In some embodiments, one gRNA targets B2M. In some embodiments, one gRNA targets HLA-A. In some embodiments, one gRNA targets CIITA.

[00423] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or B2M, wherein the two guide RNAs do not target the same gene. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions.

[00424] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or HLA-A wherein the two guide RNAs do not target the same gene. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions.

[00425] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises at least two gRNAs selected from a gRNA that targets TRAC, TRBC, HLA-A, wherein the two guide RNAs do not target

the same gene. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions.

[00426] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one guide RNA that targets TRAC, and one gRNA that targets TRBC. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions.

[00427] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one guide RNA that targets B2M, and one gRNA that targets CIITA. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions.

[00428] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one guide RNA that targets HLA-A, and one gRNA that targets CIITA. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions. In some embodiments, the cell is homozygous for HLA-B and homozygous for HLA-C.

[00429] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an

APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA that targets B2M. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions.

[00430] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA that targets HLA-A. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions. In some embodiments, the cell is homozygous for HLA-B and homozygous for HLA-C.

[00431] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets B2M, and one gRNA that targets CIITA. In some embodiments, the gRNAs are each in separate lipid nucleic acid assembly compositions.

[00432] In some embodiments, methods are provided for modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising: (a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase; (b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and (c) one or more guide RNAs, wherein the method comprises one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets HLA-A, and one gRNA that targets CIITA. In some embodiments, the gRNAs are each in separate lipid nucleic acid

assembly compositions. In some embodiments, the cell is homozygous for HLA-B and homozygous for HLA-C.

[00433] In some embodiments, a cell is provided comprising a composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA.

[00434] In some embodiments, an engineered cell is provided comprising at least one base edit and/or indel, wherein the base edit and/or indel is made by contacting a cell with a composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA.

[00435] In some embodiments the cell is a human cell. In some embodiments the genetically modified cell is referred to as an engineered cell. An engineered cell refers to a cell (or progeny of a cell) comprising an engineered genetic modification, *e.g.* that has been contacted with a gene editing system and genetically modified by the gene editing system. The terms "engineered cell" and "genetically modified cell" are used interchangeably throughout. The engineered cell may be any of the exemplary cell types disclosed herein. In some embodiments, the cell is an allogeneic cell.

[00436] In some embodiments, the cell is an immune cell. As used herein, "immune cell" refers to a cell of the immune system, including e.g., a lymphocyte (e.g., T cell, B cell, natural killer cell ("NK cell", and NKT cell, or iNKT cell)), monocyte, macrophage, mast cell, dendritic cell, or granulocyte (e.g., neutrophil, eosinophil, and basophil). In some embodiments, the cell is a primary immune cell. In some embodiments, the immune system cell may be selected from CD3+, CD4+ and CD8+ T cells, regulatory T cells (Tregs), B cells, NK cells, and dendritic cells (DC). In some embodiments, the immune cell is allogeneic. In some embodiments, the cell is a lymphocyte. In some embodiments, the cell is an adaptive immune cell. In some embodiments, the cell is a T cell. In some embodiments, the cell is a B cell. In some embodiments, the cell is a NK cell. In some embodiments, the lymphocyte is allogeneic.

[00437] In some embodiments, the genome editing or modification of the target gene is *in vivo*. In some embodiments, the genome editing or modification of the target gene is in an isolated or cultured cell.

[00438] In some embodiments, the target gene is in an organ, such as a liver, such as a mammalian liver, such as a human liver. In some embodiments, the target gene is in a liver cell, such as a mammalian liver cell, such as a human liver cell. In some embodiments, the target gene is in a hepatocyte, such as a mammalian hepatocyte, such as a human hepatocyte. In some embodiments, the liver cell or hepatocyte is in situ. In some embodiments, the liver cell or hepatocyte is isolated, e.g., in a culture, such as in a primary culture.

[00439] In some embodiments, the genome editing or modification of the target gene inactivates a splice donor or splice acceptor site.

[00440] Also provided are methods corresponding to the uses disclosed herein, which comprise administering the nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein to a subject or contacting a cell such as those described above with the nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein.

[00441] In some embodiments the nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is administered intravenously for any of the uses discussed above concerning organisms, organs, or cells in situ.

[00442] In any of the foregoing embodiments involving a subject, the subject can be mammalian. In any of the foregoing embodiments involving a subject, the subject can be human. In any of the foregoing embodiments involving a subject, the subject can be a cow, pig, monkey, sheep, dog, cat, fish, or poultry.

[00443] In some embodiments, the nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is administered intravenously or for intravenous administration.

[00444] In some embodiments, the genome editing or modification of the target gene knocks down expression of the target gene. In some embodiments, the genome editing or modification of the target gene knocks down expression of the target gene by at least 50%, 55%, 60%, 65%, 70%, 75%, or 80%. In some embodiments, the genome editing or modification of the target gene produces a missense mutation in the gene.

[00445] In some embodiments, a single administration of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is sufficient to knock down expression of the target gene product. In some embodiments, a single administration of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is sufficient to knock out expression of the target gene product. In other embodiments, more than one administration of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein may be beneficial to maximize editing via cumulative effects.

[00446] In some embodiments, the efficacy of treatment with a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein is seen at 1 year, 2 years, 3 years, 4 years, 5 years, or 10 years after delivery.

[00447] In some embodiments, treatment slows or halts disease progression.

[00448] In some embodiments, treatment results in improvement, stabilization, or slowing of change in organ function or symptoms of disease of an organ.

[00449] In some embodiments, efficacy of treatment is measured by increased survival time of the subject.

1. Exemplary Guide RNAs, Compositions, Methods, and Engineered Cells for TRAC and TRBC editing

[00450] The disclosure provides a guide RNA that target TRAC. Guide sequences targeting the TRAC gene are shown in Table 5A at SEQ ID NOs: 706-721.

[00451] The disclosure provides a guide RNA that target TRBC. Guide sequences targeting the TRBC gene are shown in Table 5B at SEQ ID NOs: 618-669.

[00452] In some embodiments, the guide sequences are complementary to the corresponding genomic region shown in the tables below, according to coordinates from human reference genome hg38. Guide sequences of further embodiments may be complementary to sequences in the close vicinity of the genomic coordinate listed in any of Tables 5A and 5B. For example, guide sequences of further embodiments may be complementary to sequences that comprise 15 consecutive nucleotides ±10 nucleotides of a genomic coordinate listed in any of Tables 5A an 5B.

[00453] As described in the preceding sections, each of the guide sequences shown in Table 5A and Table 5B may further comprise additional nucleotides to form a crRNA, e.g., with the following exemplary nucleotide sequence following the guide sequence

at its 3’ end: GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 139) in 5’ to 3’ orientation. In the case of a sgRNA, the guide sequences may further comprise additional nucleotides to form a sgRNA, e.g., with the following exemplary nucleotide sequence following the 3’ end of the guide sequence: GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 140) in 5’ to 3’ orientation. The guide sequences may further comprise additional nucleotides to form a sgRNA, e.g.,

[00454] In some embodiments, the sgRNA comprises the modification pattern shown below in SEQ ID NO: 141, where N is any natural or non-natural nucleotide, and where the totality of the N’s comprise a guide sequence as described herein and the modified sgRNA comprises the following sequence: mN*mN*mN*NNNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU (SEQ ID NO: 141), where “N” may be any natural or non-natural nucleotide. For example, encompassed herein is SEQ ID NO: 141, where the N’s are replaced with any of the guide sequences disclosed herein. The modifications remain as shown in SEQ ID NO: 141 despite the substitution of N’s for the nucleotides of a guide. That is, although the nucleotides of the guide replace the “N’s”, the first three nucleotides are 2’OMe modified and there are phosphorothioate linkages between the first and second nucleotides, the second and third nucleotides and the third and fourth nucleotides.

[00455] In some embodiments, the gRNA targeting TRAC comprises a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

[00456] Table 5A TRAC guide sequences, guide RNA sequences, and chromosomal coordinates

Guide ID	**SEQ ID NO to the Guide Sequence**	**Guide Sequence**	**Exemplary Full Sequence (SEQ ID NOS: 747 – 762)**	**Exemplary Mod Sequence (SEQ ID NOS: 853 – 868)**	**Genomic Coordinates (hg38)**

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 747 – 762)	Exemplary Mod Sequence (SEQ ID NOS: 853 – 868)	Genomic Coordinates (hg38)
G016019	706	AACAAAUGUGUCACAAAGUA	AACAAAUGUGUCACAAAGUAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mA*mA*mC*AAAUGUGUCACAAAGUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547596-22547616
G016016	707	UUUCAAAACCUGUCAGUGAU	UUUCAAAACCUGUCAGUGAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mU*mU*mU*CAAAACCUGUCAGUGAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22550570-22550590
G016026	708	CUUACCUGGGCUGGGGAAGA	CUUACCUGGGCUGGGGAAGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mU*mU*ACCUGGGCUGGGGAAGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547763-22547783
G016013	709	CCGAAUCCUCCUCCUGAAAG	CCGAAUCCUCCUCCUGAAAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mC*mG*AAUCCUCCUCCUGAAAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22550596-22550616
G016022	710	CUGACAGGUUUUGAAAGUUU	CUGACAGGUUUUGAAAGUUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mU*mG*ACAGGUUUUGAAAGUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22550566-22550586
G016023	711	CUGGGGAAGAAGGUGUCUUC	CUGGGGAAGAAGGUGUCUUCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAU	mC*mU*mG*GGGAAGAAGGUGUCUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAm	chr14: 22547753-22547773

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 747 – 762)	Exemplary Mod Sequence (SEQ ID NOS: 853 – 868)	Genomic Coordinates (hg38)
			CAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	AmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	
G016010	712	UCCUCCUCCUGAAGUGGCC	UCCUCCUCCUGAAAGUGGCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mU*mC*mC*UCCUCCUGAAAGUGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22550601-22550621
G016014	713	CCACUUUCAGGAGGAGG	CCACUUUCAGGAGGAGGAUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mC*mA*CUUUCAGGAGGAGGAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22550599-22550619
G016020	714	AUUUGUUUGAGAAUCAAAAU	AUUUGUUUGAGAAUCAAAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mA*mU*mU*UGUUUGAGAAUCAAAAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547583-22547603
G015998	715	CUUCAAGAGCAACAGUGCUG	CUUCAAGAGCAACAGUGCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mU*mU*CAAGAGCAACAGUGCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547671-22547691
G016025	716	AGCUGCCCUUACCUGGGCUG	AGCUGCCCUUACCUGGGCUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mA*mG*mC*UGCCCUUACCUGGGCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547770-22547790
G015997	717	AGAGCAACAGUG	AGAGCAACAGUGCUGUGGCCGU	mA*mG*mA*GCAACAGUGCUGUGGCCGUUUUAGAm	chr14: 22547676-

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 747 – 762)	Exemplary Mod Sequence (SEQ ID NOS: 853 – 868)	Genomic Coordinates (hg38)
		CUGUGGCC	UUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	GmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	22547696
G016029	718	AAAGCUGCCCUUACCUGGGC	AAAGCUGCCCUUACCUGGGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mA*mA*mA*GCUGCCCUUACCUGGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547772-22547792
G016028	719	AAGCUGCCCUUACCUGGGCU	AAGCUGCCCUUACCUGGGCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mA*mA*mG*CUGCCCUUACCUGGGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547771-22547791
G016021	720	UGGAAUAAUGCUGUUGUUGA	UGGAAUAAUGCUGUUGUUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mU*mG*mG*AAUAAUGCUGUUGUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547733-22547753
G016015	721	CACCAAAGCUGCCCUUACCU	CACCAAAGCUGCCCUUACCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mA*mC*CAAAGCUGCCCUUACCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr14: 22547776-22547796

[00457] In some embodiments, the guide sequence comprises SEQ ID NO: 706. In some embodiments, the guide sequence comprises SEQ ID NO: 707. In some embodiments, the guide sequence comprises SEQ ID NO: 708. In some embodiments, the

guide sequence comprises SEQ ID NO: 709. In some embodiments, the guide sequence comprises SEQ ID NO: 710. In some embodiments, the guide sequence comprises SEQ ID NO: 711. In some embodiments, the guide sequence comprises SEQ ID NO: 712. In some embodiments, the guide sequence comprises SEQ ID NO: 713. In some embodiments, the guide sequence comprises SEQ ID NO: 714. In some embodiments, the guide sequence comprises SEQ ID NO: 715. In some embodiments, the guide sequence comprises SEQ ID NO: 716. In some embodiments, the guide sequence comprises SEQ ID NO: 717. In some embodiments, the guide sequence comprises SEQ ID NO: 718. In some embodiments, the guide sequence comprises SEQ ID NO: 719. In some embodiments, the guide sequence comprises SEQ ID NO: 720. In some embodiments, the guide sequence comprises SEQ ID NO: 721.

[00458] In some embodiments, the gRNA targeting TRBC comprises a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

[00459] Table 5B TRBC guide sequences, guide RNA sequences, and chromosomal coordinates

Guide ID	**SEQ ID NO to the Guide Sequence**	**Guide Sequence**	**Exemplary Full Sequence (SEQ ID NOS: 1024-1075)**	**Exemplary Mod Sequence (SEQ ID NOS: 801-852)**	**Genomic Coordinates (hg38)**
G016200	618	CCACACCCAAAAGGCCACAC	CCACACCCAAAAGGCCACACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mC*mA*CACCCAAAAGGCCACACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791757-142791777 chr7: 142801104-142801124
G016174	619	CCCACCAGCUCAGCUCCACG	CCCACCAGCUCAGCUCCACGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGA	mC*mC*mC*ACCAGCUCAGCUCCACGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmG	chr7: 142791811-142791831; chr7: 142801158-142801178

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
			GUCGGUGCUUU U	mUmCmGmGmUmGmCmU *mU*mU*mU	
G016263	620	UCCCUA GCAGGA UCUCAU AG	UCCCUAGCAGG AUCUCAUAGGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mU*mC*mC*CUAGCAGG AUCUCAUAGGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142792728-142792748
G016270	621	GGCUCA AACACA GCGACC UC	GGCUCAAACAC AGCGACCUCGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mG*mG*mC*UCAAACAC AGCGACCUCGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142791719-142791739
G016212	622	GGCACA CCAGUG UGGCCU UU	GGCACACCAGU GUGGCCUUUGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mG*mG*mC*ACACCAGU GUGGCCUUUGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142791766-142791786; chr7: 142801113-142801133
G016207	623	AGGUG GCCGAG ACCCUC AGG	AGGUGGCCGAG ACCCUCAGGGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mA*mG*mG*UGGCCGAG ACCCUCAGGGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142791928-142791948; chr7: 142801275-142801295
G016205	624	ACCUGC UCUACC CCAGGC CU	ACCUGCUCUAC CCCAGGCCUGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mA*mC*mC*UGCUCUAC CCCAGGCCUGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142792062-142792082; chr7: 142801409-142801429
G016245	625	CAUAGA GGAUG GUGGCA GAC	CAUAGAGGAUG GUGGCAGACGU UUUAGAGCUAG AAAUAGCAAGU	mC*mA*mU*AGAGGAUG GUGGCAGACGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA	chr7: 142792713-142792733; chr7:

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
			UAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	AAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	142802126-142802146
G016232	626	CCACGUGGAGCUGAGCUGGU	CCACGUGGAGCUGAGCUGGUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mC*mA*CGUGGAGCUGAGCUGGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791808-142791828; chr7: 142801155-142801175
G016227	627	GGAGAAUGACGAGUGGACCC	GGAGAAUGACGAGUGGACCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mG*mG*mA*GAAUGACGAGUGGACCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792003-142792023; chr7: 142801350-142801370
G016273	628	CCAGUGUGGCCUUUUGGGUG	CCAGUGUGGCCUUUUGGGUGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mC*mC*mA*GUGUGGCCUUUUGGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791760-142791780
G016251	629	CAAACACAGCGACCUCGGGU	CAAACACAGCGACCUCGGGUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mC*mA*mA*ACACAGCGACCUCGGGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791715-142791735
G016276	630	UACCAUGGCCAUCAACACAA	UACCAUGGCCAUCAACACAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mU*mA*mC*CAUGGCCAUCAACACAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792781-142792801

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
G016167	631	GCGCUGACGAUCUGGGUGAC	GCGCUGACGAUCUGGGUGACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mG*mC*mG*CUGACGAUCUGGGUGACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792040-142792060; chr7: 142801387-142801407
G016253	632	CACGGACCCGCAGCCCCUCA	CACGGACCCGCAGCCCCUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mA*mC*GGACCCGCAGCCCCUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791862-142791882
G016272	633	UCAAACACAGCGACCUCGG	UCAAACACAGCGACCUCGGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mU*mC*mA*AACACAGCGACCUCGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791716-142791736
G016258	634	CAGGGAAGAAGCCUGUGGCC	CAGGGAAGAAGCCUGUGGCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mA*mG*GGAAGAAGCCUGUGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791787-142791807
G016183	635	CAGUGUGGCCUUUGGGUGU	CAGUGUGGCCUUUUGGGUGUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mC*mA*mG*UGUGGCCUUUUGGGUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791759-142791779; chr7: 142801106-142801126
G016222	636	ACCACGUGGAGCUGAGCUGG	ACCACGUGGAGCUGAGCUGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAU	mA*mC*mC*ACGUGGAGCUGAGCUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGm	chr7: 142791807-142791827; chr7: 142801154-142801174

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
			CAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	
G016233	637	GAGGGCGGGCUGCUCCUUGA	GAGGGCGGGCUGCUCCUUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mG*mA*mG*GGCGGGCUGCUCCUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791879-142791899; chr7: 142801226-142801246
G016264	638	AGCUCAGCUCCACGUGGUCA	AGCUCAGCUCCACGUGGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mA*mG*mC*UCAGCUCCACGUGGUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791805-142791825
G016255	639	GAACAAGGUGUUCCCACCCG	GAACAAGGUGUUCCCACCCGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mG*mA*mA*CAAGGUGUUCCCACCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791700-142791720
G016177	640	GCACACCAGUGUGGCCUUUU	GCACACCAGUGUGGCCUUUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mG*mC*mA*CACCAGUGUGGCCUUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791765-142791785; chr7: 142801112-142801132
G016283	641	GAGCUGGUGGGUGAAUGGGA	GAGCUGGUGGGUGAAUGGGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mG*mA*mG*CUGGUGGGUGAAUGGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791820-142791840
G016194	642	GGCUGCUCCUUG	GGCUGCUCCUUGAGGGGCUGGU	mG*mG*mC*UGCUCCUUGAGGGGCUGGUUUUAG	chr7: 142791872-

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
		AGGGGCUG	UUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	AmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	142791892; chr7: 142801219-142801239
G016266	643	CGGGUGGGAACACCUUGUUC	CGGGUGGGAACACCUUGUUCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mG*mG*GUGGGAACACCUUGUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791700-142791720
G016196	644	CAGCUCAGCUCCACGUGGUC	CAGCUCAGCUCCACGUGGUCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mC*mA*mG*CUCAGCUCCACGUGGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791806-142791826; chr7: 142801153-142801173
G016274	645	GACGAUCUGGGUGACGGGUU	GACGAUCUGGGUGACGGGUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mG*mA*mC*GAUCUGGGUGACGGGUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792035-142792055
G016282	646	UAGCAGGAUCUCAUAGAGGA	UAGCAGGAUCUCAUAGAGGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mU*mA*mG*CAGGAUCUCAUAGAGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792724-142792744
G016267	647	GACCAGCACAGCAUACAGGG	GACCAGCACAGCAUACAGGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGA	mG*mA*mC*CAGCACAGCAUACAGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmG	chr7: 142792754-142792774

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
			GUCGGUGCUUU U	mUmCmGmGmUmGmCmU *mU*mU*mU	
G016261	648	CUGACC ACGUGG AGCUGA GC	CUGACCACGUG GAGCUGAGCGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mC*mU*mG*ACCACGUG GAGCUGAGCGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142791804-142791824
G016277	649	CCAACA GUGUCC UACCAG CA	CCAACAGUGUC CUACCAGCAGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mC*mC*mA*ACAGUGUC CUACCAGCAGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142792684-142792704
G016285	650	CUGGUG GGUGA AUGGG AAGG	CUGGUGGGUGA AUGGGAAGGG UUUUAGAGCUA GAAAUAGCAAG UUAAAAUAAG GCUAGUCCGUU AUCAACUUGAA AAAGUGGCACC GAGUCGGUGCU UUU	mC*mU*mG*GUGGGUGA AUGGGAAGGGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142791823-142791843
G016275	651	CUAUGA GAUCCU GCUAGG GA	CUAUGAGAUCC UGCUAGGGAGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	mC*mU*mA*UGAGAUCC UGCUAGGGAGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142792728-142792748
G016281	652	CAGGAU CUCAUA GAGGA UGG	CAGGAUCUCAU AGAGGAUGGG UUUUAGAGCUA GAAAUAGCAAG UUAAAAUAAG GCUAGUCCGUU AUCAACUUGAA AAAGUGGCACC GAGUCGGUGCU UUU	mC*mA*mG*GAUCUCAU AGAGGAUGGGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA AAUAAGGCUAGUCCGU UAUCAmAmCmUmUmGm AmAmAmAmAmGmUmG mGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU *mU*mU*mU	chr7: 142792721-142792741
G016280	653	GGCCAC CCUGUA UGCUGU GC	GGCCACCCUGU AUGCUGUGCGU UUUAGAGCUAG AAAUAGCAAGU	mG*mG*mC*CACCCUGU AUGCUGUGCGUUUUAG AmGmCmUmAmGmAmAm AmUmAmGmCAAGUUAA	chr7: 142792749-142792769

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
			UAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	AAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	
G016279	654	CAACAGUGUCCUACCAGCAA	CAACAGUGUCCUACCAGCAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mC*mA*mA*CAGUGUCCUACCAGCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792685-142792705
G016250	655	AGCUGAGCUGGUGGGUGAAU	AGCUGAGCUGGUGGGUGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mA*mG*mC*UGAGCUGGUGGGUGAAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791816-142791836; chr7: 142801163-142801183
G016278	656	AACAGUGUCCUACCAGCAAG	AACAGUGUCCUACCAGCAAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mA*mA*mC*AGUGUCCUACCAGCAAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792686-142792706
G016259	657	AGGCUUCUUCCCUGACCACG	AGGCUUCUUCCCUGACCACGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mA*mG*mG*CUUCUUCCCUGACCACGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791793-142791813
G016284	658	UCUUCUGCAGGUCAAGAGAA	UCUUCUGCAGGUCAAGAGAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mU*mC*mU*UCUGCAGGUCAAGAGAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142793110-142793130

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
G016249	659	GAGCUGAGCUGGUGGGUGAA	GAGCUGAGCUGGUGGGUGAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU	mG*mA*mG*CUGAGCUGGUGGGUGAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791815-142791835; chr7: 142801162-142801182
G016256	660	GGUCAGCGCCCUUGUGU UGA	GGUCAGCGCCCUUGUGUUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mG*mG*mU*CAGCGCCCUUGUGUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792770-142792790
G016190	661	AGAUCGUCAGCGCCGAGGCC	AGAUCGUCAGCGCCGAGGCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mA*mG*mA*UCGUCAGCGCCGAGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792047-142792067; chr7: 142801394-142801414
G016248	662	GCUGCUCCUUGAGGGGCUGC	GCUGCUCCUUGAGGGGCUGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mG*mC*mU*GCUCCUUGAGGGGCUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791871-142791891; chr7: 142801218-142801238
G016257	663	GUAUCUGGAGUCAUUGAGGG	GUAUCUGGAGUCAUUGAGGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mG*mU*mA*UCUGGAGUCAUUGAGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791894-142791914
G016260	664	AUCCUCUAUGAGAUCCUGCU	AUCCUCUAUGAGAUCCUGCUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAU	mA*mU*mC*CUCUAUGAGAUCCUGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGm	chr7: 142792723-142792743

Guide ID	SEQ ID NO to the Guide Sequence	Guide Sequence	Exemplary Full Sequence (SEQ ID NOS: 1024-1075)	Exemplary Mod Sequence (SEQ ID NOS: 801-852)	Genomic Coordinates (hg38)
			CAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	AmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	
G016262	665	UCCUCUAUGAGAUCCUGCUA	UCCUCUAUGAGAUCCUGCUAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mU*mC*mC*UCUAUGAGAUCCUGCUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792724-142792744
G016242	666	GCAGUAUCUGGAGUCAUUGA	GCAGUAUCUGGAGUCAUUGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mG*mC*mA*GUAUCUGGAGUCAUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142791897-142791917; chr7: 142801244-142801264
G016268	667	GCUGACCAGCACAGCAUACA	GCUGACCAGCACAGCAUACAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mG*mC*mU*GACCAGCACAGCAUACAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792757-142792777
G016265	668	ACAGGGUGGCCUUCCCUAGC	ACAGGGUGGCCUUCCCUAGCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mA*mC*mA*GGGUGGCCUUCCCUAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792740-142792760
G016254	669	CGCUGACCAGCACAGCAUAC	CGCUGACCAGCACAGCAUACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU	mC*mG*mC*UGACCAGCACAGCAUACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU	chr7: 142792758-142792778

[00460] In some embodiments, the guide sequence comprises SEQ ID NO: 618. In some embodiments, the guide sequence comprises SEQ ID NO: 619. In some embodiments, the guide sequence comprises SEQ ID NO: 620. In some embodiments, the guide sequence comprises SEQ ID NO: 621. In some embodiments, the guide sequence comprises SEQ ID NO: 622. In some embodiments, the guide sequence comprises SEQ ID NO: 623. In some embodiments, the guide sequence comprises SEQ ID NO: 624. In some embodiments, the guide sequence comprises SEQ ID NO: 625. In some embodiments, the guide sequence comprises SEQ ID NO: 626. In some embodiments, the guide sequence comprises SEQ ID NO: 627. In some embodiments, the guide sequence comprises SEQ ID NO: 628. In some embodiments, the guide sequence comprises SEQ ID NO: 629. In some embodiments, the guide sequence comprises SEQ ID NO: 630. In some embodiments, the guide sequence comprises SEQ ID NO: 631. In some embodiments, the guide sequence comprises SEQ ID NO: 632. In some embodiments, the guide sequence comprises SEQ ID NO: 633. In some embodiments, the guide sequence comprises SEQ ID NO: 634. In some embodiments, the guide sequence comprises SEQ ID NO: 635. In some embodiments, the guide sequence comprises SEQ ID NO: 636. In some embodiments, the guide sequence comprises SEQ ID NO: 637. In some embodiments, the guide sequence comprises SEQ ID NO: 638. In some embodiments, the guide sequence comprises SEQ ID NO: 639. In some embodiments, the guide sequence comprises SEQ ID NO: 640. In some embodiments, the guide sequence comprises SEQ ID NO: 641. In some embodiments, the guide sequence comprises SEQ ID NO: 642. In some embodiments, the guide sequence comprises SEQ ID NO: 643. In some embodiments, the guide sequence comprises SEQ ID NO: 644. In some embodiments, the guide sequence comprises SEQ ID NO: 645. In some embodiments, the guide sequence comprises SEQ ID NO: 646. In some embodiments, the guide sequence comprises SEQ ID NO: 647. In some embodiments, the guide sequence comprises SEQ ID NO: 648. In some embodiments, the guide sequence comprises SEQ ID NO: 649. In some embodiments, the guide sequence comprises SEQ ID NO: 650. In some embodiments, the guide sequence comprises SEQ ID NO: 651. In some embodiments, the guide sequence comprises SEQ ID NO: 652. In some embodiments, the guide sequence comprises SEQ ID NO: 653. In some embodiments, the guide sequence comprises SEQ ID NO: 654. In some embodiments, the guide sequence comprises SEQ ID NO: 655. In some embodiments, the guide sequence comprises SEQ ID NO: 656. In some embodiments, the guide sequence comprises SEQ ID NO: 657. In some embodiments, the guide sequence comprises SEQ ID

NO: 658. In some embodiments, the guide sequence comprises SEQ ID NO: 659. In some embodiments, the guide sequence comprises SEQ ID NO: 660. In some embodiments, the guide sequence comprises SEQ ID NO: 661. In some embodiments, the guide sequence comprises SEQ ID NO: 662. In some embodiments, the guide sequence comprises SEQ ID NO: 663. In some embodiments, the guide sequence comprises SEQ ID NO: 664. In some embodiments, the guide sequence comprises SEQ ID NO: 665. In some embodiments, the guide sequence comprises SEQ ID NO: 666. In some embodiments, the guide sequence comprises SEQ ID NO: 667. In some embodiments, the guide sequence comprises SEQ ID NO: 668. In some embodiments, the guide sequence comprises SEQ ID NO: 669.

[00461] In some embodiments, the disclosure provides a method of altering a DNA sequence within a TRAC gene, comprising delivering a composition disclosed herein to a cell. The composition may comprise:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

[00462] In some embodiments, the disclosure provides a method of reducing the expression of a TRAC gene, comprising delivering a composition disclosed herein to a cell. The composition may comprise:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

[00463] In some embodiments, the disclosure provides a method of immunotherapy comprising administering a composition disclosed herein to a subject, an autologous cell thereof, and/or an allogeneic cell. The composition may comprise:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

[00464] In some embodiments, a cell altered by the method disclosed herein is proivded. The cell may be altered ex vivo. The cell may be a T cell, a $CD4^+$ or $CD8^+$ cell. The cell may be a mammalian, primate, or human cell. The cell may be used for immunotherapy of a subject.

[00465] In some embodiments, a compositions is provided, comprising: a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v). The composition may optionally further comprise any one of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein.

[00466] In certain embodiments, the composition disclosed herein is used for altering a DNA sequence within the TRAC gene in a cell. In certain embodiments, the composition disclosed herein is used for reducing the expression of the TRAC gene in a cell. In some embodiments, the composition disclosed herein is used for immunotherapy of a subject.

[00467] In some embodiments, the disclosure provides a method of altering a DNA sequence within a TRBC1 and/or TRBC2 gene, comprising delivering a composition disclosed herein to a cell. The composition may comprise:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

[00468] In some embodiments, the disclosure provides a method of reducing the expression of a TRBC1 and/or TRBC2 gene, comprising delivering a composition disclosed herein to a cell. The composition may comprise:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5C; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

[00469] In some embodiments, the disclosure provides a method of immunotherapy comprising administering a composition disclosed herein to a subject, an autologous cell thereof, and/or an allogeneic cell. The composition may comprise:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

[00470] In some embodiments, a cell may be provided, being altered by the method disclosed herein. The cell may be altered ex vivo. The cell is a T cell, a $CD4^+$ or $CD8^+$ cell. The cell may be a mammalian, primate, or human cell. The cell may be used for immunotherapy of a subject.

[00471] In some embodiments, a compositions is provided, comprising: a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v). The composition may optionally further comprise any one of a nucleic acid (e.g., mRNA), polypeptide, composition, or lipid nucleic acid assembly composition disclosed herein.

[00472] In certain embodiments, the composition disclosed herein is used for altering a DNA sequence within the TRBC1 and/or TRBC2 gene in a cell. In certain embodiments, the composition disclosed herein is used for reducing the expression of the TRBC1 and/or TRBC2 gene in a cell. In some embodiments, the composition disclosed herein is used for immunotherapy of a subject.

J. Exemplary DNA Molecules, Vectors, Expression Constructs, Host Cells, and Production Methods

[00473] In certain embodiments, the disclosure provides a DNA molecule comprising a sequence encoding a polypeptide described herein. In some embodiments, the DNA molecule further comprises nucleic acids that do not encode the polypeptide. Nucleic acids that do not encode the polypeptide disclosed herein include, but are not limited to, promoters, enhancers, regulatory sequences, and nucleic acids encoding a gRNA.

[00474] In some embodiments, the DNA molecule further comprises a nucleotide sequence encoding a crRNA, a trRNA, or a crRNA and trRNA. In some embodiments, the nucleotide sequence encoding the crRNA, trRNA, or crRNA and trRNA comprises or consists of a guide sequence flanked by all or a portion of a repeat sequence from a naturally-occurring CRISPR/Cas system. The nucleic acid comprising or consisting of the crRNA, trRNA, or crRNA and trRNA may further comprise a vector sequence wherein

the vector sequence comprises or consists of nucleic acids that are not naturally found together with the crRNA, trRNA, or crRNA and trRNA. In some embodiments, the crRNA and the trRNA are encoded by non-contiguous nucleic acids within one vector. In other embodiments, the crRNA and the trRNA may be encoded by a contiguous nucleic acid. In some embodiments, the crRNA and the trRNA are encoded by opposite strands of a single nucleic acid. In other embodiments, the crRNA and the trRNA are encoded by the same strand of a single nucleic acid.

[00475] In some embodiments, the DNA molecule further comprises a promoter operably linked to the sequence encoding any of the mRNAs encoding the polypeptide described herein. In some embodiments, the DNA molecule is an expression construct suitable for expression in a mammalian cell, e.g., a human cell or a mouse cell, such as a human hepatocyte or a rodent (e.g., mouse) hepatocyte. In some embodiments, the DNA molecule is an expression construct suitable for expression in a cell of a mammalian organ, e.g., a human liver or a rodent (e.g., mouse) liver. In some embodiments, the DNA molecule is a plasmid or an episome. In some embodiments, the DNA molecule is contained in a host cell, such as a bacterium or a cultured eukaryotic cell. Exemplary bacteria include proteobacteria such as E. coli. Exemplary cultured eukaryotic cells include primary hepatocytes, including hepatocytes of rodent (e.g., mouse) or human origin; hepatocyte cell lines, including hepatocytes of rodent (e.g., mouse) or human origin; human cell lines; rodent (e.g., mouse) cell lines; CHO cells; microbial fungi, such as fission or budding yeasts, e.g., Saccharomyces, such as S. cerevisiae; and insect cells.

[00476] In some embodiments, a method of producing an mRNA disclosed herein is provided. In some embodiments, such a method comprises contacting a DNA molecule described herein with an RNA polymerase under conditions permissive for transcription. In some embodiments, the contacting is performed in vitro, e.g., in a cell-free system. In some embodiments, the RNA polymerase is an RNA polymerase of bacteriophage origin, such as T7 RNA polymerase. In some embodiments, NTPs are provided that include at least one modified nucleotide as discussed above. In some embodiments, the NTPs include at least one modified nucleotide as discussed above and do not comprise UTP.

[00477] In some embodiments, an mRNA disclosed herein alone or together with one or more gRNAs, may be comprised within or delivered by a vector system of one or more vectors. In some embodiments, one or more of the vectors, or all of the vectors, may be DNA vectors. In some embodiments, one or more of the vectors, or all of the vectors, may be RNA vectors. In some embodiments, one or more of the vectors, or all of the vectors, may be

circular. In other embodiments, one or more of the vectors, or all of the vectors, may be linear. In some embodiments, one or more of the vectors, or all of the vectors, may be enclosed in a lipid nanoparticle, liposome, non-lipid nanoparticle, or viral capsid. Non-limiting exemplary vectors include plasmids, phagemids, cosmids, artificial chromosomes, minichromosomes, transposons, viral vectors, and expression vectors.

[00478] Non-limiting exemplary viral vectors include adeno-associated virus (AAV) vector, lentivirus vectors, adenovirus vectors, helper dependent adenoviral vectors (HDAd), herpes simplex virus (HSV-1) vectors, bacteriophage T4, baculovirus vectors, and retrovirus vectors. In some embodiments, the viral vector may be an AAV vector. In other embodiments, the viral vector may a lentivirus vector. In some embodiments, the lentivirus may be non-integrating. In some embodiments, the viral vector may be an adenovirus vector. In some embodiments, the adenovirus may be a high-cloning capacity or "gutless" adenovirus, where all coding viral regions apart from the 5' and 3' inverted terminal repeats (ITRs) and the packaging signal ('I') are deleted from the virus to increase its packaging capacity. In yet other embodiments, the viral vector may be an HSV-1 vector. In some embodiments, the HSV-1-based vector is helper dependent, and in other embodiments it is helper independent. For example, an amplicon vector that retains only the packaging sequence requires a helper virus with structural components for packaging, while a 30kb-deleted HSV-1 vector that removes non-essential viral functions does not require helper virus. In additional embodiments, the viral vector may be bacteriophage T4. In some embodiments, the bacteriophage T4 may be able to package any linear or circular DNA or RNA molecules when the head of the virus is emptied. In further embodiments, the viral vector may be a baculovirus vector. In yet further embodiments, the viral vector may be a retrovirus vector. In embodiments using AAV or lentiviral vectors, which have smaller cloning capacity, it may be necessary to use more than one vector to deliver all the components of a vector system as disclosed herein. For example, one AAV vector may contain sequences encoding a Cas protein, while a second AAV vector may contain one or more guide sequences.

[00479] In some embodiments, the vector may be capable of driving expression of one or more coding sequences, such as the coding sequence of an mRNA disclosed herein, in a cell. In some embodiments, the cell may be a prokaryotic cell, such as, e.g., a bacterial cell. In some embodiments, the cell may be a eukaryotic cell, such as, e.g., a yeast, plant, insect, or mammalian cell. In some embodiments, the eukaryotic cell may be a mammalian cell. In some embodiments, the eukaryotic cell may be a rodent cell. In some embodiments,

the eukaryotic cell may be a human cell. Suitable promoters to drive expression in different types of cells are known in the art. In some embodiments, the promoter may be wild type. In other embodiments, the promoter may be modified for more efficient or efficacious expression. In yet other embodiments, the promoter may be truncated yet retain its function. For example, the promoter may have a normal size or a reduced size that is suitable for proper packaging of the vector into a virus.

[00480] In some embodiments, the vector system may comprise one copy of a nucleotide sequence encoding a polypeptide disclosed herein. In other embodiments, the vector system may comprise more than one copy of a nucleotide sequence encoding a polypeptide disclosed herein. In some embodiments, the nucleotide sequence encoding a polypeptide disclosed herein may be operably linked to at least one transcriptional or translational control sequence. In some embodiments, the nucleotide sequence encoding the protein may be operably linked to at least one promoter.

[00481] In some embodiments, the promoter may be constitutive, inducible, or tissue- specific. In some embodiments, the promoter may be a constitutive promoter. Non-limiting exemplary constitutive promoters include cytomegalovirus immediate early promoter (CMV), simian virus (SV40) promoter, adenovirus major late (MLP) promoter, Rous sarcoma virus (RSV) promoter, mouse mammary tumor virus (MMTV) promoter, phosphoglycerate kinase (PGK) promoter, elongation factor-alpha (EF1a) promoter, ubiquitin promoters, actin promoters, tubulin promoters, immunoglobulin promoters, a functional fragment thereof, or a combination of any of the foregoing. In some embodiments, the promoter may be a CMV promoter. In some embodiments, the promoter may be a truncated CMV promoter. In other embodiments, the promoter may be an EF1a promoter. In some embodiments, the promoter may be an inducible promoter. Non-limiting exemplary inducible promoters include those inducible by heat shock, light, chemicals, peptides, metals, steroids, antibiotics, or alcohol. In some embodiments, the inducible promoter may be one that has a low basal (non-induced) expression level, such as, e.g., the Tet-On® promoter (Clontech).

[00482] In some embodiments, the promoter may be a tissue-specific promoter, e.g., a promoter specific for expression in the liver.

[00483] The vector may further comprise a nucleotide sequence encoding at least one gRNA. In some embodiments, the vector comprises one copy of the gRNA. In other embodiments, the vector comprises more than one copy of the gRNA. In embodiments with more than one gRNA, the gRNAs may be non-identical such that they target different target sequences, or may be identical in that they target the same target sequence. In some

embodiments where the vectors comprise more than one gRNA, each gRNA may have other different properties, such as activity or stability within a complex with the polypeptide comprising a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) and an RNA-guided nickase disclosed herein . In some embodiments, the nucleotide sequence encoding the gRNA may be operably linked to at least one transcriptional or translational control sequence, such as a promoter, a 3' UTR, or a 5' UTR. In one embodiment, the promoter may be a tRNA promoter, e.g., tRNALys3, or a tRNA chimera. See Mefferd et al., RNA. 2015 21:1683-9; Scherer et al., Nucleic Acids Res. 2007 35: 2620–2628. In some embodiments, the promoter may be recognized by RNA polymerase III (Pol III). Non-limiting examples of Pol III promoters include U6 and H1 promoters. In some embodiments, the nucleotide sequence encoding the gRNA may be operably linked to a mouse or human U6 promoter. In other embodiments, the nucleotide sequence encoding the gRNA may be operably linked to a mouse or human H1 promoter. In embodiments with more than one gRNA, the promoters used to drive expression may be the same or different. In some embodiments, the nucleotide encoding the crRNA of the gRNA and the nucleotide encoding the trRNA of the gRNA may be provided on the same vector. In some embodiments, the nucleotide encoding the crRNA and the nucleotide encoding the trRNA may be driven by the same promoter. In some embodiments, the crRNA and trRNA may be transcribed into a single transcript. For example, the crRNA and trRNA may be processed from the single transcript to form a double-molecule gRNA. Alternatively, the crRNA and trRNA may be transcribed into a single-molecule gRNA. In other embodiments, the crRNA and the trRNA may be driven by their corresponding promoters on the same vector. In yet other embodiments, the crRNA and the trRNA may be encoded by different vectors.

[00484] In some embodiments, the compositions comprise a vector system, wherein the system comprises more than one vector. In some embodiments, the vector system may comprise one single vector. In other embodiments, the vector system may comprise two vectors. In additional embodiments, the vector system may comprise three vectors. When different gRNAs are used for multiplexing, or when multiple copies of the gRNA are used, the vector system may comprise more than three vectors.

[00485] In some embodiments, the vector system may comprise inducible promoters to start expression only after it is delivered to a target cell. Non-limiting exemplary inducible promoters include those inducible by heat shock, light, chemicals, peptides, metals, steroids, antibiotics, or alcohol. In some embodiments, the inducible promoter may be one

that has a low basal (non-induced) expression level, such as, e.g., the Tet-On® promoter (Clontech).

[00486] In additional embodiments, the vector system may comprise tissue-specific promoters to start expression only after it is delivered into a specific tissue.

[00487] In some embodiments, the vector may be delivered systemically. In some embodiments, the vector may be delivered into the hepatic circulation.

TABLE 5C. SEQUENCE TABLE

[00488] The following sequence table provides a listing of certain sequences disclosed herein. It is understood that if a DNA sequence (comprising Ts) is referenced with respect to an RNA, then Ts should be replaced with Us (which may be modified or unmodified depending on the context), and vice versa. In the following table and throughout, the terms "mA," "mC," "mU," or "mG" are used to denote a nucleotide that has been modified with 2'-O-Me. In the following table, a "*" is used to depict a PS modification. In this application, the terms A*, C*, U*, or G* may be used to denote a nucleotide that is linked to the next (e.g., 3') nucleotide with a PS bond. * = PS linkage; 'm' = 2'-O-Me nucleotide. In the following table, single amino acid letter code is used to provide peptide sequences.

SEQ ID NO	Description	Sequence
1	mRNA encoding BC22n	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACC UGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGG AGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGU GCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCU ACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGG AGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGA UGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCC UGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGACAAGAAGUACU CCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGU UCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGA CCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGG AGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGG ACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGA UCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGC UGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCG CCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCA ACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGC UGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGG CCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU

		CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGC CCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGG AGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGG AGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCA UCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGA UCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCA CCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCA AGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGC UGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCG UGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACA AGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGG AGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGU ACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACA UCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCA AGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCG UGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGG AGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGU ACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACG UGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCA AGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCU UCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCA AGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGA AGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGG GCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGA AGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGA UCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGA CCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGA AGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGG AGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGA AGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACU CCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGC

		AGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGG UGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGG CCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCG ACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGA CCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGACUAGCACCAG CCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGU AGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAA AACGGAAAAAAAAAAAAAGGUAAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAACGAAAAAAAAAAAACG UAAAAAAAAAAAACUCAAAAAAAAAAAAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAGGGAA AAAAAAAAAACGCAAAAAAAAAAAAACACAAAAAAAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAA AAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAAUAGAAAAAAAAAAAAGUUAAAAAAAAA AAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
2	Open reading frame for BC22n	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUC GGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGG GGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUG GUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGG GGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUAC GACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGAC GAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCC GAGUCCGCCACCCCCGAGUCCGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUG AUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAAC CUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUAC ACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUC CACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGAC GAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCC GACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAAC GCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAG CUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCC AACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCC CAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUG CGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGC UACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGAC GGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUC

		CCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAAC CGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUC GCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCC CAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUG UCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAG GACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGC ACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGAC AUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGAC GACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUC CGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAG CACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUG AAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAG AAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCC GUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAG GAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUC GACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAG CGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCAC GUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUC ACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCAC CACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUC GUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCC AAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUG CUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCC AAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACC GUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUG GGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUG AAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCC GCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAG AUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUAC AACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCC CCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACC

		CUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCC CCCAAGAAGAAGCGGAAGGUGUGA
3	Amino acid sequence for BC22n	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDL VPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYD EFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAV ITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
4	mRNA encoding BC22n with Hibit tag	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACC UGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGG AGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGU GCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCU ACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGG AGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGA UGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCC UGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGACAAGAAGUACU CCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGU UCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGA CCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGG AGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGG

		ACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGA UCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGC UGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCG CCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCA ACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGC UGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGG CCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGC CCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGG AGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGG AGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCA UCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGA UCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCA CCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCA AGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGC UGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCG UGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACA AGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGG AGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGU ACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACA UCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCA AGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCG UGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGG AGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGU ACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACG UGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCA AGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCU UCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCA AGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGA AGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGG GCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGA AGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU

		UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGA UCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGA CCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGA AGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGG AGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGA AGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACU CCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGC AGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGG UGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGG CCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCG ACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGA CCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUCCGAGUCCGCCA CCCCCGAGUCCGUGUCCGGCUGGCGGCUGUUCAAGAAGAUCUCCUGACUAGCACCAGCCUCAAGAACACCCGAAUGGA GUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUA AUAAAAAGAAAGUUUCUUCACAUUCUCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAAACGGAAAAAAAAAAAAGGUA AAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAACUCAAAAA AAAAAAAGAUAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAACGCAAAAAAAA AAAACACAAAAAAAAAAAAAUGCAAAAAAAAAAAAAUCGAAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAA CCCAAAAAAAAAAAAAGACAAAAAAAAAAAAAUAGAAAAAAAAAAAAAGUUAAAAAAAAAAAAACUGAAAAAAAAAAAAUUU AAAAAAAAAAAAUCUAG
5	Open reading frame for BC22n with Hibit tag	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUC GGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGG GGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUG GUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGG GGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUAC GACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGAC GAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCC GAGUCCGCCACCCCCGAGUCCGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUG AUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAAC CUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUAC ACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUC CACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGAC GAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCC

		GACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAAC GCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAG CUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCC AACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCC CAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUG CGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGC UACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGAC GGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUC CCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAAC CGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUC GCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCC CAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUG UCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAG GACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGC ACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGAC AUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGAC GACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUC CGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAG CACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUG AAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAG AAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCC GUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAG GAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUC GACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAG CGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCAC GUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUC ACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCAC CACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUC GUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCC AAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUG CUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCC

		AAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACC GUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUG GGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUG AAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCC GCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAG AUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUAC AACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCC CCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACC CUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCC CCCAAGAAGAAGCGGAAGGUGUCCGAGUCCGCCACCCCCGAGUCCGUGUCCGGCUGGCGGCUGUUCAAGAAGAUCUCC UGA
6	Amino acid sequence for BC22n with Hibit tag	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDL VPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYD EFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAV ITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDGGGSPKKKRKVSESATPESVSGWRLFKKIS
7	Not used	
8	Open reading	AUGGACAAGAAGUACAGCAUCGGACUGGACAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAG

	frame for Cas9	GUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGC UUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUAC GCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAG AUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUG

		GUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUC AACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGC GAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUAC UUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAGGAGGAAGCCCGAAGAAGAAGAGAAAG GUCUAG
9	Amino acid sequence for Cas9	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV*

10	Not used	
11	Open reading frame for Cas9	AUGGACAAGAAGUACUCCAUCGGCCUGGACAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAG GUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUG UUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGG AUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCC UUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAG AAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUG GCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGAC AAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCC AAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAG AACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAG GACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUAC GCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUC ACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUG GUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGAC GGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUG GUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUG GGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAG AUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAG UCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGG AUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUG UACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAG GCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUC GAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACC CUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAG CUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGC AAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUG ACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCC GGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCAC AAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGG AUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUG CAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGG CUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACC CGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGG

		CAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAG CUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGAC UCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUG GUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUAC CUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAG GUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCC AACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAAC GGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUG AACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAG CUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUG GUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAG CGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUC AAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAG GGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCC CCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCC GAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAG CCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUAC UUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUC ACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAG GUGUGA
12	Amino acid sequence for Cas9	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS

		PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
13	mRNA encoding BE3	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUG UGUCGUUGCAGGCCUUAUUCGGAUCCGCCACCAUGAGCAGCGAAACAGGACCGGUCGCAGUCGACCCGACACUGAGAA GAAGAAUCGAACCGCACGAAUUCGAAGUCUUCUUCGACCCGAGAGAACUGAGAAAGGAAACAUGCCUGCUGUACGAAA UCAACUGGGGAGGAAGACACAGCAUCUGGAGACACACAAGCCAGAACACAAACAAGCACGUCGAAGUCAACUUCAUCG AAAAGUUCACAACAGAAAGAUACUUCUGCCCGAACACAAGAUGCAGCAUCACAUGGUUCCUGAGCUGGAGCCCGUGCG GAGAAUGCAGCAGAGCAAUCACAGAAUUCCUGAGCAGAUACCCGCACGUCACACUGUUCAUCUACAUCGCAAGACUGU ACCACCACGCAGACCCGAGAAACAGACAGGGACUGAGAGACCUGAUCAGCAGCGGAGUCACAAUCCAGAUCAUGACAG AACAGGAAAGCGGAUACUGCUGGAGAAACUUCGUCAACUACAGCCCGAGCAACGAAGCACACUGGCCGAGAUACCCGC ACCUGUGGGUCAGACUGUACGUCCUGGAACUGUACUGCAUCAUCCUGGGACUGCCGCCGUGCCUGAACAUCCUGAGAA GAAAGCAGCCGCAGCUGACAUUCUUCACAAUCGCACUGCAGAGCUGCCACUACCAGAGACUGCCGCCGCACAUCCUGU GGGCAACAGGACUGAAGAGCGGAAGCGAAACACCGGGAACAAGCGAAAGCGCAACACCGGAAAGCGACAAGAAGUACA GCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGU UCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAA CAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGG AAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAG ACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCU ACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGA UCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGC UGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCG CAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAA ACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGC UGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGG CAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGA GCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGC CGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGG AAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAG AAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAA UCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAA UCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCA CACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACA AGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAA AGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGC UGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCG UCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACA

		AGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAG AAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAU ACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACU UCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACA UCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCA AGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCG UCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAG AAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGU ACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACG UCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACA GAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAA AGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAU UCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAA AGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAA AGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCG GAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAA AGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCU UCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAA UCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGA CAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGA AGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCG AAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAA AGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACA GCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCAC UGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAAC AGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAG UCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGG CAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCG ACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAA CAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAG GAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAA GCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAU ACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGA AGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUG AAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCU UUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAAAAAAAAAAAAAAA

		AAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAU
14	Open reading frame for BE3	AUGAGCAGCGAAACAGGACCGGUCGCAGUCGACCCGACACUGAGAAGAAGAAUCGAACCGCACGAAUUCGAAGUCUUC UUCGACCCGAGAGAACUGAGAAAGGAAACAUGCCUGCUGUACGAAAUCAACUGGGGAGGAAGACACAGCAUCUGGAGA CACACAAGCCAGAACACAAACAAGCACGUCGAAGUCAACUUCAUCGAAAAGUUCACAACAGAAAGAUACUUCUGCCCG AACACAAGAUGCAGCAUCACAUGGUUCCUGAGCUGGAGCCCGUGCGGAGAAUGCAGCAGAGCAAUCACAGAAUUCCUG AGCAGAUACCCGCACGUCACACUGUUCAUCUACAUCGCAAGACUGUACCACCACGCAGACCCGAGAAACAGACAGGGA CUGAGAGACCUGAUCAGCAGCGGAGUCACAAUCCAGAUCAUGACAGAACAGGAAAGCGGAUACUGCUGGAGAAACUUC GUCAACUACAGCCCGAGCAACGAAGCACACUGGCCGAGAUACCCGCACCUGUGGGUCAGACUGUACGUCCUGGAACUG UACUGCAUCAUCCUGGGACUGCCGCCGUGCCUGAACAUCCUGAGAAGAAAGCAGCCGCAGCUGACAUUCUUCACAAUC GCACUGCAGAGCUGCCACUACCAGAGACUGCCGCCGCACAUCCUGUGGGCAACAGGACUGAAGAGCGGAAGCGAAACA CCGGGAACAAGCGAAAGCGCAACACCGGAAAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUC GGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGC AUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCA AGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGAC GACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGA AACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACA GACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGA GACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAA AACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAAC CUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCG AACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGAC AACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUG AGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACAC CACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAG AGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUG GAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGAC AACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUC CUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGA AACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAG GGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAG CACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAG CCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAG CAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAAC GCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGAC AUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCA

		CACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUG AUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAAC UUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGAC AGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUC GACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACA CAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUG AAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUG UACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAG GACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAA GUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUG ACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAG AUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAA GUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUC AACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUG GAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGA AAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAA AUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACA GUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAA AGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUC GACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUC AAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGA UACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGA AUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUG GCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCAC UACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUC CUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACA AACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUC CUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGAC AGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUG AUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAA AGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGC AACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUC
15	Amino acid sequence for BE3	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCP NTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNF VNYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA

		RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPIL EKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARG NSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQ ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKL ESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSV KELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESIL MLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
16	mRNA encoding BE3	GGGUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGCCUUAUUCGGAUCCACCAUGAG CUCAGAGACUGGCCCAGUGGCUGUGGACCCCACAUUGAGACGGCGGAUCGAGCCCCAUGAGUUUGAGGUAUUCUUCGA UCCGAGAGAGCUCCGCAAGGAGACCUGCCUGCUUUACGAAAUUAAUUGGGGGGGCCGGCACUCCAUUUGGCGACAUAC AUCACAGAACACUAACAAGCACGUCGAAGUCAACUUCAUCGAGAAGUUCACGACAGAAAGAUAUUUCUGUCCGAACAC AAGGUGCAGCAUUACCUGGUUUCUCAGCUGGAGCCCAUGCGGCGAAUGUAGUAGGGCCAUCACUGAAUUCCUGUCAAG GUAUCCCCACGUCACUCUGUUUAUUUACAUCGCAAGGCUGUACCACCACGCUGACCCCCGCAAUCGACAAGGCCUGCG GGAUUUGAUCUCUUCAGGUGUGACUAUCCAAAUUAUGACUGAGCAGGAGUCAGGAUACUGCUGGAGAAACUUUGUGAA UUAUAGCCCGAGUAAUGAAGCCCACUGGCCUAGGUAUCCCCAUCUGUGGGUACGACUGUACGUUCUUGAACUGUACUG CAUCAUACUGGGCCUGCCUCCUUGUCUCAACAUUCUGAGAAGGAAGCAGCCACAGCUGACAUUCUUUACCAUCGCUCU UCAGUCUUGUCAUUACCAGCGACUGCCCCCACACAUUCUCUGGGCCACCGGGUUGAAAAGCGGCAGCGAGACUCCGGG CACCUCAGAGUCCGCCACACCCGAAAGUGAUAAGAAGUACUCAAUCGGGCUGGCCAUCGGAACUAAUUCCGUGGGUUG GGCAGUGAUCACGGAUGAAUACAAAGUGCCGUCCAAGAAGUUCAAGGUCCUGGGGAACACCGAUAGACACAGCAUCAA GAAAAAUCUCAUCGGAGCCCUGCUGUUUGACUCCGGCGAAACCGCAGAAGCGACCCGGCUCAAACGUACCGCGAGGCG ACGCUACACCCGGCGGAAGAAUCGCAUCUGCUAUCUGCAAGAGAUCUUUUCGAACGAAAUGGCAAAGGUCGACGACAG CUUCUUCCACCGCCUGGAAGAAUCUUUCCUGGUGGAGGAGGACAAGAAGCAUGAACGGCAUCCUAUCUUUGGAAACAU CGUCGACGAAGUGGCGUACCACGAAAAGUACCCGACCAUCUACCAUCUGCGGAAGAAGUUGGUUGACUCAACUGACAA GGCCGACCUCAGAUUGAUCUACUUGGCCCUCGCCCAUAUGAUCAAAUUCCGCGGACACUUCCUGAUCGAAGGCGAUCU GAACCCUGAUAACUCCGACGUGGAUAAGCUUUUCAUUCAACUGGUGCAGACCUACAACCAACUGUUCGAAGAAAACCC

		AAUCAAUGCUAGCGGCGUCGAUGCCAAGGCCAUCCUGUCCGCCCGGCUGUCGAAGUCGCGGCGCCUCGAAAACCUGAU CGCACAGCUGCCGGGAGAGAAAAAGAACGGACUUUUCGGCAACUUGAUCGCUCUCUCACUGGGACUCACUCCCAAUUU CAAGUCCAAUUUUGACCUGGCCGAGGACGCGAAGCUGCAACUCUCAAAGGACACCUACGACGACGACUUGGACAAUUU GCUGGCACAAAUUGGCGAUCAGUACGCGGAUCUGUUCCUUGCCGCUAAGAACCUUUCGGACGCAAUCUUGCUGUCCGA UAUCCUGCGCGUGAACACCGAAAUAACCAAAGCGCCGCUUAGCGCCUCGAUGAUUAAGCGGUACGACGAGCAUCACCA GGAUCUCACGCUGCUCAAAGCGCUCGUGAGACAGCAACUGCCUGAAAAGUACAAGGAGAUCUUCUUCGACCAGUCCAA GAAUGGGUACGCAGGGUACAUCGAUGGAGGCGCUAGCCAGGAAGAGUUCUAUAAGUUCAUCAAGCCAAUCCUGGAAAA GAUGGACGGAACCGAAGAACUGCUGGUCAAGCUGAACAGGGAGGAUCUGCUCCGGAAACAGAGAACCUUUGACAACGG AUCCAUUCCCCACCAGAUCCAUCUGGGUGAGCUGCACGCCAUCUUGCGGCGCCAGGAGGACUUUUACCCAUUCCUCAA GGACAACCGGGAAAAGAUCGAGAAAAUUCUGACGUUCCGCAUCCCGUAUUACGUGGGCCCACUGGCGCGCGGCAAUUC GCGCUUCGCGUGGAUGACUAGAAAAUCAGAGGAAACCAUCACUCCUUGGAAUUUCGAGGAAGUUGUGGAUAAGGGAGC UUCGGCACAAAGCUUCAUCGAACGAAUGACCAACUUCGACAAGAAUCUCCCAAACGAGAAGGUGCUUCCUAAGCACAG CCUCCUUUACGAAUACUUCACUGUCUACAACGAACUGACUAAAGUGAAAUACGUUACUGAAGGAAUGAGGAAGCCGGC CUUUCUGUCCGGAGAACAGAAGAAAGCAAUUGUCGAUCUGCUGUUCAAGACCAACCGCAAGGUGACCGUCAAGCAGCU UAAAGAGGACUACUUCAAGAAGAUCGAGUGUUUCGACUCAGUGGAAAUCAGCGGGGUGGAGGACAGAUUCAACGCUUC GCUGGGAACCUAUCAUGAUCUCCUGAAGAUCAUCAAGGACAAGGACUUCCUUGACAACGAGGAGAACGAGGACAUCCU GGAAGAUAUCGUCCUGACCUUGACCCUUUUCGAGGAUCGCGAGAUGAUCGAGGAGAGGCUUAAGACCUACGCUCAUCU CUUCGACGAUAAGGUCAUGAAACAACUCAAGCGCCGCCGGUACACUGGUUGGGGCCGCCUCUCCCGCAAGCUGAUCAA CGGUAUUCGCGAUAAACAGAGCGGUAAAACUAUCCUGGAUUUCCUCAAAUCGGAUGGCUUCGCUAAUCGUAACUUCAU GCAAUUGAUCCACGACGACAGCCUGACCUUUAAGGAGGACAUCCAAAAAGCACAAGUGUCCGGACAGGGAGACUCACU CCAUGAACACAUCGCGAAUCUGGCCGGUUCGCCGGCGAUUAAGAAGGGAAUUCUGCAAACUGUGAAGGUGGUCGACGA GCUGGUGAAGGUCAUGGGACGGCACAAACCGGAGAAUAUCGUGAUUGAAAUGGCCCGAGAAAACCAGACUACCCAGAA GGGCCAGAAAAACUCCCGCGAAAGGAUGAAGCGGAUCGAAGAAGGAAUCAAGGAGCUGGGCAGCCAGAUCCUGAAAGA GCACCCGGUGGAAAACACGCAGCUGCAGAACGAGAAGCUCUACCUGUACUAUUUGCAAAAUGGACGGGACAUGUACGU GGACCAAGAGCUGGACAUCAAUCGGUUGUCUGAUUACGACGUGGACCACAUCGUUCCACAGUCCUUUCUGAAGGAUGA CUCGAUCGAUAACAAGGUGUUGACUCGCAGCGACAAGAACAGAGGGAAGUCAGAUAAUGUGCCAUCGGAGGAGGUCGU GAAGAAGAUGAAGAAUUACUGGCGGCAGCUCCUGAAUGCGAAGCUGAUUACCCAGAGAAAGUUUGACAAUCUCACUAA AGCCGAGCGCGGCGGACUCUCAGAGCUGGAUAAGGCUGGAUUCAUCAAACGGCAGCUGGUCGAGACUCGGCAGAUUAC CAAGCACGUGGCGCAGAUCUUGGACUCCCGCAUGAACACUAAAUACGACGAGAACGAUAAGCUCAUCCGGGAAGUGAA GGUGAUUACCCUGAAAAGCAAACUUGUGUCGGACUUUCGGAAGGACUUUCAGUUUUACAAAGUGAGAGAAAUCAACAA CUACCAUCACGCGCAUGACGCAUACCUCAACGCUGUGGUCGGUACCGCCCUGAUCAAAAAGUACCCUAAACUUGAAUC GGAGUUUGUGUACGGAGACUACAAGGUCUACGACGUGAGGAAGAUGAUAGCCAAGUCCGAACAGGAAAUCGGGAAAGC AACUGCGAAAUACUUCUUUUACUCAAACAUCAUGAACUUUUUCAAGACUGAAAUUACGCUGGCCAAUGGAGAAAUCAG GAAGAGGCCACUGAUCGAAACUAACGGAGAAACGGGCGAAAUCGUGUGGGACAAGGGCAGGGACUUCGCAACUGUUCG CAAAGUGCUCUCUAUGCCGCAAGUCAAUAUUGUGAAGAAAACCGAAGUGCAAACCGGCGGAUUUUCAAAGGAAUCGAU CCUCCCAAAGAGAAAUAGCGACAAGCUCAUUGCACGCAAGAAAGACUGGGACCCGAAGAAGUACGGAGGAUUCGAUUC

		GCCGACUGUCGCAUACUCCGUCCUCGUGGUGGCCAAGGUGGAGAAGGGAAAGAGCAAAAAGCUCAAAUCCGUCAAAGA GCUGCUGGGGAUUACCAUCAUGGAACGAUCCUCGUUCGAGAAGAACCCGAUUGAUUUCCUCGAGGCGAAGGGUUACAA GGAGGUGAAGAAGGAUCUGAUCAUCAAACUCCCCAAGUACUCACUGUUCGAACUGGAAAAUGGUCGGAAGCGCAUGCU GGCUUCGGCCGGAGAACUCCAAAAAGGAAAUGAGCUGGCCUUGCCUAGCAAGUACGUCAACUUCCUCUAUCUUGCUUC GCACUACGAAAAACUCAAAGGGUCACCGGAAGAUAACGAACAGAAGCAGCUUUUCGUGGAGCAGCACAAGCAUUAUCU GGAUGAAAUCAUCGAACAAAUCUCCGAGUUUUCAAAGCGCGUGAUCCUCGCCGACGCCAACCUCGACAAAGUCCUGUC GGCCUACAAUAAGCAUAGAGAUAAGCCGAUCAGAGAACAGGCCGAGAACAUUAUCCACUUGUUCACCCUGACUAACCU GGGAGCCCCAGCCGCCUUCAAGUACUUCGAUACUACUAUCGAUCGCAAAAGAUACACGUCCACCAAGGAAGUUCUGGA CGCGACCCUGAUCCACCAAAGCAUCACUGGACUCUACGAAACUAGGAUCGAUCUGUCGCAGCUGGGUGGCGAUUCUGG UGGUUCUACUAAUCUGUCAGAUAUUAUUGAAAAGGAGACCGGUAAGCAACUGGUUAUCCAGGAAUCCAUCCUCAUGCU CCCAGAGGAGGUGGAAGAAGUCAUUGGGAACAAGCCGGAAAGCGAUAUACUCGUGCACACCGCCUACGACGAGAGCAC CGACGAGAAUGUCAUGCUUCUGACUAGCGACGCCCCUGAAUACAAGCCUUGGGCUCUGGUCAUACAGGAUAGCAACGG UGAGAACAAGAUUAAGAUGCUCUCUGGUGGUUCUCCCAAGAAGAAGAGGAAAGUCUAAUAGUCUAGCCAUCACAUUUA AAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGU UGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAA UAAAAAAUGGAAAGAACCUCGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAA AAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAU
17	Open reading frame for BE3	AUGAGCUCAGAGACUGGCCCAGUGGCUGUGGACCCCACAUUGAGACGGCGGAUCGAGCCCCAUGAGUUUGAGGUAUUC UUCGAUCCGAGAGAGCUCCGCAAGGAGACCUGCCUGCUUUACGAAAUUAAUUGGGGGGGCCGGCACUCCAUUUGGCGA CAUACAUCACAGAACACUAACAAGCACGUCGAAGUCAACUUCAUCGAGAAGUUCACGACAGAAAGAUAUUUCUGUCCG AACACAAGGUGCAGCAUUACCUGGUUUCUCAGCUGGAGCCCAUGCGGCGAAUGUAGUAGGGCCAUCACUGAAUUCCUG UCAAGGUAUCCCCACGUCACUCUGUUUAUUUACAUCGCAAGGCUGUACCACCACGCUGACCCCCGCAAUCGACAAGGC CUGCGGGAUUUGAUCUCUUCAGGUGUGACUAUCCAAAUUAUGACUGAGCAGGAGUCAGGAUACUGCUGGAGAAACUUU GUGAAUUAUAGCCCGAGUAAUGAAGCCCACUGGCCUAGGUAUCCCCAUCUGUGGGUACGACUGUACGUUCUUGAACUG UACUGCAUCAUACUGGGCCUGCCUCCUUGUCUCAACAUUCUGAGAAGGAAGCAGCCACAGCUGACAUUCUUUACCAUC GCUCUUCAGUCUUGUCAUUACCAGCGACUGCCCCCACACAUUCUCUGGGCCACCGGGUUGAAAAGCGGCAGCGAGACU CCGGGCACCUCAGAGUCCGCCACACCCGAAAGUGAUAAGAAGUACUCAAUCGGGCUGGCCAUCGGAACUAAUUCCGUG GGUUGGGCAGUGAUCACGGAUGAAUACAAAGUGCCGUCCAAGAAGUUCAAGGUCCUGGGGAACACCGAUAGACACAGC AUCAAGAAAAAUCUCAUCGGAGCCCUGCUGUUUGACUCCGGCGAAACCGCAGAAGCGACCCGGCUCAAACGUACCGCG AGGCGACGCUACACCCGGCGGAAGAAUCGCAUCUGCUAUCUGCAAGAGAUCUUUUCGAACGAAAUGGCAAAGGUCGAC GACAGCUUCUUCCACCGCCUGGAAGAAUCUUUCCUGGUGGAGGAGGACAAGAAGCAUGAACGGCAUCCUAUCUUUGGA AACAUCGUCGACGAAGUGGCGUACCACGAAAAGUACCCGACCAUCUACCAUCUGCGGAAGAAGUUGGUUGACUCAACU GACAAGGCCGACCUCAGAUUGAUCUACUUGGCCCUCGCCCAUAUGAUCAAAUUCCGCGGACACUUCCUGAUCGAAGGC GAUCUGAACCCUGAUAACUCCGACGUGGAUAAGCUUUUCAUUCAACUGGUGCAGACCUACAACCAACUGUUCGAAGAA AACCCAAUCAAUGCUAGCGGCGUCGAUGCCAAGGCCAUCCUGUCCGCCCGGCUGUCGAAGUCGCGGCGCCUCGAAAAC CUGAUCGCACAGCUGCCGGGAGAGAAAAAGAACGGACUUUUCGGCAACUUGAUCGCUCUCUCACUGGGACUCACUCCC

		AAUUUCAAGUCCAAUUUUGACCUGGCCGAGGACGCGAAGCUGCAACUCUCAAAGGACACCUACGACGACGACUUGGAC AAUUUGCUGGCACAAAUUGGCGAUCAGUACGCGGAUCUGUUCCUUGCCGCUAAGAACCUUUCGGACGCAAUCUUGCUG UCCGAUAUCCUGCGCGUGAACACCGAAAUAACCAAAGCGCCGCUUAGCGCCUCGAUGAUUAAGCGGUACGACGAGCAU CACCAGGAUCUCACGCUGCUCAAAGCGCUCGUGAGACAGCAACUGCCUGAAAAGUACAAGGAGAUCUUCUUCGACCAG UCCAAGAAUGGGUACGCAGGGUACAUCGAUGGAGGCGCUAGCCAGGAAGAGUUCUAUAAGUUCAUCAAGCCAAUCCUG GAAAAGAUGGACGGAACCGAAGAACUGCUGGUCAAGCUGAACAGGGAGGAUCUGCUCCGGAAACAGAGAACCUUUGAC AACGGAUCCAUUCCCCACCAGAUCCAUCUGGGUGAGCUGCACGCCAUCUUGCGGCGCCAGGAGGACUUUUACCCAUUC CUCAAGGACAACCGGGAAAAGAUCGAGAAAAUUCUGACGUUCCGCAUCCCGUAUUACGUGGGCCCACUGGCGCGCGGC AAUUCGCGCUUCGCGUGGAUGACUAGAAAAUCAGAGGAAACCAUCACUCCUUGGAAUUUCGAGGAAGUUGUGGAUAAG GGAGCUUCGGCACAAAGCUUCAUCGAACGAAUGACCAACUUCGACAAGAAUCUCCCAAACGAGAAGGUGCUUCCUAAG CACAGCCUCCUUUACGAAUACUUCACUGUCUACAACGAACUGACUAAAGUGAAAUACGUUACUGAAGGAAUGAGGAAG CCGGCCUUUCUGUCCGGAGAACAGAAGAAAGCAAUUGUCGAUCUGCUGUUCAAGACCAACCGCAAGGUGACCGUCAAG CAGCUUAAAGAGGACUACUUCAAGAAGAUCGAGUGUUUCGACUCAGUGGAAAUCAGCGGGGUGGAGGACAGAUUCAAC GCUUCGCUGGGAACCUAUCAUGAUCUCCUGAAGAUCAUCAAGGACAAGGACUUCCUUGACAACGAGGAGAACGAGGAC AUCCUGGAAGAUAUCGUCCUGACCUUGACCCUUUUCGAGGAUCGCGAGAUGAUCGAGGAGAGGCUUAAGACCUACGCU CAUCUCUUCGACGAUAAGGUCAUGAAACAACUCAAGCGCCGCCGGUACACUGGUUGGGGCCGCCUCUCCCGCAAGCUG AUCAACGGUAUUCGCGAUAAACAGAGCGGUAAAACUAUCCUGGAUUUCCUCAAAUCGGAUGGCUUCGCUAAUCGUAAC UUCAUGCAAUUGAUCCACGACGACAGCCUGACCUUUAAGGAGGACAUCCAAAAAGCACAAGUGUCCGGACAGGGAGAC UCACUCCAUGAACACAUCGCGAAUCUGGCCGGUUCGCCGGCGAUUAAGAAGGGAAUUCUGCAAACUGUGAAGGUGGUC GACGAGCUGGUGAAGGUCAUGGGACGGCACAAACCGGAGAAUAUCGUGAUUGAAAUGGCCCGAGAAAACCAGACUACC CAGAAGGGCCAGAAAAACUCCCGCGAAAGGAUGAAGCGGAUCGAAGAAGGAAUCAAGGAGCUGGGCAGCCAGAUCCUG AAAGAGCACCCGGUGGAAAACACGCAGCUGCAGAACGAGAAGCUCUACCUGUACUAUUUGCAAAAUGGACGGGACAUG UACGUGGACCAAGAGCUGGACAUCAAUCGGUUGUCUGAUUACGACGUGGACCACAUCGUUCCACAGUCCUUUCUGAAG GAUGACUCGAUCGAUAACAAGGUGUUGACUCGCAGCGACAAGAACAGAGGGAAGUCAGAUAAUGUGCCAUCGGAGGAG GUCGUGAAGAAGAUGAAGAAUUACUGGCGGCAGCUCCUGAAUGCGAAGCUGAUUACCCAGAGAAAGUUUGACAAUCUC ACUAAAGCCGAGCGCGGCGGACUCUCAGAGCUGGAUAAGGCUGGAUUCAUCAAACGGCAGCUGGUCGAGACUCGGCAG AUUACCAAGCACGUGGCGCAGAUCUUGGACUCCCGCAUGAACACUAAAUACGACGAGAACGAUAAGCUCAUCCGGGAA GUGAAGGUGAUUACCCUGAAAAGCAAACUUGUGUCGGACUUUCGGAAGGACUUUCAGUUUUACAAAGUGAGAGAAAUC AACAACUACCAUCACGCGCAUGACGCAUACCUCAACGCUGUGGUCGGUACCGCCCUGAUCAAAAAGUACCCUAAACUU GAAUCGGAGUUUGUGUACGGAGACUACAAGGUCUACGACGUGAGGAAGAUGAUAGCCAAGUCCGAACAGGAAAUCGGG AAAGCAACUGCGAAAUACUUCUUUUACUCAAACAUCAUGAACUUUUUCAAGACUGAAAUUACGCUGGCCAAUGGAGAA AUCAGGAAGAGGCCACUGAUCGAAACUAACGGAGAAACGGGCGAAAUCGUGUGGGACAAGGGCAGGGACUUCGCAACU GUUCGCAAAGUGCUCUCUAUGCCGCAAGUCAAUAUUGUGAAGAAAACCGAAGUGCAAACCGGCGGAUUUUCAAAGGAA UCGAUCCUCCCAAAGAGAAAUAGCGACAAGCUCAUUGCACGCAAGAAAGACUGGGACCCGAAGAAGUACGGAGGAUUC GAUUCGCCGACUGUCGCAUACUCCGUCCUCGUGGUGGCCAAGGUGGAGAAGGGAAAGAGCAAAAAGCUCAAAUCCGUC AAAGAGCUGCUGGGGAUUACCAUCAUGGAACGAUCCUCGUUCGAGAAGAACCCGAUUGAUUUCCUCGAGGCGAAGGGU

		UACAAGGAGGUGAAGAAGGAUCUGAUCAUCAAACUCCCCAAGUACUCACUGUUCGAACUGGAAAAUGGUCGGAAGCGC AUGCUGGCUUCGGCCGGAGAACUCCAAAAAGGAAAUGAGCUGGCCUUGCCUAGCAAGUACGUCAACUUCCUCUAUCUU GCUUCGCACUACGAAAAACUCAAAGGGUCACCGGAAGAUAACGAACAGAAGCAGCUUUUCGUGGAGCAGCACAAGCAU UAUCUGGAUGAAAUCAUCGAACAAAUCUCCGAGUUUUCAAAGCGCGUGAUCCUCGCCGACGCCAACCUCGACAAAGUC CUGUCGGCCUACAAUAAGCAUAGAGAUAAGCCGAUCAGAGAACAGGCCGAGAACAUUAUCCACUUGUUCACCCUGACU AACCUGGGAGCCCCAGCCGCCUUCAAGUACUUCGAUACUACUAUCGAUCGCAAAAGAUACACGUCCACCAAGGAAGUU CUGGACGCGACCCUGAUCCACCAAAGCAUCACUGGACUCUACGAAACUAGGAUCGAUCUGUCGCAGCUGGGUGGCGAU UCUGGUGGUUCUACUAAUCUGUCAGAUAUUAUUGAAAAGGAGACCGGUAAGCAACUGGUUAUCCAGGAAUCCAUCCUC AUGCUCCCAGAGGAGGUGGAAGAAGUCAUUGGGAACAAGCCGGAAAGCGAUAUACUCGUGCACACCGCCUACGACGAG AGCACCGACGAGAAUGUCAUGCUUCUGACUAGCGACGCCCCUGAAUACAAGCCUUGGGCUCUGGUCAUACAGGAUAGC AACGGUGAGAACAAGAUUAAGAUGCUCUCUGGUGGUUCUCCCAAGAAGAAGAGGAAAGUCUAA
18	Amino acid sequence for BE3	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCP NTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNF VNYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA RRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPIL EKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARG NSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRK PAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQ ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKL ESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSV KELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESIL MLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
19	mRNA encoding BC22	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUG UGUCGUUGCAGGCCUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGC ACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACA

		ACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACG GAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGAGUCACAU GGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACG UCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCC CGUUCCAGCCGUGGGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGG GAAACAGCGGAAGCGAAACACCGGGAACAAGCGAAAGCGCAACACCGGAAAGCGACAAGAAGUACAGCAUCGGACUGG CCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGG GAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAA CAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCA ACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACG AAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAA AGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAG GACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAU ACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCA AGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCAC UGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACA CAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACC UGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGA UCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACA AGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACA AGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGA GAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGAC AGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACG UCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACU UCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGA ACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACG UCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAA ACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCG GAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGG ACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAG AAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGG GAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCG ACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCAC AGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCC UGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGG

		CAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGG AACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACC UGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCG UCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCG ACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACAC AGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGAC AGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAA ACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGU UCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGA UCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAA AGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAA UCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACA AGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGA CAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACC CGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGA GCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCG ACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAAC UGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGU ACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGU UCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAG ACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCA UCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAU ACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUC UGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGG UCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGG UCCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGG CACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGG UCUAAUAGUCUAGACAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGC UUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUG CCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA GCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAU
20	Open reading frame for BC22	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUC GGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGA GGAUUCCUGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUG GUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGG GGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUAC

		GACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAUCAUGACAUACGAC GAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACAGCGGAAGCGAAACACCGGGAACAAGC GAAAGCGCAACACCGGAAAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUC AUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAAC CUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUAC ACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUC CACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGAC GAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGAC CUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCG GACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAAC GCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAG CUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGC AACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCA CAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUG AGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUG ACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGA UACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGAC GGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUC CCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAAC AGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUC GCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCA CAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUG UACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUG AGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAA GACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGA ACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGAC AUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGAC GACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUC AGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUG AUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAA CACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUC AAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAG AAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCG GUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAG GAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUC

		GACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAG AUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAG AGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCAC GUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUC ACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCAC CACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUC GUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCA AAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGA CCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUC CUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCG AAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACA GUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUG GGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUC AAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGC GCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUAC GAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAA AUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUAC AACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCA CCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACA CUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGC ACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAA GAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAA AACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAAC AAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
21	Amino acid sequence for BC22	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDL VPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYD EFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAV ITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE

		HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
22	Not used	
23	Open reading frame for Cas9 with Hibit tag	AUGGACAAGAAGUACUCCAUCGGCCUGGACAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAG GUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUG UUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGG AUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCC UUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAG AAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUG GCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGAC AAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCC AAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAG AACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAG GACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUAC GCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUC ACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUG GUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGAC GGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUG GUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUG GGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAG AUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAG UCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGG AUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUG UACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAG GCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUC GAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACC

		CUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAG CUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGC AAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUG ACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCC GGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCAC AAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGG AUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUG CAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGG CUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACC CGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGG CAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAG CUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGAC UCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUG GUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUAC CUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAG GUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCC AACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAAC GGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUG AACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAG CUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUG GUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAG CGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUC AAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAG GGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCC CCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCC GAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAG CCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUAC UUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUC ACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAG GUGUCCGAGUCCGCCACCCCCGAGUCCGUGUCCGGCUGGCGGCUGUUCAAGAAGAUCUCCUGA
24	Amino acid sequence for Cas9 with Hibit tag	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK

		SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKVSESATPESVSGWRLFKKIS*
25	mRNA encoding UGI	GGGAGACCCAAGCUGGCUAGCUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGCCUU AUUCGGAUCCGCCACCAUGGGACCGAAGAAGAAGAGAAAGGUCGGAGGAGGAAGCACAAACCUGUCGGACAUCAUCGA AAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAUCGAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAA CAAGCCGGAAUCGGACAUCCUGGUCCACACAGCAUACGACGAAUCGACAGACGAAAACGUCAUGCUGCUGACAUCGGA CGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACUCGAACGGAGAAAACAAGAUCAAGAUGCUGUGAUAGUC UAGACAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCU CUUUUUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCU CUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGUCUAG
26	Open reading frame for UGI	AUGGGACCGAAGAAGAAGAGAAAGGUCGGAGGAGGAAGCACAAACCUGUCGGACAUCAUCGAAAAGGAAACAGGAAAG CAGCUGGUCAUCCAGGAAUCGAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAUCGGAC AUCCUGGUCCACACAGCAUACGACGAAUCGACAGACGAAAACGUCAUGCUGCUGACAUCGGACGCACCGGAAUACAAG CCGUGGGCACUGGUCAUCCAGGACUCGAACGGAGAAAACAAGAUCAAGAUGCUGUGA
27	Amino acid sequence for UGI	MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGE NKIKMLSGGSKRTADGSEFESPKKKRKVE
28	mRNA encoding BC22 with 2x UGI	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACC UGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGG AGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGU GCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCU ACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGG AGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGA UGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCC UGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGACAAGAAGUACU

		CCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGU UCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGA CCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGG AGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGG ACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGA UCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGC UGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCG CCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCA ACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGC UGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGG CCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGC CCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGG AGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGG AGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCA UCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGA UCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCA CCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCA AGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGC UGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCG UGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACA AGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGG AGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGU ACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACA UCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCA AGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCG UGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGG AGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGU ACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACG UGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCA AGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCU UCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCA

		AGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGA AGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGG GCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGA AGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGA UCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGA CCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGA AGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGG AGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGA AGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACU CCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGC AGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGG UGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGG CCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCG ACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGA CCCGGAUCGACCUGUCCCAGCUGGGCGGCGACUCCGGCGGCUCCGGCGGCUCCGGCGGCUCCACCAACCUGUCCGACA UCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGA UCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGA CCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGU CCGGCGGCUCCGGCGGCUCCGGCGGCUCCACCAACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGA UCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGC ACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCC UGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCAAGCGGACCGCCGACGGCUCCG AGUUCGAGCCCAAGAAGAAGCGGAAGGUGUGAUAGCUAGCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUA CAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAG UUUCUUCACAUUCUCUCGAGAAAAAAAAAAAAAUGGAAAAAAAAAAAAACGGAAAAAAAAAAAAAGGUAAAAAAAAAAAAU AUAAAAAAAAAAAAACAUAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAAACUCAAAAAAAAAAAAGAUAA AAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAA AAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAA AAAGACAAAAAAAAAAAAUAGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAA UCUAG
29	Open reading frame for BC22 with 2x UGI	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUC GGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGG GGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUG GUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGG

		GGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUAC GACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGAC GAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCC GAGUCCGCCACCCCCGAGUCCGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUG AUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAAC CUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUAC ACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUC CACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGAC GAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCC GACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAAC GCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAG CUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCC AACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCC CAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUG CGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGC UACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGAC GGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUC CCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAAC CGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUC GCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCC CAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUG UCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAG GACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGC ACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGAC AUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGAC GACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUC CGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAG CACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUG AAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAG AAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCC GUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAG

		GAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUC GACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAG CGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCAC GUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUC ACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCAC CACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUC GUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCC AAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUG CUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCC AAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACC GUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUG GGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUG AAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCC GCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAG AUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUAC AACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCC CCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACC CUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACUCCGGCGGCUCC GGCGGCUCCGGCGGCUCCACCAACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCC AUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUAC GACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAG GACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCGGCGGCUCCGGCGGCUCCACCAACCUGUCCGAC AUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUG AUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGCUG ACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUG UCCGGCGGCUCCAAGCGGACCGCCGACGGCUCCGAGUUCGAGCCCAAGAAGAAGCGGAAGGUGUGAUAG
30	Amino acid sequence for BC22 with 2x UGI	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDL VPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYD EFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAV ITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL

		TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILP KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAY DESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEV IGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV
31	mRNA encoding BE4MAX protein	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGAAGCGGACCGCCGACGGCUCCGAGUUCGAGU CCCCCAAGAAGAAGCGGAAGGUGUCCUCCGAGACCGGCCCCGUGGCCGUGGACCCCACCCUGCGGCGGCGGAUCGAGC CCCACGAGUUCGAGGUGUUCUUCGACCCCCGGGAGCUGCGGAAGGAGACCUGCCUGCUGUACGAGAUCAACUGGGGCG GCCGGCACUCCAUCUGGCGGCACACCUCCCAGAACACCAACAAGCACGUGGAGGUGAACUUCAUCGAGAAGUUCACCA CCGAGCGGUACUUCUGCCCCAACACCCGGUGCUCCAUCACCUGGUUCCUGUCCUGGUCCCCCUGCGGCGAGUGCUCCC GGGCCAUCACCGAGUUCCUGUCCCGGUACCCCCACGUGACCCUGUUCAUCUACAUCGCCCGGCUGUACCACCACGCCG ACCCCCGGAACCGGCAGGGCCUGCGGGACCUGAUCUCCUCCGGCGUGACCAUCCAGAUCAUGACCGAGCAGGAGUCCG GCUACUGCUGGCGGAACUUCGUGAACUACUCCCCCUCCAACGAGGCCCACUGGCCCCGGUACCCCCACCUGUGGGUGC GGCUGUACGUGCUGGAGCUGUACUGCAUCAUCCUGGGCCUGCCCCCCUGCCUGAACAUCCUGCGGCGGAAGCAGCCCC AGCUGACCUUCUUCACCAUCGCCCUGCAGUCCUGCCACUACCAGCGGCUGCCCCCCCACAUCCUGUGGGCCACCGGCC UGAAGUCCGGCGGCUCCUCCGGCGGCUCCUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCU CCGGCGGCUCCUCCGGCGGCUCCGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCG UGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGA ACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGU ACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCU UCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGG ACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCG ACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACC CCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCA ACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCC

		AGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGU CCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGG CCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCC UGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACC UGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACG GCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGG ACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCA UCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACA ACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGU UCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCG CCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGC UGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCC UGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGG AGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGG GCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGG ACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCG ACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCA UCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGC UGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACG AGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGG UGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCC AGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACC CCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACC AGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCA UCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGA AGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCG AGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGC ACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGA UCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACC ACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGU UCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCG CCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGC GGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGG UGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGC CCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCA CCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGC

		UGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGG UGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCU CCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACU ACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACG AGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCU ACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCG CCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCA CCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACUCCGGCGGCU CCGGCGGCUCCGGCGGCUCCACCAACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGU CCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCU ACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCC AGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCGGCGGCUCCGGCGGCUCCACCAACCUGUCCG ACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGG UGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGC UGACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGC UGUCCGGCGGCUCCAAGCGGACCGCCGACGGCUCCGAGUUCGAGCCCAAGAAGAAGCGGAAGGUGUGAUAGCUAGCAC CAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAA UGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUCUCGAGAAAAAAAAAAAAUGGAAAAAAA AAAAACGGAAAAAAAAAAAAGGUAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAACGAAAAAAAAAA ACGUAAAAAAAAAAAAACUCAAAAAAAAAAAAAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAGG GAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAA AAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAAUAGAAAAAAAAAAAAGUUAAAAAA AAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
32	Open reading frame for BE4MAX protein	AUGAAGCGGACCGCCGACGGCUCCGAGUUCGAGUCCCCCAAGAAGAAGCGGAAGGUGUCCUCCGAGACCGGCCCCGUG GCCGUGGACCCCACCCUGCGGCGGCGGAUCGAGCCCCACGAGUUCGAGGUGUUCUUCGACCCCCGGGAGCUGCGGAAG GAGACCUGCCUGCUGUACGAGAUCAACUGGGGCGGCCGGCACUCCAUCUGGCGGCACACCUCCCAGAACACCAACAAG CACGUGGAGGUGAACUUCAUCGAGAAGUUCACCACCGAGCGGUACUUCUGCCCCAACACCCGGUGCUCCAUCACCUGG UUCCUGUCCUGGUCCCCCUGCGGCGAGUGCUCCCGGGCCAUCACCGAGUUCCUGUCCCGGUACCCCCACGUGACCCUG UUCAUCUACAUCGCCCGGCUGUACCACCACGCCGACCCCCGGAACCGGCAGGGCCUGCGGGACCUGAUCUCCUCCGGC GUGACCAUCCAGAUCAUGACCGAGCAGGAGUCCGGCUACUGCUGGCGGAACUUCGUGAACUACUCCCCCUCCAACGAG GCCCACUGGCCCCGGUACCCCCACCUGUGGGUGCGGCUGUACGUGCUGGAGCUGUACUGCAUCAUCCUGGGCCUGCCC CCCUGCCUGAACAUCCUGCGGCGGAAGCAGCCCCAGCUGACCUUCUUCACCAUCGCCCUGCAGUCCUGCCACUACCAG CGGCUGCCCCCCCACAUCCUGUGGGCCACCGGCCUGAAGUCCGGCGGCUCCUCCGGCGGCUCCUCCGGCUCCGAGACC CCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCUCCGGCGGCUCCUCCGGCGGCUCCGACAAGAAGUACUCCAUCGGC CUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUG CUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAG

		GCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUC UCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAG CACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUG CGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUC CGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAG ACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUG UCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUC GCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAG GACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAG AACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCC AUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAG UACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUC UACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUG CUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGG CGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUAC UACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGG AACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUG CCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAG UACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAG ACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUC UCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUC CUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUC GAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGC UGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAG UCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAG GCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGC AUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAG AUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUC AAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUAC UACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCAC AUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAG UCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUC ACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAG CGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGAC GAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUC CAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCC

		CUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUC GCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACC GAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGG GACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUG CAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGG GACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGC AAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCC AUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUC GAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCC AAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAG CUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUG GCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAAC AUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAG CGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUC GACCUGUCCCAGCUGGGCGGCGACUCCGGCGGCUCCGGCGGCUCCGGCGGCUCCACCAACCUGUCCGACAUCAUCGAG AAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGAUCGGCAAC AAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGAC GCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGC UCCGGCGGCUCCGGCGGCUCCACCAACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAG UCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCC UACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUC CAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCAAGCGGACCGCCGACGGCUCCGAGUUCGAG CCCAAGAAGAAGCGGAAGGUGUGAUAG
33	Amino acid sequence for BE4MAX protein	MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSG VTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQ RLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQ TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNL PNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG

		ILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLI TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKT EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVIL ADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSD APEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTA YDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV**
34	mRNA sequence encoding UGI	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGACCAACCUGUCCGACAUCAUCGAGAAGGAGA CCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGAUCGGCAACAAGCCCG AGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGACGCCCCCG AGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCAAGC GGACCGCCGACGGCUCCGAGUUCGAGUCCCCCAAGAAGAAGCGGAAGGUGGAGUGAUAGCUAGCACCAGCCUCAAGAA CACCCGAAUGGAGUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCG UAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUCUCGAGAAAAAAAAAAAAAUGGAAAAAAAAAAAACGGAAAA AAAAAAAAGGUAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAACGAAAAAAAAAAAACGUAAAAAAAA AAAACUCAAAAAAAAAAAAAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAA ACGCAAAAAAAAAAAAACACAAAAAAAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAAACG AAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAAUAGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAA AAAAAAAAUUUAAAAAAAAAAAUCUAG
35	Open reading frame for UGI	AUGACCAACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCC GAGGAGGUGGAGGAGGUGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGAC GAGAACGUGAUGCUGCUGACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAG AACAAGAUCAAGAUGCUGUCCGGCGGCUCCAAGCGGACCGCCGACGGCUCCGAGUUCGAGUCCCCCAAGAAGAAGCGG AAGGUGGAGUGA
36	amino acid sequence for recombinant Cas9-NLS	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT

		LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
37	not used	NOT USED
38	not used	NOT USED
39	not used	NOT USED
40	Amino acid sequence of *H. sapiens* APOBEC3A deaminase (A3A) see TABLE 58,BC22	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDL VPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYD EFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGN
41	Amino acid sequence of *R. norvegicus* Apobec1 see TABLE 58,BC27	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCP NTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNF VNYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK
42	exemplary coding sequence for UGI (SEQ ID NO: 43)	ACTAATCTGTCAGATATTATTGAAAAGGAGACCGGTAAGCAACTGGTTATCCAGGAATCCATCCTCATGCTCCCAGAG GAGGTGGAAGAAGTCATTGGGAACAAGCCGGAAAGCGATATACTCGTGCACACCGCCTACGACGAGAGCACCGACGAG AATGTCATGCTTCTGACTAGCGACGCCCCTGAATACAAGCCTTGGGCTCTGGTCATACAGGATAGCAACGGTGAGAAC AAGATTAAGATGCTC
43	exemplary UGI	TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGEN KIKML
44	exemplary coding sequence for XTEN SEQ ID NO. 46	AGCGGCAGCGAGACTCCGGGCACCTCAGAGTCCGCCACACCCGAAAGT

45	exemplary coding sequence for XTEN SEQ ID NO. 46	AGCGGAAGCGAAACACCGGGAACAAGCGAAAGCGCAACACCGGAAAGC
46	exemplary XTEN	SGSETPGTSESATPES
47	exemplary XTEN	SGSETPGTSESA
48	exemplary XTEN	SGSETPGTSESATPEGGSGGS
49	amino acid sequence for exemplary linker	GGGGSEAAAKEAAAK
50	amino acid sequence for exemplary linker	EAAAKGGGGSGGGGS
51	amino acid sequence for exemplary linker	EAAAKEAAAKEAAAK
52	amino acid sequence for exemplary linker	GGGGSGGGGSGGGGSGGGGS
53	amino acid sequence for exemplary linker	GGGGSGGGGSEAAAKEAAAK
54	amino acid sequence for exemplary linker	GGGGSEAAAKGGGGSGGGGS
55	amino acid sequence for exemplary linker	EAAAKEAAAKEAAAKGGGGSGGGGS
56	amino acid sequence for exemplary linker	EAAAKEAAAKEAAAKEAAAK
57	amino acid sequence for exemplary linker	GGGGSEAAAKEAAAKGGGGSEAAAK

58	amino acid sequence for exemplary linker	EAAAKEAAAKGGGGSGGGGSGGGGS
59	amino acid sequence for exemplary linker	EAAAKEAAAKGGGGSGGGGSEAAAK
60	nucleic acid sequence for exemplary linker SSGS	TCTGGTGGTTCT
61	amino acid sequence for exemplary linker SGGS	SGGS
62	nucleic acid sequence for SV40 NLS	CCCAAGAAGAAGAGGAAAGTC
63	amino acid acid sequence for SV40 NLS	PKKKRKV
64	pCl-Neo	TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTT GTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTGATTATTGACTAG TTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAA TGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT GCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGG GCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCAC AGTTAAATTGCTAACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCC TTGCAGAAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACT GGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCT CTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACTATAGGCTAGCCT CGAGAATTCACGCGTGGTACCTCTAGAGTCGACCCGGGCGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGA

		CATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTG TGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTT TCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGATC GATCCGGGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGG ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCC TAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGG GGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTA GTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCC AAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGT TAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTA TTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACCTCT GAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAG TCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCA GGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCC ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCC GAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT GATTCTTCTGACACAACAGTCTCGAACTTAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCACGCAGGTTCTC CGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCC GGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGG CAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGG ACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCA TGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCG AGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAG CCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGC CGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGG ACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTA TCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGA AATGACCGACCAAGCGACGCCCAACCTGCCATCACGATGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTG TTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCCGCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCAT AGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACA GACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGG GCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGG GAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCT GATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTG CGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCAC GAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGA

		TGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCA TACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG AATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGG AGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCA TACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTAC TTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCC TTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGC CAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGA TCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATT TAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTG AGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAA TCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTC CGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCA AGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGT GTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACAC AGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG AAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAA ACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGG CTCGACAGATCT
65	Screening plasmid - invariant sequence	GATAAGAAGTACTCAATCGGGCTGGATATCGGAACTAATTCCGTGGGTTGGGCAGTGATCACGGATGAATACAAAGTG CCGTCCAAGAAGTTCAAGGTCCTGGGGAACACCGATAGACACAGCATCAAGAAAAATCTCATCGGAGCCCTGCTGTTT GACTCCGGCGAAACCGCAGAAGCGACCCGGCTCAAACGTACCGCGAGGCGACGCTACACCCGGCGGAAGAATCGCATC TGCTATCTGCAAGAGATCTTTTCGAACGAAATGGCAAAGGTCGACGACAGCTTCTTCCACCGCCTGGAAGAATCTTTC CTGGTGGAGGAGGACAAGAAGCATGAACGGCATCCTATCTTTGGAAACATCGTCGACGAAGTGGCGTACCACGAAAAG TACCCGACCATCTACCATCTGCGGAAGAAGTTGGTTGACTCAACTGACAAGGCCGACCTCAGATTGATCTACTTGGCC CTCGCCCATATGATCAAATTCCGCGGACACTTCCTGATCGAAGGCGATCTGAACCCTGATAACTCCGACGTGGATAAG CTTTTCATTCAACTGGTGCAGACCTACAACCAACTGTTCGAAGAAAACCCAATCAATGCTAGCGGCGTCGATGCCAAG GCCATCCTGTCCGCCCGGCTGTCGAAGTCGCGGCGCCTCGAAAACCTGATCGCACAGCTGCCGGGAGAGAAAAAGAAC GGACTTTTCGGCAACTTGATCGCTCTCTCACTGGGACTCACTCCCAATTTCAAGTCCAATTTTGACCTGGCCGAGGAC GCGAAGCTGCAACTCTCAAAGGACACCTACGACGACGACTTGGACAATTTGCTGGCACAAATTGGCGATCAGTACGCG GATCTGTTCCTTGCCGCTAAGAACCTTTCGGACGCAATCTTGCTGTCCGATATCCTGCGCGTGAACACCGAAATAACC AAAGCGCCGCTTAGCGCCTCGATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCAAAGCGCTCGTG AGACAGCAACTGCCTGAAAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAATGGGTACGCAGGGTACATCGATGGA GGCGCTAGCCAGGAAGAGTTCTATAAGTTCATCAAGCCAATCCTGGAAAAGATGGACGGAACCGAAGAACTGCTGGTC AAGCTGAACAGGGAGGATCTGCTCCGGAAACAGAGAACCTTTGACAACGGATCCATTCCCCACCAGATCCATCTGGGT

		GAGCTGCACGCCATCTTGCGGCGCCAGGAGGACTTTTACCCATTCCTCAAGGACAACCGGGAAAAGATCGAGAAAATT CTGACGTTCCGCATCCCGTATTACGTGGGCCCACTGGCGCGCGGCAATTCGCGCTTCGCGTGGATGACTAGAAAATCA GAGGAAACCATCACTCCTTGGAATTTCGAGGAAGTTGTGGATAAGGGAGCTTCGGCACAAAGCTTCATCGAACGAATG ACCAACTTCGACAAGAATCTCCCAAACGAGAAGGTGCTTCCTAAGCACAGCCTCCTTTACGAATACTTCACTGTCTAC AACGAACTGACTAAAGTGAAATACGTTACTGAAGGAATGAGGAAGCCGGCCTTTCTGTCCGGAGAACAGAAGAAAGCA ATTGTCGATCTGCTGTTCAAGACCAACCGCAAGGTGACCGTCAAGCAGCTTAAAGAGGACTACTTCAAGAAGATCGAG TGTTTCGACTCAGTGGAAATCAGCGGGGTGGAGGACAGATTCAACGCTTCGCTGGGAACCTATCATGATCTCCTGAAG ATCATCAAGGACAAGGACTTCCTTGACAACGAGGAGAACGAGGACATCCTGGAAGATATCGTCCTGACCTTGACCCTT TTCGAGGATCGCGAGATGATCGAGGAGAGGCTTAAGACCTACGCTCATCTCTTCGACGATAAGGTCATGAAACAACTC AAGCGCCGCCGGTACACTGGTTGGGGCCGCCTCTCCCGCAAGCTGATCAACGGTATTCGCGATAAACAGAGCGGTAAA ACTATCCTGGATTTCCTCAAATCGGATGGCTTCGCTAATCGTAACTTCATGCAATTGATCCACGACGACAGCCTGACC TTTAAGGAGGACATCCAAAAAGCACAAGTGTCCGGACAGGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGT TCGCCGGCGATTAAGAAGGGAATTCTGCAAACTGTGAAGGTGGTCGACGAGCTGGTGAAGGTCATGGGACGGCACAAA CCGGAGAATATCGTGATTGAAATGGCCCGAGAAAACCAGACTACCCAGAAGGGCCAGAAAAACTCCCGCGAAAGGATG AAGCGGATCGAAGAAGGAATCAAGGAGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGTGGAAAACACGCAGCTGCAG AACGAGAAGCTCTACCTGTACTATTTGCAAAATGGACGGGACATGTACGTGGACCAAGAGCTGGACATCAATCGGTTG TCTGATTACGACGTGGACCACATCGTTCCACAGTCCTTTCTGAAGGATGACTCGATCGATAACAAGGTGTTGACTCGC AGCGACAAGAACAGAGGGAAGTCAGATAATGTGCCATCGGAGGAGGTCGTGAAGAAGATGAAGAATTACTGGCGGCAG CTCCTGAATGCGAAGCTGATTACCCAGAGAAAGTTTGACAATCTCACTAAAGCCGAGCGCGGCGGACTCTCAGAGCTG GATAAGGCTGGATTCATCAAACGGCAGCTGGTCGAGACTCGGCAGATTACCAAGCACGTGGCGCAGATCTTGGACTCC CGCATGAACACTAAATACGACGAGAACGATAAGCTCATCCGGGAAGTGAAGGTGATTACCCTGAAAAGCAAACTTGTG TCGGACTTTCGGAAGGACTTTCAGTTTTACAAAGTGAGAGAAATCAACAACTACCATCACGCGCATGACGCATACCTC AACGCTGTGGTCGGTACCGCCCTGATCAAAAAGTACCCTAAACTTGAATCGGAGTTTGTGTACGGAGACTACAAGGTC TACGACGTGAGGAAGATGATAGCCAAGTCCGAACAGGAAATCGGGAAAGCAACTGCGAAATACTTCTTTTACTCAAAC ATCATGAACTTTTTCAAGACTGAAATTACGCTGGCCAATGGAGAAATCAGGAAGAGGCCACTGATCGAAACTAACGGA GAAACGGGCGAAATCGTGTGGGACAAGGGCAGGGACTTCGCAACTGTTCGCAAAGTGCTCTCTATGCCGCAAGTCAAT ATTGTGAAGAAAACCGAAGTGCAAACCGGCGGATTTTCAAAGGAATCGATCCTCCCAAAGAGAAATAGCGACAAGCTC ATTGCACGCAAGAAAGACTGGGACCCGAAGAAGTACGGAGGATTCGATTCGCCGACTGTCGCATACTCCGTCCTCGTG GTGGCCAAGGTGGAGAAGGGAAAGAGCAAAAAGCTCAAATCCGTCAAAGAGCTGCTGGGGATTACCATCATGGAACGA TCCTCGTTCGAGAAGAACCCGATTGATTTCCTCGAGGCGAAGGGTTACAAGGAGGTGAAGAAGGATCTGATCATCAAA CTCCCCAAGTACTCACTGTTCGAACTGGAAAATGGTCGGAAGCGCATGCTGGCTTCGGCCGGAGAACTCCAAAAAGGA AATGAGCTGGCCTTGCCTAGCAAGTACGTCAACTTCCTCTATCTTGCTTCGCACTACGAAAAACTCAAAGGGTCACCG GAAGATAACGAACAGAAGCAGCTTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAATCATCGAACAAATCTCCGAG TTTTCAAAGCGCGTGATCCTCGCCGACGCCAACCTCGACAAAGTCCTGTCGGCCTACAATAAGCATAGAGATAAGCCG ATCAGAGAACAGGCCGAGAACATTATCCACTTGTTCACCCTGACTAACCTGGGAGCCCCAGCCGCCTTCAAGTACTTC GATACTACTATCGATCGCAAAAGATACACGTCCACCAAGGAAGTTCTGGACGCGACCCTGATCCACCAAAGCATCACT

		GGACTCTACGAAACTAGGATCGATCTGTCGCAGCTGGGTGGCGATTCTGGTGGTTCTACTAATCTGTCAGATATTATT GAAAAGGAGACCGGTAAGCAACTGGTTATCCAGGAATCCATCCTCATGCTCCCAGAGGAGGTGGAAGAAGTCATTGGG AACAAGCCGGAAAGCGATATACTCGTGCACACCGCCTACGACGAGAGCACCGACGAGAATGTCATGCTTCTGACTAGC GACGCCCCTGAATACAAGCCTTGGGCTCTGGTCATACAGGATAGCAACGGTGAGAACAAGATTAAGATGCTCTCTGGT GGTTCTCCCAAGAAGAAGAGGAAAGTCTAATAG
66	pUC19	GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCA ACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGC AACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAA CGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGA AGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCG GCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCT TTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGG AAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAA GAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGAT GACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAA GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTG AGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTC AGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCT GCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTCGAGCTC GGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG AAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAG GCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT TCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAA GCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAA

		GCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA CGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTAT CTCAGCGATCTGTCTATTTCGTTCATCCATAGTT
67	U6 promoter	TTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAAC ACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTT AAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACG
68	CMV promoter	ATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCC GCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAG TCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCC TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGG ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATC
69	3' UTR from human albumin gene	CAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUU UUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGU GCUUCAAUUAAUAAAAAAUGGAAAGAA
70	Amino acid sequence of Cas9 nickase (D10A) with 1x NLS as the C-terminal 7 amino acids	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY

		FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
71	Cas9 nickase (D10A) mRNA ORF encoding SEQ ID NO: 70 using minimal uridine codons as listed in Table 3, with start and stop codons	AUGGACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAG GUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGC UUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUAC GCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAG AUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACA

		AGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUC AACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGC GAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUAC UUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAGGAGGAAGCCCGAAGAAGAAGAGAAAG GUCUAG
72	Cas9 nickase (D10A) mRNA coding sequence using minimal uridine codons as listed in Table 3 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUC CUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCA GACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUC

		AAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAA UGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGC AGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC

		GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAGGAGGAAGCCCGAAGAAGAAGAGAAAGGUC
73	Amino acid sequence of Cas9 nickase (without NLS)	MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQV NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
74	Cas9 nickase mRNA ORF encoding SEQ ID NO: 73 using minimal uridine codons as listed in Table 3, with start and stop codons	AUGGACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAG GUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGC UUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUAC GCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGAC

		GGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAG AUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACA AGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUC AACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGC GAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAG

		CCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUAC UUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACUAG
75	Cas9 nickase coding sequence encoding SEQ ID NO: 73 using minimal uridine codons as listed in Table 3 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUC CUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCA GACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUC AAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAA UGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGA

		AGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGC AGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGAC
76	Amino acid sequence of Cas9 nickase with two nuclear localization signals as the C-terminal amino acids	DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKN GLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKS EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLT FKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVN IVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP

		EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGSGSPKKKRKVDGSPKKKRKVDSG
77	Cas9 nickase mRNA ORF encoding SEQ ID NO: 76 using minimal uridine codons as listed in Table 3, with start and stop codons	AUGGACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAG GUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUG UUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGA AUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGC UUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAA AAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUG GCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGAC AAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCA AAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAG AACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAA GACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUAC GCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUC ACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGAC GGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUG GUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUG GGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAG AUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAG AGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGA AUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUC UACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAG GCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUC GAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUG AAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACA CUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAG CUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGA AAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUG ACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCA GGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACAC AAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGA AUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUG CAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGA CUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACA

		AGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGA CAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAA CUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGAC AGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUG GUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUAC CUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAG GUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGC AACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAAC GGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUC AACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAG CUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUG GUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAA AGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUC AAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAG GGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGC CCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGC GAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAG CCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUAC UUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUC ACAGGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAAGCGGAAGCCCGAAGAAGAAGAGAAAG GUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUCGACAGCGGAUAG
78	Cas9 nickase coding sequence encoding SEQ ID NO: 76 using minimal uridine codons as listed in Table 3 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGUACAGCAUCGGACUGGCAAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUC CCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUC GACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUC CUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAG UACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCA CUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAG CUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAG GCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAAC GGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCA GACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACA AAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUC AGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGA GGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUC

		AAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGA GAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGC GAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUG ACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUAC AACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCA AUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAA UGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAG AUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUG AAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAG ACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACA UUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGA AGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAG CCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUG AAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUG AGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGA AGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAG CUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUG GACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGC AGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUC AGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUC UACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAAC AUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGA GAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAAC AUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUG AUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUC GUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAG CUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGA AACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCG GAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAA UUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCG AUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUC

		GACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGAC GGAAGCGGAAGCCCGAAGAAGAAGAGAAAGGUCGACGGAAGCCCGAAGAAGAAGAGAAAGGUCGACAGCGGA
79	Cas9 nickase ORF using low A codons of Table 4, with start and stop codons	ATGGACAAGAAGTACTCCATCGGCCTGGCCATCGGCACCAACTCCGTGGGCTGGGCCGTGATCACCGACGAGTACAAG GTGCCCTCCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACTCCATCAAGAAGAACCTGATCGGCGCCCTGCTG TTCGACTCCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGG ATCTGCTACCTGCAGGAGATCTTCTCCAACGAGATGGCCAAGGTGGACGACTCCTTCTTCCACCGGCTGGAGGAGTCC TTCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAG AAGTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACTCCACCGACAAGGCCGACCTGCGGCTGATCTACCTG GCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACTCCGACGTGGAC AAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCTCCGGCGTGGACGCC AAGGCCATCCTGTCCGCCCGGCTGTCCAAGTCCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAG AACGGCCTGTTCGGCAACCTGATCGCCCTGTCCCTGGGCCTGACCCCCAACTTCAAGTCCAACTTCGACCTGGCCGAG GACGCCAAGCTGCAGCTGTCCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTAC GCCGACCTGTTCCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGTCCGACATCCTGCGGGTGAACACCGAGATC ACCAAGGCCCCCCTGTCCGCCTCCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTG GTGCGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAACGGCTACGCCGGCTACATCGAC GGCGGCGCCTCCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTG GTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCTCCATCCCCCACCAGATCCACCTG GGCGAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAG ATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACTCCCGGTTCGCCTGGATGACCCGGAAG TCCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCTCCGCCCAGTCCTTCATCGAGCGG ATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACTCCCTGCTGTACGAGTACTTCACCGTG TACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGTCCGGCGAGCAGAAGAAG GCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATC GAGTGCTTCGACTCCGTGGAGATCTCCGGCGTGGAGGACCGGTTCAACGCCTCCCTGGGCACCTACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACC CTGTTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAG CTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGTCCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGC AAGACCATCCTGGACTTCCTGAAGTCCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACTCCCTG ACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGTCCGGCCAGGGCGACTCCCTGCACGAGCACATCGCCAACCTGGCC GGCTCCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCAC AAGCCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACTCCCGGGAGCGG ATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCTCCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGG

		CTGTCCGACTACGACGTGGACCACATCGTGCCCCAGTCCTTCCTGAAGGACGACTCCATCGACAACAAGGTGCTGACC CGGTCCGACAAGAACCGGGGCAAGTCCGACAACGTGCCCTCCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGG CAGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGTCCGAG CTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGAC TCCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGTCCAAGCTG GTGTCCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTAC CTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGTCCGAGTTCGTGTACGGCGACTACAAG GTGTACGACGTGCGGAAGATGATCGCCAAGTCCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACTCC AACATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAAC GGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGTCCATGCCCCAGGTG AACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCTCCAAGGAGTCCATCCTGCCCAAGCGGAACTCCGACAAG CTGATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACTCCCCCACCGTGGCCTACTCCGTGCTG GTGGTGGCCAAGGTGGAGAAGGGCAAGTCCAAGAAGCTGAAGTCCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG CGGTCCTCCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATC AAGCTGCCCAAGTACTCCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCTCCGCCGGCGAGCTGCAGAAG GGCAACGAGCTGGCCCTGCCCTCCAAGTACGTGAACTTCCTGTACCTGGCCTCCCACTACGAGAAGCTGAAGGGCTCC CCCGAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCTCC GAGTTCTCCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGTCCGCCTACAACAAGCACCGGGACAAG CCCATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTAC TTCGACACCACCATCGACCGGAAGCGGTACACCTCCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGTCCATC ACCGGCCTGTACGAGACCCGGATCGACCTGTCCCAGCTGGGCGGCGACGGCGGCGGCTCCCCCAAGAAGAAGCGGAAG GTGTGA
80	Cas9 nickase ORF using low A codons of Table 4, with start and stop codons and no NLS	ATGGACAAGAAGTACTCCATCGGCCTGGCCATCGGCACCAACTCCGTGGGCTGGGCCGTGATCACCGACGAGTACAAG GTGCCCTCCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACTCCATCAAGAAGAACCTGATCGGCGCCCTGCTG TTCGACTCCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGG ATCTGCTACCTGCAGGAGATCTTCTCCAACGAGATGGCCAAGGTGGACGACTCCTTCTTCCACCGGCTGGAGGAGTCC TTCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAG AAGTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACTCCACCGACAAGGCCGACCTGCGGCTGATCTACCTG GCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACTCCGACGTGGAC AAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCTCCGGCGTGGACGCC AAGGCCATCCTGTCCGCCCGGCTGTCCAAGTCCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAG AACGGCCTGTTCGGCAACCTGATCGCCCTGTCCCTGGGCCTGACCCCCAACTTCAAGTCCAACTTCGACCTGGCCGAG GACGCCAAGCTGCAGCTGTCCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTAC GCCGACCTGTTCCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGTCCGACATCCTGCGGGTGAACACCGAGATC ACCAAGGCCCCCCTGTCCGCCTCCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTG GTGCGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAACGGCTACGCCGGCTACATCGAC

		GGCGGCGCCTCCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTG GTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCTCCATCCCCCACCAGATCCACCTG GGCGAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAG ATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACTCCCGGTTCGCCTGGATGACCCGGAAG TCCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCTCCGCCCAGTCCTTCATCGAGCGG ATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACTCCCTGCTGTACGAGTACTTCACCGTG TACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGTCCGGCGAGCAGAAGAAG GCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATC GAGTGCTTCGACTCCGTGGAGATCTCCGGCGTGGAGGACCGGTTCAACGCCTCCCTGGGCACCTACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACC CTGTTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAG CTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGTCCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGC AAGACCATCCTGGACTTCCTGAAGTCCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACTCCCTG ACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGTCCGGCCAGGGCGACTCCCTGCACGAGCACATCGCCAACCTGGCC GGCTCCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCAC AAGCCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACTCCCGGGAGCGG ATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCTCCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGG CTGTCCGACTACGACGTGGACCACATCGTGCCCCAGTCCTTCCTGAAGGACGACTCCATCGACAACAAGGTGCTGACC CGGTCCGACAAGAACCGGGGCAAGTCCGACAACGTGCCCTCCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGG CAGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGTCCGAG CTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGAC TCCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGTCCAAGCTG GTGTCCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTAC CTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGTCCGAGTTCGTGTACGGCGACTACAAG GTGTACGACGTGCGGAAGATGATCGCCAAGTCCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACTCC AACATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAAC GGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGTCCATGCCCCAGGTG AACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCTCCAAGGAGTCCATCCTGCCCAAGCGGAACTCCGACAAG CTGATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACTCCCCCACCGTGGCCTACTCCGTGCTG GTGGTGGCCAAGGTGGAGAAGGGCAAGTCCAAGAAGCTGAAGTCCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG CGGTCCTCCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATC AAGCTGCCCAAGTACTCCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCTCCGCCGGCGAGCTGCAGAAG GGCAACGAGCTGGCCCTGCCCTCCAAGTACGTGAACTTCCTGTACCTGGCCTCCCACTACGAGAAGCTGAAGGGCTCC CCCGAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCTCC GAGTTCTCCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGTCCGCCTACAACAAGCACCGGGACAAG

		CCCATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTAC TTCGACACCACCATCGACCGGAAGCGGTACACCTCCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGTCCATC ACCGGCCTGTACGAGACCCGGATCGACCTGTCCCAGCTGGGCGGCGACTGA
81	Cas9 nickase ORF using low A codons of Table 4, with two C-terminal NLS sequences and start and stop codons	ATGGACAAGAAGTACTCCATCGGCCTGGCCATCGGCACCAACTCCGTGGGCTGGGCCGTGATCACCGACGAGTACAAG GTGCCCTCCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACTCCATCAAGAAGAACCTGATCGGCGCCCTGCTG TTCGACTCCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGG ATCTGCTACCTGCAGGAGATCTTCTCCAACGAGATGGCCAAGGTGGACGACTCCTTCTTCCACCGGCTGGAGGAGTCC TTCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAG AAGTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACTCCACCGACAAGGCCGACCTGCGGCTGATCTACCTG GCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACTCCGACGTGGAC AAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCTCCGGCGTGGACGCC AAGGCCATCCTGTCCGCCCGGCTGTCCAAGTCCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAG AACGGCCTGTTCGGCAACCTGATCGCCCTGTCCCTGGGCCTGACCCCCAACTTCAAGTCCAACTTCGACCTGGCCGAG GACGCCAAGCTGCAGCTGTCCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTAC GCCGACCTGTTCCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGTCCGACATCCTGCGGGTGAACACCGAGATC ACCAAGGCCCCCCTGTCCGCCTCCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTG GTGCGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAACGGCTACGCCGGCTACATCGAC GGCGGCGCCTCCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTG GTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCTCCATCCCCCACCAGATCCACCTG GGCGAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAG ATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACTCCCGGTTCGCCTGGATGACCCGGAAG TCCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCTCCGCCCAGTCCTTCATCGAGCGG ATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACTCCCTGCTGTACGAGTACTTCACCGTG TACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGTCCGGCGAGCAGAAGAAG GCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATC GAGTGCTTCGACTCCGTGGAGATCTCCGGCGTGGAGGACCGGTTCAACGCCTCCCTGGGCACCTACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACC CTGTTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAG CTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGTCCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGC AAGACCATCCTGGACTTCCTGAAGTCCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACTCCCTG ACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGTCCGGCCAGGGCGACTCCCTGCACGAGCACATCGCCAACCTGGCC GGCTCCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCAC AAGCCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACTCCCGGGAGCGG ATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCTCCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGG CTGTCCGACTACGACGTGGACCACATCGTGCCCCAGTCCTTCCTGAAGGACGACTCCATCGACAACAAGGTGCTGACC

		CGGTCCGACAAGAACCGGGGCAAGTCCGACAACGTGCCCTCCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGG CAGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGTCCGAG CTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGAC TCCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGTCCAAGCTG GTGTCCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTAC CTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGTCCGAGTTCGTGTACGGCGACTACAAG GTGTACGACGTGCGGAAGATGATCGCCAAGTCCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACTCC AACATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAAC GGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGTCCATGCCCCAGGTG AACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCTCCAAGGAGTCCATCCTGCCCAAGCGGAACTCCGACAAG CTGATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACTCCCCCACCGTGGCCTACTCCGTGCTG GTGGTGGCCAAGGTGGAGAAGGGCAAGTCCAAGAAGCTGAAGTCCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG CGGTCCTCCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATC AAGCTGCCCAAGTACTCCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCTCCGCCGGCGAGCTGCAGAAG GGCAACGAGCTGGCCCTGCCCTCCAAGTACGTGAACTTCCTGTACCTGGCCTCCCACTACGAGAAGCTGAAGGGCTCC CCCGAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCTCC GAGTTCTCCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGTCCGCCTACAACAAGCACCGGGACAAG CCCATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTAC TTCGACACCACCATCGACCGGAAGCGGTACACCTCCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGTCCATC ACCGGCCTGTACGAGACCCGGATCGACCTGTCCCAGCTGGGCGGCGACGGCTCCGGCTCCCCCAAGAAGAAGCGGAAG GTGGACGGCTCCCCCAAGAAGAAGCGGAAGGTGGACTCCGGCTGA
82	Cas9 nickase ORF using low A/U codons of Table 4, with start and stop codons	ATGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCACCGACGAGTACAAG GTGCCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTG TTCGACAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGG ATCTGCTACCTGCAGGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGAGGAGAGC TTCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAG AAGTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTACCTG GCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGAC AAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCC AAGGCCATCCTGAGCGCCCGGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAG AACGGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAG GACGCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTAC GCCGACCTGTTCCTGGCCGCCAAGAACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCGAGATC ACCAAGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTG GTGCGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATCGAC GGCGGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTG

		GTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTG GGCGAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAG ATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAG AGCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGG ATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTG TACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAG GCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATC GAGTGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACC CTGTTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAG CTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGC AAGACCATCCTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACAGCCTG ACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGGCC GGCAGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCAC AAGCCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCCGGGAGCGG ATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGG CTGAGCGACTACGACGTGGACCACATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTGCTGACC CGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGG CAGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGAGCGAG CTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGAC AGCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTG GTGAGCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTAC CTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAG GTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACAGC AACATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAAC GGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTG AACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTGCCCAAGCGGAACAGCGACAAG CTGATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTG GTGGTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG CGGAGCAGCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATC AAGCTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAG GGCAACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGC CCCGAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGC GAGTTCAGCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAG CCCATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTAC

		TTCGACACCACCATCGACCGGAAGCGGTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATC ACCGGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGGCGACGGCGGCGGCAGCCCCAAGAAGAAGCGGAAG GTGTGA
83	Cas9 nickase ORF using low A/U codons of Table 4, with two C-terminal NLS sequences and start and stop codons	ATGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCACCGACGAGTACAAG GTGCCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTG TTCGACAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGG ATCTGCTACCTGCAGGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGAGGAGAGC TTCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAG AAGTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTACCTG GCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGAC AAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCC AAGGCCATCCTGAGCGCCCGGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAG AACGGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAG GACGCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTAC GCCGACCTGTTCCTGGCCGCCAAGAACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCGAGATC ACCAAGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTG GTGCGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATCGAC GGCGGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTG GTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTG GGCGAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAG ATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAG AGCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGG ATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTG TACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAG GCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATC GAGTGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACC CTGTTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAG CTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGC AAGACCATCCTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACAGCCTG ACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGGCC GGCAGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCAC AAGCCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCCGGGAGCGG ATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGG CTGAGCGACTACGACGTGGACCACATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTGCTGACC

		CGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGG CAGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGAGCGAG CTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGAC AGCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTG GTGAGCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTAC CTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAG GTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACAGC AACATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAAC GGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTG AACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTGCCCAAGCGGAACAGCGACAAG CTGATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTG GTGGTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG CGGAGCAGCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATC AAGCTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAG GGCAACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGC CCCGAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGC GAGTTCAGCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAG CCCATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTAC TTCGACACCACCATCGACCGGAAGCGGTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATC ACCGGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGGCGACGGCAGCGGCAGCCCCAAGAAGAAGCGGAAG GTGGACGGCAGCCCCAAGAAGAAGCGGAAGGTGGACAGCGGCTGA
84	Cas9 nickase ORF using low A/U codons of Table 4, with start and stop codons and no NLS	ATGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCACCGACGAGTACAAG GTGCCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTG TTCGACAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGG ATCTGCTACCTGCAGGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGAGGAGAGC TTCCTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAG AAGTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTACCTG GCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGAC AAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCC AAGGCCATCCTGAGCGCCCGGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAG AACGGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAG GACGCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTAC GCCGACCTGTTCCTGGCCGCCAAGAACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCGAGATC ACCAAGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTG GTGCGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATCGAC GGCGGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTG

		GTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTG GGCGAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAG ATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAG AGCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGG ATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTG TACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAG GCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATC GAGTGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACCTGCTG AAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACC CTGTTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAG CTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGC AAGACCATCCTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACAGCCTG ACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGGCC GGCAGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCAC AAGCCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCCGGGAGCGG ATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTG CAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGG CTGAGCGACTACGACGTGGACCACATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTGCTGACC CGGAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGG CAGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGAGCGAG CTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGAC AGCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTG GTGAGCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTAC CTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAG GTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACAGC AACATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAAC GGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTG AACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTGCCCAAGCGGAACAGCGACAAG CTGATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTG GTGGTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAG CGGAGCAGCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATC AAGCTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAG GGCAACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGC CCCGAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGC GAGTTCAGCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAG CCCATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTAC

		TTCGACACCACCATCGACCGGAAGCGGTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATC ACCGGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGGCGACTGA
85	Cas9 nickase ORF using low A codons of Table 4 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGTACTCCATCGGCCTGGCCATCGGCACCAACTCCGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTG CCCTCCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACTCCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTC GACTCCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGGATC TGCTACCTGCAGGAGATCTTCTCCAACGAGATGGCCAAGGTGGACGACTCCTTCTTCCACCGGCTGGAGGAGTCCTTC CTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAG TACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACTCCACCGACAAGGCCGACCTGCGGCTGATCTACCTGGCC CTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACTCCGACGTGGACAAG CTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCTCCGGCGTGGACGCCAAG GCCATCCTGTCCGCCCGGCTGTCCAAGTCCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAAC GGCCTGTTCGGCAACCTGATCGCCCTGTCCCTGGGCCTGACCCCCAACTTCAAGTCCAACTTCGACCTGGCCGAGGAC GCCAAGCTGCAGCTGTCCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCC GACCTGTTCCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGTCCGACATCCTGCGGGTGAACACCGAGATCACC AAGGCCCCCCTGTCCGCCTCCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTGGTG CGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAACGGCTACGCCGGCTACATCGACGGC GGCGCCTCCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTGGTG AAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCTCCATCCCCCACCAGATCCACCTGGGC GAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAGATC CTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACTCCCGGTTCGCCTGGATGACCCGGAAGTCC GAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCTCCGCCCAGTCCTTCATCGAGCGGATG ACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACTCCCTGCTGTACGAGTACTTCACCGTGTAC AACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGTCCGGCGAGCAGAAGAAGGCC ATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATCGAG TGCTTCGACTCCGTGGAGATCTCCGGCGTGGAGGACCGGTTCAACGCCTCCCTGGGCACCTACCACGACCTGCTGAAG ATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACCCTG TTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAGCTG AAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGTCCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAG ACCATCCTGGACTTCCTGAAGTCCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACTCCCTGACC TTCAAGGAGGACATCCAGAAGGCCCAGGTGTCCGGCCAGGGCGACTCCCTGCACGAGCACATCGCCAACCTGGCCGGC TCCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCACAAG CCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACTCCCGGGAGCGGATG AAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCTCCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCAG AACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGGCTG TCCGACTACGACGTGGACCACATCGTGCCCCAGTCCTTCCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCCGG

		TCCGACAAGAACCGGGGCAAGTCCGACAACGTGCCCTCCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGGCAG CTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGTCCGAGCTG GACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGACTCC CGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGTCCAAGCTGGTG TCCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTACCTG AACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGTCCGAGTTCGTGTACGGCGACTACAAGGTG TACGACGTGCGGAAGATGATCGCCAAGTCCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACTCCAAC ATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAACGGC GAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGTCCATGCCCCAGGTGAAC ATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCTCCAAGGAGTCCATCCTGCCCAAGCGGAACTCCGACAAGCTG ATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACTCCCCCACCGTGGCCTACTCCGTGCTGGTG GTGGCCAAGGTGGAGAAGGGCAAGTCCAAGAAGCTGAAGTCCGTGAAGGAGCTGCTGGGCATCACCATCATGGAGCGG TCCTCCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATCAAG CTGCCCAAGTACTCCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCTCCGCCGGCGAGCTGCAGAAGGGC AACGAGCTGGCCCTGCCCTCCAAGTACGTGAACTTCCTGTACCTGGCCTCCCACTACGAGAAGCTGAAGGGCTCCCCC GAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCTCCGAG TTCTCCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGTCCGCCTACAACAAGCACCGGGACAAGCCC ATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTACTTC GACACCACCATCGACCGGAAGCGGTACACCTCCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGTCCATCACC GGCCTGTACGAGACCCGGATCGACCTGTCCCAGCTGGGCGGCGACGGCGGCGGCTCCCCCAAGAAGAAGCGGAAGGTG
86	Cas9 nickase ORF using low A codons of Table 4 (no NLS and no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGTACTCCATCGGCCTGGCCATCGGCACCAACTCCGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTG CCCTCCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACTCCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTC GACTCCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGGATC TGCTACCTGCAGGAGATCTTCTCCAACGAGATGGCCAAGGTGGACGACTCCTTCTTCCACCGGCTGGAGGAGTCCTTC CTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAG TACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACTCCACCGACAAGGCCGACCTGCGGCTGATCTACCTGGCC CTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACTCCGACGTGGACAAG CTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCTCCGGCGTGGACGCCAAG GCCATCCTGTCCGCCCGGCTGTCCAAGTCCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAAC GGCCTGTTCGGCAACCTGATCGCCCTGTCCCTGGGCCTGACCCCCAACTTCAAGTCCAACTTCGACCTGGCCGAGGAC GCCAAGCTGCAGCTGTCCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCC GACCTGTTCCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGTCCGACATCCTGCGGGTGAACACCGAGATCACC AAGGCCCCCCTGTCCGCCTCCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTGGTG CGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAACGGCTACGCCGGCTACATCGACGGC GGCGCCTCCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTGGTG AAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCTCCATCCCCCACCAGATCCACCTGGGC

		GAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAGATC CTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACTCCCGGTTCGCCTGGATGACCCGGAAGTCC GAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCTCCGCCCAGTCCTTCATCGAGCGGATG ACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACTCCCTGCTGTACGAGTACTTCACCGTGTAC AACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGTCCGGCGAGCAGAAGAAGGCC ATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATCGAG TGCTTCGACTCCGTGGAGATCTCCGGCGTGGAGGACCGGTTCAACGCCTCCCTGGGCACCTACCACGACCTGCTGAAG ATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACCCTG TTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAGCTG AAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGTCCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAG ACCATCCTGGACTTCCTGAAGTCCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACTCCCTGACC TTCAAGGAGGACATCCAGAAGGCCCAGGTGTCCGGCCAGGGCGACTCCCTGCACGAGCACATCGCCAACCTGGCCGGC TCCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCACAAG CCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACTCCCGGGAGCGGATG AAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCTCCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCAG AACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGGCTG TCCGACTACGACGTGGACCACATCGTGCCCCAGTCCTTCCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCCGG TCCGACAAGAACCGGGGCAAGTCCGACAACGTGCCCTCCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGGCAG CTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGTCCGAGCTG GACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGACTCC CGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGTCCAAGCTGGTG TCCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTACCTG AACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGTCCGAGTTCGTGTACGGCGACTACAAGGTG TACGACGTGCGGAAGATGATCGCCAAGTCCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACTCCAAC ATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAACGGC GAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGTCCATGCCCCAGGTGAAC ATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCTCCAAGGAGTCCATCCTGCCCAAGCGGAACTCCGACAAGCTG ATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACTCCCCCACCGTGGCCTACTCCGTGCTGGTG GTGGCCAAGGTGGAGAAGGGCAAGTCCAAGAAGCTGAAGTCCGTGAAGGAGCTGCTGGGCATCACCATCATGGAGCGG TCCTCCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATCAAG CTGCCCAAGTACTCCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCTCCGCCGGCGAGCTGCAGAAGGGC AACGAGCTGGCCCTGCCCTCCAAGTACGTGAACTTCCTGTACCTGGCCTCCCACTACGAGAAGCTGAAGGGCTCCCCC GAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCTCCGAG TTCTCCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGTCCGCCTACAACAAGCACCGGGACAAGCCC ATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTACTTC GACACCACCATCGACCGGAAGCGGTACACCTCCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGTCCATCACC

		GGCCTGTACGAGACCCGGATCGACCTGTCCCAGCTGGGCGGCGAC
87	Cas9 nickase ORF using low A codons of Table 4, with two C-terminal NLS sequences (no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGTACTCCATCGGCCTGGCCATCGGCACCAACTCCGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTG CCCTCCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACTCCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTC GACTCCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGGATC TGCTACCTGCAGGAGATCTTCTCCAACGAGATGGCCAAGGTGGACGACTCCTTCTTCCACCGGCTGGAGGAGTCCTTC CTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTAC CACGAGAAGTACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACTCCACCGACAAGGCCGACCTGCGGCTGATC TACCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACTCCGAC GTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCTCCGGCGTG GACGCCAAGGCCATCCTGTCCGCCCGGCTGTCCAAGTCCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAG AAGAAGAACGGCCTGTTCGGCAACCTGATCGCCCTGTCCCTGGGCCTGACCCCCAACTTCAAGTCCAACTTCGACCTG GCCGAGGACGCCAAGCTGCAGCTGTCCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGAC CAGTACGCCGACCTGTTCCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGTCCGACATCCTGCGGGTGAACACC GAGATCACCAAGGCCCCCCTGTCCGCCTCCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAG GCCCTGGTGCGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGTCCAAGAACGGCTACGCCGGCTAC ATCGACGGCGGCGCCTCCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAG CTGCTGGTGAAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCTCCATCCCCCACCAGATC CACCTGGGCGAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATC GAGAAGATCCTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACTCCCGGTTCGCCTGGATGACC CGGAAGTCCGAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCTCCGCCCAGTCCTTCATC GAGCGGATGACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACTCCCTGCTGTACGAGTACTTC ACCGTGTACAACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGTCCGGCGAGCAG AAGAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAG AAGATCGAGTGCTTCGACTCCGTGGAGATCTCCGGCGTGGAGGACCGGTTCAACGCCTCCCTGGGCACCTACCACGAC CTGCTGAAGATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACC CTGACCCTGTTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATG AAGCAGCTGAAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGTCCCGGAAGCTGATCAACGGCATCCGGGACAAGCAG TCCGGCAAGACCATCCTGGACTTCCTGAAGTCCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGAC TCCCTGACCTTCAAGGAGGACATCCAGAAGGCCCAGGTGTCCGGCCAGGGCGACTCCCTGCACGAGCACATCGCCAAC CTGGCCGGCTCCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGC CGGCACAAGCCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACTCCCGG GAGCGGATGAAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCTCCCAGATCCTGAAGGAGCACCCCGTGGAGAACACC CAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATC AACCGGCTGTCCGACTACGACGTGGACCACATCGTGCCCCAGTCCTTCCTGAAGGACGACTCCATCGACAACAAGGTG

		CTGACCCGGTCCGACAAGAACCGGGGCAAGTCCGACAACGTGCCCTCCGAGGAGGTGGTGAAGAAGATGAAGAACTAC TGGCGGCAGCTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTG TCCGAGCTGGACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATC CTGGACTCCCGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGTCC AAGCTGGTGTCCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGAC GCCTACCTGAACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGTCCGAGTTCGTGTACGGCGAC TACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGTCCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTC TACTCCAACATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAG ACCAACGGCGAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGTCCATGCCC CAGGTGAACATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCTCCAAGGAGTCCATCCTGCCCAAGCGGAACTCC GACAAGCTGATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACTCCCCCACCGTGGCCTACTCC GTGCTGGTGGTGGCCAAGGTGGAGAAGGGCAAGTCCAAGAAGCTGAAGTCCGTGAAGGAGCTGCTGGGCATCACCATC ATGGAGCGGTCCTCCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTG ATCATCAAGCTGCCCAAGTACTCCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCTCCGCCGGCGAGCTG CAGAAGGGCAACGAGCTGGCCCTGCCCTCCAAGTACGTGAACTTCCTGTACCTGGCCTCCCACTACGAGAAGCTGAAG GGCTCCCCCGAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAG ATCTCCGAGTTCTCCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGTCCGCCTACAACAAGCACCGG GACAAGCCCATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTC AAGTACTTCGACACCACCATCGACCGGAAGCGGTACACCTCCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAG TCCATCACCGGCCTGTACGAGACCCGGATCGACCTGTCCCAGCTGGGCGGCGACGGCTCCGGCTCCCCCAAGAAGAAG CGGAAGGTGGACGGCTCCCCCAAGAAGAAGCGGAAGGTGGACTCCGGC
88	Cas9 nickase ORF using low A/U codons of Table 4 (no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTG CCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTC GACAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGGATC TGCTACCTGCAGGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGAGGAGAGCTTC CTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAG TACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTACCTGGCC CTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAG CTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCCAAG GCCATCCTGAGCGCCCGGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAAC GGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGAC GCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCC GACCTGTTCCTGGCCGCCAAGAACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCGAGATCACC AAGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTGGTG CGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATCGACGGC GGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTGGTG

		AAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGC GAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAGATC CTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAGAGC GAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATG ACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTAC AACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAGGCC ATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATCGAG TGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACCTGCTGAAG ATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACCCTG TTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAGCTG AAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGCAAG ACCATCCTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACAGCCTGACC TTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGGCCGGC AGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCACAAG CCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCCGGGAGCGGATG AAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCAG AACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGGCTG AGCGACTACGACGTGGACCACATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTGCTGACCCGG AGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGGCAG CTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGAGCGAGCTG GACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGACAGC CGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTGGTG AGCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTACCTG AACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAGGTG TACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACAGCAAC ATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAACGGC GAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTGAAC ATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTGCCCAAGCGGAACAGCGACAAGCTG ATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTGGTG GTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAGCGG AGCAGCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATCAAG CTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAGGGC AACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGCCCC GAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAG TTCAGCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAGCCC ATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTACTTC

		GACACCACCATCGACCGGAAGCGGTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACC GGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGGCGACGGCGGCGGCAGCCCCAAGAAGAAGCGGAAGGTG
89	Cas9 nickase ORF using low A/U codons of Table 4, with two C-terminal NLS sequences (no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTG CCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTC GACAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGGATC TGCTACCTGCAGGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGAGGAGAGCTTC CTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAG TACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTACCTGGCC CTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAG CTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCCAAG GCCATCCTGAGCGCCCGGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAAC GGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGAC GCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCC GACCTGTTCCTGGCCGCCAAGAACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCGAGATCACC AAGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTGGTG CGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATCGACGGC GGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTGGTG AAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGC GAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAGATC CTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAGAGC GAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATG ACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTAC AACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAGGCC ATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATCGAG TGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACCTGCTGAAG ATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACCCTG TTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAGCTG AAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGCAAG ACCATCCTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACAGCCTGACC TTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGGCCGGC AGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCACAAG CCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCCGGGAGCGGATG AAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCAG AACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGGCTG AGCGACTACGACGTGGACCACATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTGCTGACCCGG

		AGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGGCAG CTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGAGCGAGCTG GACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGACAGC CGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTGGTG AGCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTACCTG AACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAGGTG TACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACAGCAAC ATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAACGGC GAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTGAAC ATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTGCCCAAGCGGAACAGCGACAAGCTG ATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTGGTG GTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAGCGG AGCAGCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATCAAG CTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAGGGC AACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGCCCC GAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAG TTCAGCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAGCCC ATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTACTTC GACACCACCATCGACCGGAAGCGGTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACC GGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGGCGACGGCAGCGGCAGCCCCAAGAAGAAGCGGAAGGTG GACGGCAGCCCCAAGAAGAAGCGGAAGGTGGACAGCGGC
90	Cas9 nickase ORF using low A/U codons of Table 4 (no NLS and no start or stop codons; suitable for inclusion in fusion protein coding sequence)	GACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACAGCGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTG CCCAGCAAGAAGTTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGCGCCCTGCTGTTC GACAGCGGCGAGACCGCCGAGGCCACCCGGCTGAAGCGGACCGCCCGGCGGCGGTACACCCGGCGGAAGAACCGGATC TGCTACCTGCAGGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACCGGCTGGAGGAGAGCTTC CTGGTGGAGGAGGACAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAG TACCCCACCATCTACCACCTGCGGAAGAAGCTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTACCTGGCC CTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAG CTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAGAACCCCATCAACGCCAGCGGCGTGGACGCCAAG GCCATCCTGAGCGCCCGGCTGAGCAAGAGCCGGCGGCTGGAGAACCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAAC GGCCTGTTCGGCAACCTGATCGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGAC GCCAAGCTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCC GACCTGTTCCTGGCCGCCAAGAACCTGAGCGACGCCATCCTGCTGAGCGACATCCTGCGGGTGAACACCGAGATCACC AAGGCCCCCCTGAGCGCCAGCATGATCAAGCGGTACGACGAGCACCACCAGGACCTGACCCTGCTGAAGGCCCTGGTG CGGCAGCAGCTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATCGACGGC GGCGCCAGCCAGGAGGAGTTCTACAAGTTCATCAAGCCCATCCTGGAGAAGATGGACGGCACCGAGGAGCTGCTGGTG

		AAGCTGAACCGGGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGC GAGCTGCACGCCATCCTGCGGCGGCAGGAGGACTTCTACCCCTTCCTGAAGGACAACCGGGAGAAGATCGAGAAGATC CTGACCTTCCGGATCCCCTACTACGTGGGCCCCCTGGCCCGGGGCAACAGCCGGTTCGCCTGGATGACCCGGAAGAGC GAGGAGACCATCACCCCCTGGAACTTCGAGGAGGTGGTGGACAAGGGCGCCAGCGCCCAGAGCTTCATCGAGCGGATG ACCAACTTCGACAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTAC AACGAGCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGGAAGCCCGCCTTCCTGAGCGGCGAGCAGAAGAAGGCC ATCGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTGAAGCAGCTGAAGGAGGACTACTTCAAGAAGATCGAG TGCTTCGACAGCGTGGAGATCAGCGGCGTGGAGGACCGGTTCAACGCCAGCCTGGGCACCTACCACGACCTGCTGAAG ATCATCAAGGACAAGGACTTCCTGGACAACGAGGAGAACGAGGACATCCTGGAGGACATCGTGCTGACCCTGACCCTG TTCGAGGACCGGGAGATGATCGAGGAGCGGCTGAAGACCTACGCCCACCTGTTCGACGACAAGGTGATGAAGCAGCTG AAGCGGCGGCGGTACACCGGCTGGGGCCGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGAGCGGCAAG ACCATCCTGGACTTCCTGAAGAGCGACGGCTTCGCCAACCGGAACTTCATGCAGCTGATCCACGACGACAGCCTGACC TTCAAGGAGGACATCCAGAAGGCCCAGGTGAGCGGCCAGGGCGACAGCCTGCACGAGCACATCGCCAACCTGGCCGGC AGCCCCGCCATCAAGAAGGGCATCCTGCAGACCGTGAAGGTGGTGGACGAGCTGGTGAAGGTGATGGGCCGGCACAAG CCCGAGAACATCGTGATCGAGATGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACAGCCGGGAGCGGATG AAGCGGATCGAGGAGGGCATCAAGGAGCTGGGCAGCCAGATCCTGAAGGAGCACCCCGTGGAGAACACCCAGCTGCAG AACGAGAAGCTGTACCTGTACTACCTGCAGAACGGCCGGGACATGTACGTGGACCAGGAGCTGGACATCAACCGGCTG AGCGACTACGACGTGGACCACATCGTGCCCCAGAGCTTCCTGAAGGACGACAGCATCGACAACAAGGTGCTGACCCGG AGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCAGCGAGGAGGTGGTGAAGAAGATGAAGAACTACTGGCGGCAG CTGCTGAACGCCAAGCTGATCACCCAGCGGAAGTTCGACAACCTGACCAAGGCCGAGCGGGGCGGCCTGAGCGAGCTG GACAAGGCCGGCTTCATCAAGCGGCAGCTGGTGGAGACCCGGCAGATCACCAAGCACGTGGCCCAGATCCTGGACAGC CGGATGAACACCAAGTACGACGAGAACGACAAGCTGATCCGGGAGGTGAAGGTGATCACCCTGAAGAGCAAGCTGGTG AGCGACTTCCGGAAGGACTTCCAGTTCTACAAGGTGCGGGAGATCAACAACTACCACCACGCCCACGACGCCTACCTG AACGCCGTGGTGGGCACCGCCCTGATCAAGAAGTACCCCAAGCTGGAGAGCGAGTTCGTGTACGGCGACTACAAGGTG TACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAGATCGGCAAGGCCACCGCCAAGTACTTCTTCTACAGCAAC ATCATGAACTTCTTCAAGACCGAGATCACCCTGGCCAACGGCGAGATCCGGAAGCGGCCCCTGATCGAGACCAACGGC GAGACCGGCGAGATCGTGTGGGACAAGGGCCGGGACTTCGCCACCGTGCGGAAGGTGCTGAGCATGCCCCAGGTGAAC ATCGTGAAGAAGACCGAGGTGCAGACCGGCGGCTTCAGCAAGGAGAGCATCCTGCCCAAGCGGAACAGCGACAAGCTG ATCGCCCGGAAGAAGGACTGGGACCCCAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTACAGCGTGCTGGTG GTGGCCAAGGTGGAGAAGGGCAAGAGCAAGAAGCTGAAGAGCGTGAAGGAGCTGCTGGGCATCACCATCATGGAGCGG AGCAGCTTCGAGAAGAACCCCATCGACTTCCTGGAGGCCAAGGGCTACAAGGAGGTGAAGAAGGACCTGATCATCAAG CTGCCCAAGTACAGCCTGTTCGAGCTGGAGAACGGCCGGAAGCGGATGCTGGCCAGCGCCGGCGAGCTGCAGAAGGGC AACGAGCTGGCCCTGCCCAGCAAGTACGTGAACTTCCTGTACCTGGCCAGCCACTACGAGAAGCTGAAGGGCAGCCCC GAGGACAACGAGCAGAAGCAGCTGTTCGTGGAGCAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAG TTCAGCAAGCGGGTGATCCTGGCCGACGCCAACCTGGACAAGGTGCTGAGCGCCTACAACAAGCACCGGGACAAGCCC ATCCGGGAGCAGGCCGAGAACATCATCCACCTGTTCACCCTGACCAACCTGGGCGCCCCCGCCGCCTTCAAGTACTTC

		GACACCACCATCGACCGGAAGCGGTACACCAGCACCAAGGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACC GGCCTGTACGAGACCCGGATCGACCTGAGCCAGCTGGGCGGCGAC
91	Exemplary 5′ UTR	ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACC
92	Exemplary 5′ UTR	CATAAACCCTGGCGCGCTCGCGGCCCGGCACTCTTCTGGTCCCCACAGACTCAGAGAGAACCCACC
93	Exemplary 5′ UTR	AAGCTCAGAATAAACGCTCAACTTTGGCC
94	Exemplary 5′ UTR	CAGGGTCCTGTGGACAGCTCACCAGCT
95	Exemplary 5′ UTR	TCCCGCAGTCGGCGTCCAGCGGCTCTGCTTGTTCGTGTGTGTGTCGTTGCAGGCCTTATTC
96	Exemplary 5′ UTR	CAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCAT
97	Exemplary 5′ UTR	AGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGG
98	Exemplary 5′ UTR	TGCATTGGAACGCGGATTCCCCGTGCCAAGAGTGACTCACCG
99	Exemplary 3′ UTR	GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTAT GAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGC
100	Exemplary 3′ UTR	GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCCAGCCCCTCCTCCCCTTCCTGCACCCGTACCCC CGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC
101	Exemplary 3′ UTR	ACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAA AATGTAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCACATTCT
102	Exemplary 3′ UTR	TTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTA ATAAAATGAGGAAATTGCATCGCA
103	Exemplary 3′ UTR	GCTGCCTTCTGCGGGGCTTGCCTTCTGGCCATGCCCTTCTTCTCTCCCTTGCACCTGTACCTCTTGGTCTTTGAATAA AGCCTGAGTAGGAAG
104	Exemplary 3′ UTR	CTGGTACTGCATGCACGCAATGCTAGCTGCCCCTTTCCCGTCCTGGGTACCCCGAGTCTCCCCCGACCTCGGGTCCCA GGTATGCTCCCACCTCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGC TCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAA GCTATACTAACCCCAGGGTTGGTCAATTTCGTGCCAGCCACACC
105	Exemplary 3′ UTR	CAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCA ATAAACGAAAGTTTAACTAAGCTATACTAACCCCAGGGTTGGTCAATTTCGTGCCAGCCACACCCTGGTACTGCATGC ACGCAATGCTAGCTGCCCCTTTCCCGTCCTGGGTACCCCGAGTCTCCCCCGACCTCGGGTCCCAGGTATGCTCCCACC TCCACCTGCCCCACTCACCACCTCTGCTAGTTCCAGACACCTCC
106	Exemplary 3′ UTR	ACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAA AATGTAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCACATTCTACCAGCCTCAAGAACACCCGAATGGA GTCTCTAAGCTACATAATACCAACTTACACTTTACAAAATGTTGTCCCCCAAAATGTAGCCATTCGTATCTGCTCCTA ATAAAAAGAAAGTTTCTTCACATTCT

107	Exemplary Kozak sequence	GCCRCCAUGG
108	Exemplary Kozak sequence	GCCGCCRCCAUGG
109	Exemplary poly-A sequence	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAA
110	Exemplary NLS 1	LAAKRSRTT
111	Exemplary NLS 2	QAAKRSRTT
112	Exemplary NLS 3	PAPAKRERTT
113	Exemplary NLS 4	QAAKRPRTT
114	Exemplary NLS 5	RAAKRPRTT
115	Exemplary NLS 6	AAAKRSWSMAA
116	Exemplary NLS 7	AAAKRVWSMAF
117	Exemplary NLS 8	AAAKRSWSMAF
118	Exemplary NLS 9	AAAKRKYFAA
119	Exemplary NLS 10	RAAKRKAFAA
120	Exemplary NLS 11	RAAKRKYFAV
121	Alternate SV40 NLS	PKKKRRV
122	Nucleoplasmin NLS	KRPAATKKAGQAKKKK
123	Exemplary coding sequence for SV40 NLS	CCGAAGAAGAAGAGAAAGGTC
124	Exemplary coding sequence for NLS1	CTGGCAGCAAAGAGAAGCAGAACAACA
125	Exemplary coding sequence for NLS2	CAGGCAGCAAAGAGAAGCAGAACAACA
126	Exemplary coding sequence for	CCGGCACCGGCAAAGAGAGAAAGAACAACA

	NLS3	
127	Exemplary coding sequence for NLS4	CAGGCAGCAAAGAGACCGAGAACAACA
128	Exemplary coding sequence for NLS5	AGAGCAGCAAAGAGACCGAGAACAACA
129	Exemplary coding sequence for NLS6	GCAGCAGCAAAGAGAAGCTGGAGCATGGCAGCA
130	Exemplary coding sequence for NLS7	GCAGCAGCAAAGAGAGTCTGGAGCATGGCATTC
131	Exemplary coding sequence for NLS8	GCAGCAGCAAAGAGAAGCTGGAGCATGGCATTC
132	Exemplary coding sequence for NLS9	GCAGCAGCAAAGAGAAAGTACTTCGCAGCA
133	Exemplary coding sequence for NLS10	AGAGCAGCAAAGAGAAAGGCATTCGCAGCA
134	Exemplary coding sequence for NLS11	AGAGCAGCAAAGAGAAAGTACTTCGCAGTC
135	Exemplary coding sequence for alternate SV40 NLS	CCGAAGAAGAAGAGAAGAGTC
136-138	not used	NOT USED
139	exemplary nucleotide sequence following the 3′ end of the guide	GUUUUAGAGCUAUGCUGUUUUG

	sequence to form a crRNA	
140	Conserved Portion of a spyCas9 sgRNA	GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
141	Modified sgRNA pattern, where N are nucleotides encoding a guide sequence	mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
142	exemplary guide constant region modification pattern(G282-C)	GUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
143	exemplary guide modification pattern (G282-mN3Nx)	mN*mN*mN*(N)xGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
144	exemplary guide modification pattern (G282-Nx)	(N)xGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
145	exemplary guide modification pattern (G282-N20)	mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
146-150		NOT USED
151	exemplary guide sequence	UAAGGCCAGUGGAAAGAAUU
152	exemplary guide sequence for B2M gene	UUACCCCACUUAACUAUCUU
153	exemplary guide sequence for TTR gene	UUACAGCCACGUCUACAGCA

154	exemplary guide sequence for TRAC gene	UUCAAAACCUGUCAGUGAUU
155	exemplary guide sequence for TRBC1/2 gene	CGCUGUCAAGUCCAGUUCUA
156	exemplary guide sequence	CCUUCCGAAAGAGGCCCCCC
157	exemplary guide sequence for SERPINA1 gene	UCCCUGGCUGAGGAUCCCCA
158	exemplary guide sequence for SERPINA1 gene	ACUCACGAUGAAAUCCUGGA
159	exemplary guide sequence	CCCCCCGCCGUGUUUGUGGG
160	exemplary guide sequence for CIITA gene	GAGCCCCCCACUGUGGUGAC
161	exemplary target sequence for CIITA gene	ACCGGCUCUGCAAAGGCCAG
162	exemplary target sequence for CIITA gene	CACCGGCUCUGCAAAGGCCA
163	exemplary target sequence for CIITA gene	CCACCGGCUCUGCAAAGGCC
164	exemplary target sequence for CIITA gene	CUGCUCCACCGGCUCUGCAA
165	exemplary target sequence for CIITA gene	CUGUGUCACCCGUUUCAGGU
166	exemplary target	UGUGUCACCCGUUUCAGGUG

	sequence for CIITA gene	
167	exemplary target sequence for CIITA gene	ACCCGUUUCAGGUGGGGUGA
168	exemplary target sequence for CIITA gene	CCCGUUUCAGGUGGGGUGAG
169	exemplary target sequence for CIITA gene	UGUGCAGACUCAGAGGUGAG
170	exemplary target sequence for CIITA gene	CAGCGCAUCCAGGCUGCAGG
171	exemplary target sequence for CIITA gene	GCGUCCACAUCCUGCAAGGG
172	exemplary target sequence for CIITA gene	GGCGUCCACAUCCUGCAAGG
173	exemplary target sequence for CIITA gene	UGGGCGUCCACAUCCUGCAA
174-176	not used	
177	G013009 guide RNA targeting TRAC	mU*mA*mG*GCAGACAGACUUGUCACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
178	G016016 guide RNA targeting TRAC	mU*mU*mU*CAAAACCUGUCAGUGAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
179	G015991 guide RNA targeting B2M	mA*mC*mU*CACGCUGGAUAGCCUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
180	G015996 guide RNA targeting	mC*mU*mU*ACCCCACUUAACUAUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	B2M	
181	G000297 guide RNA	mU*mA*mA*GGCCAGUGGAAAGAAUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGCmU*mU*mU*U
182	G015995 guide RNA	mU*mU*mA*CCCCACUUAACUAUCUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
183	G000282 guide RNA	mU*mU*mA*CAGCCACGUCUACAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
184	G016017 guide RNA targeting TRAC with guide sequence SEQ ID NO: 154	mU*mU*mC*AAAACCUGUCAGUGAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
185	G016206 guide RNA targeting TRBC1/2 with guide sequence SEQ ID NO: 155	mC*mG*mC*UGUCAAGUCCAGUUCUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
186	SG000296 guide RNA	CCUUCCGAAAGAGGCCCCCCGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUUUU
187	SG001373 guide RNA	UCCCUGGCUGAGGAUCCCCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUUUU
188	SG001400 guide RNA	ACUCACGAUGAAAUCCUGGAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUUUU
189	SG005883 guide RNA	CCCCCCGCCGUGUUUGUGGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUUUU
190	SG003018 guide RNA targeting CIITA with guide sequence SEQ ID NO: 160	GAGCCCCCCACUGUGGUGACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUUUU
191	G018075 guide RNA targeting CIITA with guide sequence SEQ ID NO: 161	mA*mC*mC*GGCUCUGCAAAGGCCAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

192	G018076 guide RNA targeting CIITA with guide sequence SEQ ID NO: 162	mC*mA*mC*CGGCUCUGCAAAGGCCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
193	G018077 guide RNA targeting CIITA with guide sequence SEQ ID NO: 163	mC*mC*mA*CCGGCUCUGCAAAGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
194	G018078 guide RNA targeting CIITA with guide sequence SEQ ID NO: 164	mC*mU*mG*CUCCACCGGCUCUGCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
195	G018081 guide RNA targeting CIITA with guide sequence SEQ ID NO: 165	mC*mU*mG*UGUCACCCGUUUCAGGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
196	G018082 guide RNA targeting CIITA with guide sequence SEQ ID NO: 166	mU*mG*mU*GUCACCCGUUUCAGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
197	G018084 guide RNA targeting CIITA with guide sequence SEQ ID NO: 167	mA*mC*mC*CGUUUCAGGUGGGGUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
198	G018085 guide RNA targeting CIITA with guide sequence SEQ ID NO: 168	mC*mC*mC*GUUUCAGGUGGGGUGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

199	G018091 guide RNA targeting CIITA with guide sequence SEQ ID NO: 169	mU*mG*mU*GCAGACUCAGAGGUGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
200	G018100 guide RNA targeting CIITA with guide sequence SEQ ID NO: 170	mC*mA*mG*CGCAUCCAGGCUGCAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
201	G018117 guide RNA targeting CIITA with guide sequence SEQ ID NO: 171	mG*mC*mG*UCCACAUCCUGCAAGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
202	G018118 guide RNA targeting CIITA with guide sequence SEQ ID NO: 172	mG*mG*mC*GUCCACAUCCUGCAAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
203	G018120 guide RNA targeting CIITA with guide sequence SEQ ID NO: 173	mU*mG*mG*GCGUCCACAUCCUGCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUC CGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
204-210		Not used
211	amino acid sequence for exemplary linker	GGS
212	amino acid sequence for exemplary linker	GGGGS
213	amino acid sequence for exemplary linker	EAAAK

214	amino acid sequence for exemplary linker	SEGSA
215	amino acid sequence for exemplary linker	SEGSAGTST
216	amino acid sequence for exemplary linker	GGGGSGGGGS
217	amino acid sequence for exemplary linker	GGGGSEAAAK
218	amino acid sequence for exemplary linker	EAAAKGGGGS
219	amino acid sequence for exemplary linker	EAAAKEAAAK
220	amino acid sequence for exemplary linker	SEGSAGTSTESEGSA
221	amino acid sequence for exemplary linker	GGGGSGGGGSGGGGS
222	amino acid sequence for exemplary linker	GGGGSGGGGSEAAAK
223	amino acid sequence for exemplary linker	GGGGSEAAAKGGGGS
224	amino acid sequence for exemplary linker	EAAAKGGGGSEAAAK
225	amino acid sequence for	EAAAKEAAAKGGGGS

	exemplary linker	
226	amino acid sequence for exemplary linker	SEGSAGTSTESEGSAGTSTE
227	amino acid sequence for exemplary linker	GGGGSGGGGSGGGGSEAAAK
228	amino acid sequence for exemplary linker	GGGGSGGGGSEAAAKGGGGS
229	amino acid sequence for exemplary linker	GGGGSEAAAKGGGGSEAAAK
230	amino acid sequence for exemplary linker	GGGGSEAAAKEAAAKGGGGS
231	amino acid sequence for exemplary linker	GGGGSEAAAKEAAAKEAAAK
232	amino acid sequence for exemplary linker	EAAAKGGGGSGGGGSGGGGS
233	amino acid sequence for exemplary linker	EAAAKGGGGSGGGGSEAAAK
234	amino acid sequence for exemplary linker	EAAAKGGGGSEAAAKGGGGS
235	amino acid sequence for exemplary linker	EAAAKGGGGSEAAAKEAAAK
236	amino acid sequence for exemplary linker	EAAAKEAAAKGGGGSGGGGS
237	amino acid	EAAAKEAAAKGGGGSEAAAK

	sequence for exemplary linker	
238	amino acid sequence for exemplary linker	EAAAKEAAAKEAAAKGGGGS
239	amino acid sequence for exemplary linker	SEGSAGTSTESEGSAGTSTESEGSA
240	amino acid sequence for exemplary linker	GGGGSGGGGSGGGGSGGGGSGGGGS
241	amino acid sequence for exemplary linker	GGGGSGGGGSGGGGSGGGGSEAAAK
242	amino acid sequence for exemplary linker	GGGGSGGGGSGGGGSEAAAKGGGGS
243	amino acid sequence for exemplary linker	GGGGSGGGGSGGGGSEAAAKEAAAK
244	amino acid sequence for exemplary linker	GGGGSGGGGSEAAAKGGGGSGGGGS
245	amino acid sequence for exemplary linker	GGGGSGGGGSEAAAKGGGGSEAAAK
246	amino acid sequence for exemplary linker	GGGGSGGGGSEAAAKEAAAKGGGGS
247	amino acid sequence for exemplary linker	GGGGSGGGGSEAAAKEAAAKEAAAK
248	amino acid sequence for exemplary linker	GGGGSEAAAKGGGGSGGGGSGGGGS

249	amino acid sequence for exemplary linker	GGGGSEAAAKGGGGSGGGGSEAAAK
250	amino acid sequence for exemplary linker	GGGGSEAAAKGGGGSEAAAKGGGGS
251	amino acid sequence for exemplary linker	GGGGSEAAAKGGGGSEAAAKEAAAK
252	amino acid sequence for exemplary linker	GGGGSEAAAKEAAAKGGGGSGGGGS
253	amino acid sequence for exemplary linker	GGGGSEAAAKEAAAKEAAAKGGGGS
254	amino acid sequence for exemplary linker	GGGGSEAAAKEAAAKEAAAKEAAAK
255	amino acid sequence for exemplary linker	EAAAKGGGGSGGGGSGGGGSGGGGS
256	amino acid sequence for exemplary linker	EAAAKGGGGSGGGGSGGGGSEAAAK
257	amino acid sequence for exemplary linker	EAAAKGGGGSGGGGSEAAAKGGGGS
258	amino acid sequence for exemplary linker	EAAAKGGGGSGGGGSEAAAKEAAAK
259	amino acid sequence for exemplary linker	EAAAKGGGGSEAAAKGGGGSGGGGS
260	amino acid sequence for	EAAAKGGGGSEAAAKGGGGSEAAAK

	exemplary linker	
261	amino acid sequence for exemplary linker	EAAAKGGGGSEAAAKEAAAKGGGGS
262	amino acid sequence for exemplary linker	EAAAKGGGGSEAAAKEAAAKEAAAK
263	amino acid sequence for exemplary linker	EAAAKEAAAKGGGGSEAAAKGGGGS
264	amino acid sequence for exemplary linker	EAAAKEAAAKGGGGSEAAAKEAAAK
265	amino acid sequence for exemplary linker	EAAAKEAAAKEAAAKGGGGSEAAAK
266	amino acid sequence for exemplary linker	EAAAKEAAAKEAAAKEAAAKGGGGS
267	amino acid sequence for exemplary linker	EAAAKEAAAKEAAAKEAAAKEAAAK
268	amino acid sequence for exemplary linker	GTKDSTKDIPETPSKD
269	amino acid sequence for exemplary linker	GRDVRQPEVKEEKPES
270	amino acid sequence for exemplary linker	EGKSSGSGSESKSTAG
271	amino acid sequence for exemplary linker	TPGSPAGSPTSTEEGT
272	amino acid	GSEPATSGSETPGTST

	sequence for exemplary linker	
273-300		Not Used
301	Exemplary mRNA encoding APOBEC3A-Nme2D16A	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG CAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAAAUCCCAUCAACUACAUCCUGGGC CUGGCCAUCGGCAUCGCCUCCGUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCGGCUGAUCGACCUGGGCGUGCGGG UGUUCGAGCGGGCCGAGGUGCCCAAGACCGGCGACUCCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCCGUGCGGCGGCUGACCCGGCGGCG GGCCCACCGGCUGCUGCGGGCCCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCGACGAGAACGGCCUGAUCAAGUCC CUGCCCAACACCCCCUGGCAGCUGCGGGCCGCCGCCCCUGGACCGGAAGCUGACCCCCCCUGGAGUGGUCCGCCGUGCUGCUGCACCUGAUCA AGCACCGGGGCUACCUGUCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUGCUGAAGGGCGUGGCCAACAACGC CCACGCCCUGCAGACCGGCGACUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGUCCGGCCACAUCCGGAACCAGCGG GGCGACUACUCCCACACCUUCUCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCAGAAGGAGUUCGGCAACCCCCACG UGUCCGGCGGCCUGAAGGAGGGCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGACGCCGUGCAGAAGAUGCUGGGCCA CUGCACCUUCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGCUGACCAAGCUGAACAACCUGCGG AUCCUGGAGCAGGGCUCCGAGCGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUACCGGAAGUCCAAGCUGACCUACG CCCAGGCCCGGAAGCUGCUGGGCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGACAACGCCGAGGCCUCCACCCUGAU GGAGAUGAAGGCCUACCACGCCAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCCCCCUGAACCUGUCCUCCGAGCUG CAGGACGAGAUCGGCACCGCCUUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGACCGGGUGCAGCCCGAGAUCCUGG AGGCCCUGCUGAAGCACAUCUCCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUCGUGCCCCUGAUGGAGCAGGGCAA GCGGUACGACGAGGCCUGCGCCGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGAUCUACCUGCCCCCCAUCCCCGCC GACGAGAUCCGGAACCCCGUGGUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGUGCGGCGGUACGGCUCCCCCGCCC GGAUCCACAUCGAGACCGCCCGGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCG GGAGAAGGCCGCCGCCAAGUUCCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACAUCCUGAAGCUGCGGCUGUACGAG CAGCAGCACGGCAAGUGCCUGUACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUACGUGGAGAUCGACCACGCCCUGC CCUUCUCCCGGACCUGGGACGACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAACAAGGGCAACCAGACCCCCUACGA GUACUUCAACGGCAAGGACAACUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGUUCCCCCGGUCCAAGAAGCAGCGG AUCCUGCUGCAGAAGUUCGACGAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAACCGCUUCCUGUGCCAGUUCGUGG CCGACCACAUCCUGCUGACCGGCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAACCUGCUGCGGGGCUUCUGGGGCCU GCGGAAGGUGCGGGCCGAGAACGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCGUGGCCAUGCAGCAGAAGAUCACC CGGUUCGUGCGGUACAAGGAGAUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGUGCUGCACCAGAAGACCCACUUCC CCCAGCCCUGGGAGUUCUUCGCCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCCGAGUUCGAGGAGGCCGACACCCC CGAGAAGCUGCGGACCCUGCUGGCCGAGAAGCUGUCCUCCCGGCCCGAGGCCGUGCACGAGUACGUGACCCCCCUGUUCGUGUCCCGGGCC CCCAACCGGAAGAUGUCCGGCGCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAACGAGAAGAUCUCCGUGAAGCGGG

		UGUGGCUGACCGAGAUCAAGCUGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUCGAGCUGUACGAGGCCCUGAAGGC CCGGCUGGAGGCCUACGGCGGCAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGAAGGGCGGCCAGCUGGUGAAGGCC GUGCGGGUGGAGAAGACCCAGGAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGACAACGGCGACAUGGUGCGGGUGG ACGUGUUCUGCAAGGUGGACAAGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAGGUGGCCGAGAACAUCCUGCCCGA CAUCGACUGCAAGGGCUACCGGAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACCUGAUCGCCUUCCAGAAGGACGAG AAGUCCAAGGUGGAGUUCGCCUACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUGGCACGACAAGGGCUCCAAGGAGC AGCAGUUCCGGAUCUCCACCCAGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAGGAGAUCCGGCCCUGCCGGCUGAA GAAGCGGCCCCCCGUGCGGUCCGGAAAGCGGACCGCCGACGGCUCCGAGUUCGAGUCCCCCAAGAAGAAGCGGAAGGUGGAGUAGUGACUA GCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAU UCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCU
302	Exemplary open reading frame for APOBEC3A-Nme2D16A	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGA CCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAA CCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGG GUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGC UGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAU CAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCC CCGAGUCCGCAGCGUUCAAACCAAAUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCCGUGGGCUGGGCCAUGGUGGAGAU CGACGAGGAGGAGAACCCCAUCCGGCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGCCCAAGACCGGCGACUCCCUGGCC AUGGCCCGGCGGCUGGCCCGGUCCGUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGCCCGGCGGCUGCUGAAGCGGGAGG GCGUGCUGCAGGCCGCCGACUUCGACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAGCUGCGGGCCGCCGCCCUGGACCG GAAGCUGACCCCCCUGGAGUGGUCCGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCCAGCGGAAGAACGAGGGCGAGACC GCCGACAAGGAGCUGGGCGCCCUGCUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGACUUCCGGACCCCCGCCGAGCUGG CCCUGAACAAGUUCGAGAAGGAGUCCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUCUCCCGGAAGGACCUGCAGGCCGA GCUGAUCCUGCUGUUCGAGAAGCAGAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGGGCAUCGAGACCCUGCUGAUGACC CAGCGGCCCGCCCUGUCCGGCGACGCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCU ACACCGCCGAGCGGUUCAUCUGGCUGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAGCGGCCCCUGACCGACACCGAGCG GGCCACCCUGAUGGACGAGCCCUACCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGGGCCUGGAGGACACCGCCUUCUUC AAGGGCCUGCGGUACGGCAAGGACAACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGCCAUCUCCCGGGCCCUGGAGAAGG AGGGCCUGAAGGACAAGAAGUCCCCCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCCUUCUCCCUGUUCAAGACCGACGA GGACAUCACCGGCCGGCUGAAGGACCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCUCCUUCGACAAGUUCGUGCAGAUC UCCCUGAAGGCCCUGCGGCGGAUCGUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGCCGAGAUCUACGGCGACCACUACG GCAAGAAGAACACCGAGGAGAAGAUCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUGGUGCUGCGGGCCCUGUCCCAGGC CCGGAAGGUGAUCAACGGCGUGGUGCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCCGGGAGGUGGGCAAGUCCUUCAAG GACCGGAAGGAGAUCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUUCCGGGAGUACUUCCCCAACUUCG UGGGCGAGCCCAAGUCCAAGGACAUCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUGUACUCCGGCAAGGAGAUCAACCU GGUGCGGCUGAACGAGAAGGGCUACGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACGACUCCUUCAACAACAAGGUGCUG GUGCUGGGCUCCGAGAACCAGAACAAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAACUCCCGGGAGUGGCAGGAGUUCA

		AGGCCCGGGUGGAGACCUCCCGGUUCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGACGAGGACGGCUUCAAGGAGUGCAA CCUGAACGACACCCGGUACGUGAACCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCGGCAAGGGCAAGCGGCGGGUGUUC GCCUCCAACGGCCAGAUCACCAACCUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAACGACCGGCACCACGCCCUGGACG CCGUGGUGGUGGCCUGCUCCACCGUGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAGAUGAACGCCUUCGACGGCAAGAC CAUCGACAAGGAGACCGGCAAGGUGCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCGCCCAGGAGGUGAUGAUCCGGGUG UUCGGCAAGCCCGACGGCAAGCCCGAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCUGGCCGAGAAGCUGUCCUCCCGGC CCGAGGCCGUGCACGAGUACGUGACCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGCGCCCACAAGGACACCCUGCGGUC CGCCAAGCGGUUCGUGAAGCACAACGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGCUGGCCGACCUGGAGAACAUGGUG AACUACAAGAACGGCCGGGAGAUCGAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGGCAACGCCAAGCAGGCCUUCGACC CCAAGGACAACCCCUUCUACAAGAAGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAGGAGUCCGGCGUGCUGCUGAACAA GAAGAACGCCUACACCAUCGCCGACAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACAAGAAGGGCAAGAACCAGUACUUC AUCGUGCCCAUCUACGCCUGGCAGGUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCGGAUCGACGACUCCUACACCUUCU GCUUCUCCCUGCACAAGUACGACCUGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCCUACUACAUCAACUGCGACUCCUC CAACGGCCGGUUCUACCUGGCCUGGCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCCAGAACCUGGUGCUGAUCCAGAAG UACCAGGUGAACGAGCUGGGCAAGGAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGUCCGGAAAGCGGACCGCCGACGGCU CCGAGUUCGAGUCCCCCAAGAAGAAGCGGAAGGUGGAGUAG
303	Exemplary amino acid sequence for APOBEC3A-Nme2D16A	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNSGSETPGTSESATPESAAFKPNPINYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLA MARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGET ADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMT QRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFF KGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQI SLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFK DRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVL VLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVF ASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRV FGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMV NYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKGKNQYF IVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQK YQVNELGKEIRPCRLKKRPPVRSGKRTADGSEFESPKKKRKVE*
304	Exemplary mRNA encoding APOBEC3A-Nme2D16A	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGA GGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGAAGAAGGGCGGCUCCGGCGGCGGCGAGGCCUCCCCCGCCUCCGGC CCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGC GGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCG GCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCC

		CCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCU ACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCA CUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUG CGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAA AUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCCGUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCG GCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGCCCAAGACCGGCGACUCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCC GUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGCCCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCG ACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAGCUGCGGGCCGCCGCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUC CGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUG CUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGACUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGU CCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUCUCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCA GAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGGGCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGAC GCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGC UGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAGCGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUA CCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGGGCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGAC AACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGCCAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCC CCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCCUUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGA CCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCUCCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUC GUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGCCGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGA UCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUGGUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGU GCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCCGGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGG CAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUUCCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACA UCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUGUACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUA CGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACGACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAAC AAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAACUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGU UCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGACGAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAA CCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCGGCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAAC CUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAACGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCG UGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAGAUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGU GCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCGCCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCC GAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCUGGCCGAGAAGCUGUCCUCCCGGCCCGAGGCCGUGCACGAGUACGUGA CCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGCGCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAA CGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGCUGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUC GAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGGCAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGA AGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAGGAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGA CAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACAAGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAG GUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCGGAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACC UGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCCUACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUG GCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCCAGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAG

		GAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGUCCGGAAAGCGGACCGCCGACGGCUCCGAGUUCGAGUCCCCCAAGAAGA AGCGGAAGGUGGAGUAGUGACUAGCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAU GUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCU
305	Exemplary open reading frame for APOBEC3A-Nme2D16A	AUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGAGGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGA AGAAGAAGGGCGGCUCCGGCGGCGGCGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAA CAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGC UUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGC UGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCU GCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGG GACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGC CCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCC CGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAAAUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCC GUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCGGCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGC CCAAGACCGGCGACUCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCCGUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGC CCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCGACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAG CUGCGGGCCGCCGCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUCCGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCC AGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUGCUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGA CUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGUCCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUC UCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCAGAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGG GCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGACGCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGA GCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGCUGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAG CGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUACCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGG GCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGACAACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGC CAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCCCCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCC UUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGACCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCU CCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUCGUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGC CGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGAUCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUG GUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGUGCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCC GGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUU CCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACAUCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUG UACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUACGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACG ACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAACAAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAA CUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGUUCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGAC GAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAACCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCG GCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAACCUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAA CGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCGUGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAG AUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGUGCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCG CCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCCGAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCU

		GGCCGAGAAGCUGUCCUCCCGGCCCGAGGCCGUGCACGAGUACGUGACCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGC GCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAACGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGC UGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUCGAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGG CAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGAAGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAG GAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGACAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACA AGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAGGUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCG GAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACCUGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCC UACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUGGCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCC AGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAGGAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGUC CGGAAAGCGGACCGCCGACGGCUCCGAGUUCGAGUCCCCCAAGAAGAAGCGGAAGGUGGAGUAG
306	Exemplary amino acid sequence for APOBEC3A-Nme2D16A	MDGSGGGSPKKKRKVEDKRPAATKKAGQAKKKKGGSGGGEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRG FLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLR DAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNSGSETPGTSESATPESAAFKPNPINYILGLAIGIAS VGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQ LRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTF SRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSE RPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTA FSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPV VLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCL YSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD EDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKE MNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSG AHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQ ESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFA YYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVRSGKRTADGSEFESPKKKRKVE*
307	Exemplary mRNA encoding APOBEC3A-Nme2D16A	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGA GGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGAAGAAGGGCGGCUCCGGCGGCGGCGAGGCCUCCCCCGCCUCCGGC CCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGC GGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCG GCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCC CCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCU ACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCA CUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUG CGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAA AUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCCGUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCG GCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGCCCAAGACCGGCGACUCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCC GUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGCCCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCG

		ACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAGCUGCGGGCCGCCGCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUC CGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUG CUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGACUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGU CCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUCUCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCA GAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGGGCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGAC GCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGC UGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAGCGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUA CCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGGGCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGAC AACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGCCAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCC CCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCCUUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGA CCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCUCCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUC GUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGCCGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGA UCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUGGUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGU GCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCCGGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGG CAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUUCCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACA UCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUGUACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUA CGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACGACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAAC AAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAACUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGU UCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGACGAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAA CCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCGGCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAAC CUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAACGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCG UGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAGAUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGU GCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCGCCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCC GAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCUGGCCGAGAAGCUGUCCUCCCGGCCCGAGGCCGUGCACGAGUACGUGA CCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGCGCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAA CGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGCUGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUC GAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGGCAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGA AGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAGGAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGA CAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACAAGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAG GUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCGGAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACC UGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCCUACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUG GCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCCAGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAG GAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGGAGGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGA AGAAGUACCCCUACGACGUGCCCGACUACGCCGGCUACCCCUACGACGUGCCCGACUACGCCGGCUCCUACCCCUACGACGUGCCCGACUA CGCCGCCGCCCCCGCCGCCAAGAAGAAGAAGCUGGACUAGCUAGUGACUAGCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUAC AUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCU
308	Exemplary open reading frame	AUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGAGGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGA AGAAGAAGGGCGGCUCCGGCGGCGGCGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAA

	for APOBEC3A-Nme2D16A	CAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGC UUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGC UGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCU GCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGG GACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGC CCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCC CGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAAAUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCC GUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCGGCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGC CCAAGACCGGCGACUCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCCGUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGC CCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCGACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAG CUGCGGGCCGCCGCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUCCGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCC AGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUGCUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGA CUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGUCCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUC UCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCAGAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGG GCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGACGCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGA GCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGCUGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAG CGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUACCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGG GCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGACAACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGC CAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCCCCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCC UUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGACCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCU CCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUCGUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGC CGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGAUCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUG GUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGUGCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCC GGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUU CCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACAUCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUG UACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUACGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACG ACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAACAAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAA CUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGUUCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGAC GAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAACCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCG GCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAACCUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAA CGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCGUGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAG AUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGUGCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCG CCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCCGAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCU GGCCGAGAAGCUGUCCUCCCGGCCCGAGGCCGUGCACGAGUACGUGACCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGC GCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAACGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGC UGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUCGAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGG CAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGAAGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAG GAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGACAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACA AGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAGGUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCG

		GAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACCUGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCC UACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUGGCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCC AGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAGGAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGGA GGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGAAGAAGUACCCCUACGACGUGCCCGACUACGCCGGCUACCCCUAC GACGUGCCCGACUACGCCGGCUCCUACCCCUACGACGUGCCCGACUACGCCGCCGCCCCCGCCGCCAAGAAGAAGAAGCUGGACU
309	Exemplary amino acid sequence for APOBEC3A-Nme2D16A	MDGSGGGSPKKKRKVEDKRPAATKKAGQAKKKKGGSGGGEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRG FLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLR DAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNSGSETPGTSESATPESAAFKPNPINYILGLAIGIAS VGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQ LRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTF SRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSE RPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTA FSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPV VLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCL YSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD EDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKE MNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSG AHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQ ESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFA YYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVREDKRPAATKKAGQAKKKKYPYDVPDYAGYPY DVPDYAGSYPYDVPDYAAAPAAKKKKLD
310	EXEMPLARY MRNA ENCODING APOBEC3A-NME2D16A	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGA GGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGAAGAAGGGCGGCUCCGGCGGCGGCGAGGCCUCCCCCGCCUCCGGC CCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGC GGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCG GCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCC CCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCU ACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCA CUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUG CGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAA AUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCCGUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCG GCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGCCCAAGACCGGCGACUCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCC GUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGCCCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCG ACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAGCUGCGGGCCGCCGCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUC CGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUG CUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGACUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGU

		CCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUCUCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCA GAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGGGCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGAC GCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGC UGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAGCGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUA CCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGGGCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGAC AACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGCCAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCC CCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCCUUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGA CCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCUCCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUC GUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGCCGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGA UCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUGGUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGU GCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCCGGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGG CAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUUCCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACA UCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUGUACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUA CGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACGACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAAC AAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAACUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGU UCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGACGAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAA CCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCGGCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAAC CUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAACGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCG UGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAGAUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGU GCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCGCCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCC GAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCUGGCCGAGAAGCUGUCCUCCCGGCCCGAGGCCGUGCACGAGUACGUGA CCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGCGCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAA CGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGCUGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUC GAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGGCAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGA AGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAGGAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGA CAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACAAGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAG GUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCGGAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACC UGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCCUACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUG GCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCCAGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAG GAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGUAGUGACUAGCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUA CAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUC U
311	Exemplary open reading frame for APOBEC3A-Nme2D16A	AUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGAGGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGA AGAAGAAGGGCGGCUCCGGCGGCGGCGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAA CAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGC UUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGC UGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCU GCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGG

		GACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGC CCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCC CGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAAAUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCC GUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCGGCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGC CCAAGACCGGCGACUCCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCCGUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGC CCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCGACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAG CUGCGGGCCGCCGCCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUCCGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCC AGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCCUGCUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGA CUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGUCCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUC UCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCAGAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGG GCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGACGCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGA GCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGCUGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAG CGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUACCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGG GCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGACAACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGC CAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCCCCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCC UUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGACCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCU CCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUCGUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGC CGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGAUCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUG GUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGUGCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCC GGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUU CCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACAUCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUG UACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUACGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACG ACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAACAAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAA CUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGUUCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGAC GAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAACCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCG GCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAACCUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAA CGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCGUGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAG AUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGUGCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCG CCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCCGAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCU GGCCGAGAAGCUGUCCUCCCCGGCCCGAGGCCGUGCACGAGUACGUGACCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGC GCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAACGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGC UGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUCGAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGG CAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGAAGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAG GAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGACAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACA AGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAGGUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCG GAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACCUGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCC UACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUGGCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCC AGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAGGAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGUA G

312	Exemplary amino acid sequence for APOBEC3A-Nme2D16A	MDGSGGGSPKKKRKVEDKRPAATKKAGQAKKKKGGSGGGEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRG FLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLR DAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNSGSETPGTSESATPESAAFKPNPINYILGLAIGIAS VGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQ LRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTF SRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSE RPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTA FSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPV VLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCL YSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD EDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKE MNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSG AHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQ ESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFA YYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVR*
313	Exemplary amino acid sequence for NLS-NLS-APOBEC3A-L070-Nme2D16A	MDGSGGGSPKKKRKVEDKRPAATKKAGQAKKKKGGSGGGEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRG FLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLR DAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGNGTKDSTKDIPETPSKDAAFKPNPINYILGLAIGIAS VGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQ LRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTF SRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSE RPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTA FSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPV VLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCL YSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFD EDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKE MNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSG AHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQ ESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFA YYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVR
314	mRNA encoding BC22-2XUGI with a C-terminal HiBiT tag (BC22-2XUGI-HibIT)	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG

		CAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAAC UCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGA AGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCG GAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUC CUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGG CCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUC GAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCG CCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGA CCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCC GACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAA GGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCC AUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCU CCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAA GAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCC GAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGU GACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACC GUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCC UGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCU GACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUG AAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCA GGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAG GUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGG GCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAA CACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUG UCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCA GCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACC CGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGG UGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCA CGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUG UACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAA GGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCC AAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCC CCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCAC

		CAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAG CUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUU CGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGAC AAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGG GCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCA CCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACUCCGGCGGCUCCGGCGGCUCCGGCGGCUCCACC AACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGA UCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGACGCCCC CGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCGGCGGCUCCGGCGGC UCCACCAACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCUGGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGG AGGUGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCCUACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGA CGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACGGCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCAAGCGGACC GCCGACGGCUCCGAGUUCGAGCCCAAGAAGAAGCGGAAGGUGUCCGAGUCCGCCACCCCCGAGUCCGUGUCCGGCUGGCGGCUGUUCAAGA AGAUCUCCUGAUAGCUAGCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAUGUUGUC CCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUCUCGAGAAAAAAAAAAAAAUGGAAAAAAAAAAAA CGGAAAAAAAAAAAAAGGUAAAAAAAAAAAAUAUAAAAAAAAAAAAAACAUAAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAAACU CAAAAAAAAAAAAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAACGCAAAAAAAAAAAACAC AAAAAAAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAA AAAAAAAAAAUAGAAAAAAAAAAAAGUUAAAAAAAAAAAAACUGAAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
315	mRNA encoding BC22-Nme2D16A (Nme2 BC22n)	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGA GGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGAAGAAGGGCGGCUCCGGCGGCGGCGAGGCCUCCCCCGCCUCCGGC CCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGC GGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCG GCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCC CCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCU ACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCA CUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUG CGGGCCAUCCUGCAGAACCAGGGCAACUCCGGCUCCGAGACCCCCGGCACCUCCGAGUCCGCCACCCCCGAGUCCGCAGCGUUCAAACCAA AUCCCAUCAACUACAUCCUGGGCCUGGCCAUCGGCAUCGCCUCCGUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCG GCUGAUCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGCCCAAGACCGGCGACUCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCC GUGCGGCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGCCCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCG ACGAGAACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAGCUGCGGGCCGCCGCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUC CGCCGUGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUG CUGAAGGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGACUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGU CCGGCCACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUCUCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCA GAAGGAGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGGGCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGAC GCCGUGCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGC UGACCAAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAGCGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUA

		CCGGAAGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGGGCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGAC AACGCCGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGCCAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCC CCCUGAACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCCUUCUCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGA CCGGGUGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCUCCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUC GUGCCCCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGCCGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGA UCUACCUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUGGUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGU GCGGCGGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCCGGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGG CAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUUCCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACA UCCUGAAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUGUACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUA CGUGGAGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACGACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAAC AAGGGCAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAACUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGU UCCCCCGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGACGAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAA CCGCUUCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCGGCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAAC CUGCUGCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAACGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCG UGGCCAUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAGAUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGU GCUGCACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCGCCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCC GAGUUCGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCUGGCCGAGAAGCUGUCCUCCCGGCCCGAGGCCGUGCACGAGUACGUGA CCCCCCUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGCGCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAA CGAGAAGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGCUGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUC GAGCUGUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGGCAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGA AGGGCGGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAGGAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGA CAACGGCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACAAGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAG GUGGCCGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCGGAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACC UGAUCGCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCCUACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUG GCACGACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCCAGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAG GAGAUCCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGUAGUGACUAGCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUA CAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUC UCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAACGGAAAAAAAAAAAAGGUAAAAAAAAAAAAUAUAAAAAAAAAAACAUAAAAAAAAAAA ACGAAAAAAAAAAAACGUAAAAAAAAAAAACUCAAAAAAAAAAAGAUAAAAAAAAAAAACCUAAAAAAAAAAAAUGUAAAAAAAAAAAGG GAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAACGAA AAAAAAAAACCCAAAAAAAAAAAGACAAAAAAAAAAAUAGAAAAAAAAAAGUUAAAAAAAAAAACUGAAAAAAAAAAUUUAAAA AAAAAAAUCUAG
316	mRNA encoding UGI with a C-terminal HiBiT tag (UGI-HiBiT)	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGACCAACCUGUCCGACAUCAUCGAGAAGGAGACCGGCAAGCAGCU GGUGAUCCAGGAGUCCAUCCUGAUGCUGCCCGAGGAGGUGGAGGAGGUGAUCGGCAACAAGCCCGAGUCCGACAUCCUGGUGCACACCGCC UACGACGAGUCCACCGACGAGAACGUGAUGCUGCUGACCUCCGACGCCCCCGAGUACAAGCCCUGGGCCCUGGUGAUCCAGGACUCCAACG GCGAGAACAAGAUCAAGAUGCUGUCCGGCGGCUCCAAGCGGACCGCCGACGGCUCCGAGUUCGAGUCCCCCAAGAAGAAGCGGAAGGUGGA GUCCGAGUCCGCCACCCCCGAGUCCGUGUCCGGCUGGCGGCUGUUCAAGAAGAUCUCCUGAUAGCUAGCACCAGCCUCAAGAACACCCGAA UGGAGUCUCUAAGCUACAUAAUACCAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGA AAGUUUCUUCACAUUCUCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAACGGAAAAAAAAAAAAGGUAAAAAAAAAAAAUAUAAAAAAAA

		AAAACAUAAAAAAAAAAAACGAAAAAAAAAAAACGUAAAAAAAAAAAACUCAAAAAAAAAAAAGAUAAAAAAAAAAAACCUAAAAAAAAAA AAUGUAAAAAAAAAAAAGGGAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAA AUCUAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAAUAGAAAAAAAAAAAAGUUAAAAAAAAAAAAC UGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
317-319		Not Used
320	amino acid sequence for Nme2Cas9	MDGSGGGSPKKKRKVEDKRPAATKKAGQAKKKKGGSGGGAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPK TGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQR KNEGETADKELGALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGI ETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGL EDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISF DKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETARE VGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDS FNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLCQFVADHILLTGK GKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQ EVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGAHKDTLRSAKRFVKHNEKISVKRVWLTEIKLA DLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCKVDKK GKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQN LVLIQKYQVNELGKEIRPCRLKKRPPVR
321	Amino acid sequence for base editor with linker L070	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNGTKDSTKDIPETPSKDDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKD FQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
322	Amino acid sequence for	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH

	base editor with linker L071	SQALSGRLRAILQNQGNGRDVRQPEVKEEKPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKD FQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
323	Amino acid sequence for base editor with linker L072	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEGKSSGSGSESKSTAGDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKD FQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
324	Amino acid sequence for base editor with linker L073	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNTPGSPAGSPTSTEEGTDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL

		IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKD FQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
325	Amino acid sequence for base editor with linker L074	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNGSEPATSGSETPGTSTDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKD FLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQL IHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKD FQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
326	Amino acid sequence for base editor with linker L019	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEAAAKEAAAKEAAAKDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY

		EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE KLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGGSPKKKRKV
327	Amino acid sequence for base editor with linker L015	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNGGGGSEAAAKEAAAKDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE KLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDEN VMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
328	Amino acid sequence for base editor with linker L016	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEAAAKGGGGSGGGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI

		HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE KLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDEN VMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV*
329	Amino acid sequence for base editor with linker L019	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEAAAKEAAAKEAAAKDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGI KELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYE KLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDEN VMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV*
330	Amino acid sequence for base editor with linker L021	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNGGGGSGGGGSGGGGSGGGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL HAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR

		IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDE STDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
331	Amino acid sequence for base editor with linker L024	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNGGGGSGGGGSEAAAKEAAAKDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL HAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDE STDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV*
332	Amino acid sequence for base editor with linker L025	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNGGGGSEAAAKGGGGSGGGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL HAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE

		VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDE STDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
333	Amino acid sequence for base editor with linker L036	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEAAAKEAAAKEAAAKEAAAKDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL HAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDT TIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDE STDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
334	Amino acid sequence for base editor with linker L051	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNGGGGSEAAAKEAAAKGGGGSEAAAKDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDN VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS

		KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVH TAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
335	Amino acid sequence for base editor with linker L062	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEAAAKEAAAKGGGGSGGGGSGGGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDN VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVH TAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
336	Amino acid sequence for base editor with linker L063	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEAAAKEAAAKGGGGSGGGGSEAAAKDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDN VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT

		LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVH TAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
337	Amino acid sequence for base editor with linker L066	MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYR VTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEH SQALSGRLRAILQNQGNEAAAKEAAAKEAAAKGGGGSGGGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSR ERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDN VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVH TAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV
338-339		Not Used
340	mRNA for Nme2 Cas9 nickase	GGGAAGCUCAGAAUAAACGCUCAACUUUGGCCGGAUCUGCCACCAUGGACGGCUCCGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGGA GGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGAAGAAGAAGGGCGGCUCCGGCGGCGGCGCCGCCUUCAAGCCCAACCCC AUCAACUACAUCCUGGGCCUGGACAUCGGCAUCGCCUCCGUGGGCUGGGCCAUGGUGGAGAUCGACGAGGAGGAGAACCCCAUCCGGCUGA UCGACCUGGGCGUGCGGGUGUUCGAGCGGGCCGAGGUGCCCAAGACCGGCGACUCCCUGGCCAUGGCCCGGCGGCUGGCCCGGUCCGUGCG GCGGCUGACCCGGCGGCGGGCCCACCGGCUGCUGCGGGCCCGGCGGCUGCUGAAGCGGGAGGGCGUGCUGCAGGCCGCCGACUUCGACGAG AACGGCCUGAUCAAGUCCCUGCCCAACACCCCCUGGCAGCUGCGGGCCGCCGCCCUGGACCGGAAGCUGACCCCCCUGGAGUGGUCCGCCG UGCUGCUGCACCUGAUCAAGCACCGGGGCUACCUGUCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAAGGAGCUGGGCGCCCUGCUGAA GGGCGUGGCCAACAACGCCCACGCCCUGCAGACCGGCGACUUCCGGACCCCCGCCGAGCUGGCCCUGAACAAGUUCGAGAAGGAGUCCGGC CACAUCCGGAACCAGCGGGGCGACUACUCCCACACCUUCUCCCGGAAGGACCUGCAGGCCGAGCUGAUCCUGCUGUUCGAGAAGCAGAAGG AGUUCGGCAACCCCCACGUGUCCGGCGGCCUGAAGGAGGGCAUCGAGACCCUGCUGAUGACCCAGCGGCCCGCCCUGUCCGGCGACGCCGU GCAGAAGAUGCUGGGCCACUGCACCUUCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCUACACCGCCGAGCGGUUCAUCUGGCUGACC AAGCUGAACAACCUGCGGAUCCUGGAGCAGGGCUCCGAGCGGCCCCUGACCGACACCGAGCGGGCCACCCUGAUGGACGAGCCCUACCGGA AGUCCAAGCUGACCUACGCCCAGGCCCGGAAGCUGCUGGGCCUGGAGGACACCGCCUUCUUCAAGGGCCUGCGGUACGGCAAGGACAACGC

		CGAGGCCUCCACCCUGAUGGAGAUGAAGGCCUACCACGCCAUCUCCCGGGCCCUGGAGAAGGAGGGCCUGAAGGACAAGAAGUCCCCCCUG AACCUGUCCUCCGAGCUGCAGGACGAGAUCGGCACCGCCUUCUCCCCUGUUCAAGACCGACGAGGACAUCACCGGCCGGCUGAAGGACCGGG UGCAGCCCGAGAUCCUGGAGGCCCUGCUGAAGCACAUCUCCUUCGACAAGUUCGUGCAGAUCUCCCUGAAGGCCCUGCGGCGGAUCGUGCC CCUGAUGGAGCAGGGCAAGCGGUACGACGAGGCCUGCGCCGAGAUCUACGGCGACCACUACGGCAAGAAGAACACCGAGGAGAAGAUCUAC CUGCCCCCCAUCCCCGCCGACGAGAUCCGGAACCCCGUGGUGCUGCGGGCCCUGUCCCAGGCCCGGAAGGUGAUCAACGGCGUGGUGCGGC GGUACGGCUCCCCCGCCCGGAUCCACAUCGAGACCGCCCGGGAGGUGGGCAAGUCCUUCAAGGACCGGAAGGAGAUCGAGAAGCGGCAGGA GGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGUUCCGGGAGUACUUCCCCAACUUCGUGGGCGAGCCCAAGUCCAAGGACAUCCUG AAGCUGCGGCUGUACGAGCAGCAGCACGGCAAGUGCCUGUACUCCGGCAAGGAGAUCAACCUGGUGCGGCUGAACGAGAAGGGCUACGUGG AGAUCGACCACGCCCUGCCCUUCUCCCGGACCUGGGACGACUCCUUCAACAACAAGGUGCUGGUGCUGGGCUCCGAGAACCAGAACAAGGG CAACCAGACCCCCUACGAGUACUUCAACGGCAAGGACAACUCCCGGGAGUGGCAGGAGUUCAAGGCCCGGGUGGAGACCUCCCGGUUCCCC CGGUCCAAGAAGCAGCGGAUCCUGCUGCAGAAGUUCGACGAGGACGGCUUCAAGGAGUGCAACCUGAACGACACCCGGUACGUGAACCGGU UCCUGUGCCAGUUCGUGGCCGACCACAUCCUGCUGACCGGCAAGGGCAAGCGGCGGGUGUUCGCCUCCAACGGCCAGAUCACCAACCUGCU GCGGGGCUUCUGGGGCCUGCGGAAGGUGCGGGCCGAGAACGACCGGCACCACGCCCUGGACGCCGUGGUGGUGGCCUGCUCCACCGUGGCC AUGCAGCAGAAGAUCACCCGGUUCGUGCGGUACAAGGAGAUGAACGCCUUCGACGGCAAGACCAUCGACAAGGAGACCGGCAAGGUGCUGC ACCAGAAGACCCACUUCCCCCAGCCCUGGGAGUUCUUCGCCCAGGAGGUGAUGAUCCGGGUGUUCGGCAAGCCCGACGGCAAGCCCGAGUU CGAGGAGGCCGACACCCCCGAGAAGCUGCGGACCCUGCUGGCCGAGAAGCUGUCCUCCCCGGCCCGAGGCCGUGCACGAGUACGUGACCCCC CUGUUCGUGUCCCGGGCCCCCAACCGGAAGAUGUCCGGCGCCCACAAGGACACCCUGCGGUCCGCCAAGCGGUUCGUGAAGCACAACGAGA AGAUCUCCGUGAAGCGGGUGUGGCUGACCGAGAUCAAGCUGGCCGACCUGGAGAACAUGGUGAACUACAAGAACGGCCGGGAGAUCGAGCU GUACGAGGCCCUGAAGGCCCGGCUGGAGGCCUACGGCGGCAACGCCAAGCAGGCCUUCGACCCCAAGGACAACCCCUUCUACAAGAAGGGC GGCCAGCUGGUGAAGGCCGUGCGGGUGGAGAAGACCCAGGAGUCCGGCGUGCUGCUGAACAAGAAGAACGCCUACACCAUCGCCGACAACG GCGACAUGGUGCGGGUGGACGUGUUCUGCAAGGUGGACAAGAAGGGCAAGAACCAGUACUUCAUCGUGCCCAUCUACGCCUGGCAGGUGGC CGAGAACAUCCUGCCCGACAUCGACUGCAAGGGCUACCGGAUCGACGACUCCUACACCUUCUGCUUCUCCCUGCACAAGUACGACCUGAUC GCCUUCCAGAAGGACGAGAAGUCCAAGGUGGAGUUCGCCUACUACAUCAACUGCGACUCCUCCAACGGCCGGUUCUACCUGGCCUGGCACG ACAAGGGCUCCAAGGAGCAGCAGUUCCGGAUCUCCACCCAGAACCUGGUGCUGAUCCAGAAGUACCAGGUGAACGAGCUGGGCAAGGAGAU CCGGCCCUGCCGGCUGAAGAAGCGGCCCCCCGUGCGGUAGCUAGCACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAAUAC CAACUUACACUUUACAAAAUGUUGUCCCCCAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUCUCGAGA AAAAAAAAAAAUGGAAAAAAAAAAAACGGAAAAAAAAAAAAAGGUAAAAAAAAAAAAAUAUAAAAAAAAAAAAAACAUAAAAAAAAAAAAACGAAA AAAAAAAAACGUAAAAAAAAAAAAACUCAAAAAAAAAAAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAGGGAAAAA AAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAAAAUGCAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAAACGAAAAAAA AAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAAUAGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAA AAAUCUAG
341	mRNA encoding base editor with linker L070	GGGaagcucagaauaaacgcucaacuuuggccggaucugccacCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG CAACggcaccaaggacuccaccaaggacaucccccgagacccccuccaaggacGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAAC

		UCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGA AGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCG GAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUC CUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGG CCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUC GAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCG CCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGA CCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCC GACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAA GGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCC AUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCU CCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAA GAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCC GAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGU GACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACC GUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCC UGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCU GACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUG AAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCA GGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAG GUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGG GCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAA CACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUG UCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCA GCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACC CGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGG UGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCA CGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUG UACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAA GGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCC AAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCC CCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCAC CAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAG

		CUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUU CGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGAC AAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGG GCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCA CCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUG UGACUAGCaccagccucaagaacacccgaauggagucucuaagcuacauaauaccaacuuacacuuuacaaaauguuguccccccaaaaugu agccauucguaucugcuccuaauaaaaagaaaguuucuucacauucuCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAAACGGAAAAAAAA AAAAGGUAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAACUCAAAAAAAAAA AAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAA AUGCAAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAU AGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
342	mRNA encoding base editor with linker L071	GGGaagcucagaauaaacgcucaacuuuggccggaucugccacCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG CAACggccgggacgugcggcagcccgaggugaaggaggagaagcccgaguccGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAAC UCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGA AGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCG GAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUC CUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGG CCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUC GAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCG CCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGA CCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCC GACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAA GGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCC AUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCU CCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAA GAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCC GAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGU GACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACC GUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCC

		UGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCU GACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUG AAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCA GGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAG GUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGG GCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAA CACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUG UCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCA GCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACC CGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGG UGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCA CGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUG UACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAA GGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCC AAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCC CCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCAC CAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAG CUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUU CGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGAC AAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGG GCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCA CCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUG UGACUAGCaccagccucaagaacacccgaauggagucucuaagcuacauaauaccaacuuacacuuuacaaaauguuguccccccaaaaugu agccauucguaucugcuccuaauaaaaagaaaguuucuucacauucuCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAAACGGAAAAAAAA AAAAGGUAAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAAACUCAAAAAAAAAA AAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAAGGGAAAAAAAAAAAAACGCAAAAAAAAAAAAACACAAAAAAAAAAA AUGCAAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAU AGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
343	mRNA encoding base editor with linker L072	GGGaagcucagaauaaacgcucaacuuuggccggaucugccacCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG

		CAACgagggcaaguccuccggcuccggcuccgaguccaaguccaccgccggcGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAAC UCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGA AGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCG GAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUC CUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGG CCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUC GAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCG CCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGA CCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCC GACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAA GGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCC AUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCU CCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAA GAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCC GAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGU GACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACC GUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCC UGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCU GACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUG AAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCA GGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAG GUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGG GCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAA CACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUG UCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCA GCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACC CGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGG UGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCA CGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUG UACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAA GGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCC AAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCC CCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCAC

		CAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAG CUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUU CGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGAC AAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGG GCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCA CCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUG UGACUAGCaccagccucaagaacacccgaauggagucucuaagcuacauaauaccaacuuacacuuuacaaaauguuguccccccaaaaugu agccauucguaucugcuccuaauaaaaagaaaguuucuucacauucuCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAACGGAAAAAAAA AAAAGGUAAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAACUCAAAAAAAAAA AAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAA AUGCAAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAU AGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
344	mRNA encoding base editor with linker L073	GGGaagcucagaauaaacgcucaacuuuggccggaucugccacCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG CAACaccccggcucccccgccggcucccccaccuccaccgaggagggcaccGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAAC UCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGA AGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCG GAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUC CUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGG CCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUC GAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCG CCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGA CCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCC GACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAA GGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCC AUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCU CCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAA GAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCC GAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGU GACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACC

		GUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCC UGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCU GACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUG AAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCA GGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAG GUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGG GCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAA CACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUG UCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCA GCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACC CGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGG UGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCA CGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUG UACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAA GGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCC AAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCC CCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCAC CAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAG CUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUU CGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGAC AAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGG GCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCA CCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUG UGACUAGCaccagccucaagaacacccgaauggagucucuaagcuacauaauaccaacuuacacuuuacaaaauguuguccccccaaaaugu agccauucguaucugcuccuaauaaaaagaaaguuucuucacauucuCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAAACGGAAAAAAAA AAAAGGUAAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAAACUCAAAAAAAAAA AAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAA AUGCAAAAAAAAAAAAAUCGAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAU AGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
345	mRNA encoding base editor with linker L074	GGGaagcucagaauaaacgcucaacuuuggccggaucugccacCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC

		ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG CAACggcuccgagcccgccaccuccggcuccgagacccccggcaccuccaccGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAAC UCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGA AGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCG GAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUC CUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCU ACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGG CCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUC GAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCG CCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGA CCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCC GACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGU CCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAA GGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCC AUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCU CCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAA GAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCC GAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACA AGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGU GACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACC GUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCC UGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCU GACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUG AAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACU UCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCA GGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAG GUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGG GCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAA CACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUG UCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACC GGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCA GCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACC CGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGG UGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCA CGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUG UACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCU UCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAA GGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCC AAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCC

		CCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCAC CAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAG CUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCC UGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUU CGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGAC AAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGG GCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCA CCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUG UGACUAGCaccagccucaagaacacccgaauggagucucuaagcuacauaauaccaacuuacacuuuacaaaauguuguccccccaaaaugu agccauucguaucugcuccuaauaaaaagaaaguuucuucacauucuCUCGAGAAAAAAAAAAAAUGGAAAAAAAAAAAAACGGAAAAAAAA AAAAGGUAAAAAAAAAAAAAUAUAAAAAAAAAAAAACAUAAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAAACUCAAAAAAAAAA AAGAUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAAA AUGCAAAAAAAAAAAAAUCGAAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAAU AGAAAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
346	mRNA encoding base editor with linker L019	GGGaagcucagaauaaacgcucaacuuuggccggaucugccacCAUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCA CAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGACCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUG AAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAACCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGG ACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGGGUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGC CGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGCUGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAG GAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAUCAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACC ACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCACUCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGG CAACgaggccgccgccaaggaggccgccgccaaggaggccgccgccaagGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCC GUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGA ACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAA GAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUG GUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACC ACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCA CUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAG GAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCC AGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCU GGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGAC CUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCG CCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGA GAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUC CUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCA UCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAU CGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAG GAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGA ACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGAC

		CGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUG AAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGG GCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGAC CCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAG CGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCC UGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGU GUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUG GUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCC AGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACAC CCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCC GACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGG GCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCG GAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGG CAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGA UCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGA CGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUAC GACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCA AGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGG CCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAG GAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCA CCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAU CAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUG CCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGC CCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGU GGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAG GUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCG CCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCA GUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGA CUAGCaccagccucaagaacacccgaauggagucucuaagcuacauaaauaccaacuuacacuuuacaaaauguuguccccccaaaauguagc cauucguaucugcuccuaauaaaaagaaaguuucuucacauucuCUCGAGAAAAAAAAAAAAAUGGAAAAAAAAAAAAACGGAAAAAAAAAAAA AGGUAAAAAAAAAAAAAUAUAAAAAAAAAAAACAUAAAAAAAAAAAAAACGAAAAAAAAAAAAACGUAAAAAAAAAAAAACUCAAAAAAAAAAAAG AUAAAAAAAAAAAAACCUAAAAAAAAAAAAAUGUAAAAAAAAAAAAAGGGAAAAAAAAAAAAACGCAAAAAAAAAAAACACAAAAAAAAAAAAUG CAAAAAAAAAAAAUCGAAAAAAAAAAAAAUCUAAAAAAAAAAAAACGAAAAAAAAAAAAACCCAAAAAAAAAAAAGACAAAAAAAAAAAAUAGA AAAAAAAAAAAGUUAAAAAAAAAAAACUGAAAAAAAAAAAAUUUAAAAAAAAAAAAUCUAG
347	mRNA encoding base editor with linker L015	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG

		GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGGAGGAGGAGGAAGCGAAGCA GCAGCAAAGGAAGCAGCAGCAAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACG AAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGA CAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAA AUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAA GACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAG CACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAAC CCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAG UCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGG ACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUG AGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGA GCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGA ACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAAC GGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAG AACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGA ACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUC CCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCG AAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCC GAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUC CUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCA AGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAU CAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAA AUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAA GACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAG AAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCAC GAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGG GAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAA GAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUAC CUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCC CGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGA AGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCA GAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGA UCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAG CGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGA ACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGA GCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAG

		GUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACA GCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGU CGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAG AACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGG AAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUA CCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGAC GAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACA GAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGA CACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACA AGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCA UCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGA CGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAA AACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAAAGCAUCUCAGCCUAC CAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAA CAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAAAAAAAAAAAAAAAA AAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAU
348	mRNA encoding base editor with linker L016	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAGGCCGCCGCCAAGGGCGGC GGCGGCAGCGGCGGCGGCGGCAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACG AAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGA CAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAA AUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAA GACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAG CACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAAC CCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAG UCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGG ACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUG AGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGA GCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGA ACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAAC GGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAG AACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGA ACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUC

		CCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCG AAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCC GAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUC CUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCA AGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAU CAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAA AUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAA GACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAG AAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCAC GAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGG GAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAA GAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUAC CUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCC CGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGA AGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCA GAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGA UCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAG CGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGA ACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGA GCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAG GUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACA GCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGU CGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAG AACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGG AAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUA CCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGAC GAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACA GAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGA CACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACA AGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCA UCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGA CGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAA AACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAAAGCAUCUCAGCCUAC CAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAA CAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAU
349	mRNA encoding base editor with	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC

	linker L019	GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAGGCCGCCGCCAAGGAGGCC GCCGCCAAGGAGGCCGCCGCCAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACG AAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGA CAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAA AUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAA GACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAG CACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAAC CCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAG UCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGG ACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUG AGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGA GCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGA ACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAAC GGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAG AACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGA ACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUC CCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCG AAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCC GAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUC CUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCA AGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAU CAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAA AUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAA GACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAG AAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCAC GAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGG GAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAA GAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUAC CUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCC CGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGA AGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCA GAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGA UCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAG CGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGA

		ACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGA GCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGG AGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAG GUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACA GCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGU CGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAG AACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGG AAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUA CCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGAC GAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACA GAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGA CACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACA AGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCA UCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGA CGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAA AACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAAAGCAUCUCAGCCUAC CAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAAA CAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAAAAAAAAAAAAAAAA AAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAU
350	mRNA encoding base editor with linker L021	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGGCGGCGGCGGCAGCGGCGGC GGCGGCAGCGGCGGCGGCGGCAGCGGCGGCGGCGGCAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGG AGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAG ACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGG AGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGG CAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUC

		GACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCA GAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCA CACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUG AGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCA CGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAU ACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAG CAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACG UCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGAC AACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAA AGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAA GAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAA AGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUC GCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUA CAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACA GCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCAC AAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGC AUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAG GAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGU CCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAG GACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAA AGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAAC ACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAA

		AAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAU
351	mRNA encoding base editor with linker L024	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGGCGGCGGCGGCAGCGGAGGA GGAGGAAGCGAAGCCGCCGCCAAAGAGGCCGCCGCCAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGG AGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAG ACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGG AGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGG CAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUC GACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCA GAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCA CACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUG AGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCA CGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAU ACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAG CAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACG UCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGA

		CAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGAC AACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAA AGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAA GAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAA AGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUC GCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUA CAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACA GCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCAC AAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGC AUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAG GAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGU CCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAG GACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAA AGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAAC ACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAU
352	mRNA encoding base editor with linker L025	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGGCGGCGGCGGAAGCGAGGCC GCUGCCAAAGGCGGAGGAGGAAGCGGAGGAGGCGGCAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGG AGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAG ACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGA

		UCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGG AGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGG CAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUC GACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCA GAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCA CACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUG AGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCA CGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAU ACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAG CAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACG UCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGAC AACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAA AGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAA GAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAA AGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUC GCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUA CAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACA GCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCAC AAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGC AUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAG

		GAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGU CCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAG GACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAA AGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAAC ACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAU
353	mRNA encoding base editor with linker L036	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAGGCCGCCGCCAAGGAGGCC GCCGCCAAGGAGGCCGCCGCCAAGGAGGCCGCCGCCAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGG CAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGG AGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUC UGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAG ACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAA GAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUC GAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGA UCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGG AGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGAC GCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGG CAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAU CAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUC GACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGA UGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCA GAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUC CUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCA CACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAA CGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUG AGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGA AGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCA CGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUG UUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAU ACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGA CGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAG

		GGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAAC UGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAG CAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAG AACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACG UCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGA CAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGAC AACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAA AGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAA GAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUG AACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAA AGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAU CACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUC GCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCC UGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUA CAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGA AGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACA GCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUA CGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCAC AAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCG CAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGC AUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACA GGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAG GAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGU CCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAG GACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUAGACAUCACAUUUAAA AGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGUUGGUGUAAAGCCAAC ACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAAGAACCUCGAGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAU
354	mRNA encoding base editor with linker L051	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGGAGGAGGAGGAAGCGAAGCC GCCGCCAAAGAGGCCGCCGCCAAGGGAGGCGGCGGCAGCGAGGCCGCCGCCAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAA ACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAA

		GAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGA AGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCU UCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAU CUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGA GGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGU UCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAU CGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUC GACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACG CAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCU GAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUAC AAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGC CGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGG AAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAA AAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGA GCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGA CAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUAC GUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCA CAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAG CCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUC CUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGC UGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGA CUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCA CAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCA AGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAA GGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAA AACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGAC UGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAA CAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACA CAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAA CAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAA GGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCA CACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGG UCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUU CUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCA GCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAG CCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUC ACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCA AGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGC

		ACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUG UUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGG ACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCU GGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUC CACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCA UCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGA AAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGG GCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUA GACAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGU UGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAA GAACCUCGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAU
355	mRNA encoding base editor with linker L062	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAGGCCGCCGCCAAGGAGGCC GCCGCCAAGGGCGGCGGCGGCAGCGGCGGCGGCGGCAGCGGCGGCGGCGGCAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAA ACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAA GAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGA AGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCU UCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAU CUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGA GGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGU UCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAU CGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUC GACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACG CAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCU GAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUAC AAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGC CGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGG AAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAA AAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGA GCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGA CAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUAC GUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCA

		CAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAG CCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUC CUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGC UGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGA CUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCA CAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCA AGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAA GGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAA AACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGAC UGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAA CAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACA CAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAA CAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAA GGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCA CACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGG UCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUU CUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCA GCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAG CCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUC ACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCA AGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGC ACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUG UUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGG ACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCU GGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUC CACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCA UCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGA AAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGG GCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUA GACAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGU UGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAA GAACCUCGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAU
356	mRNA encoding base editor with linker L063	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG

		GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAAGCAGCAGCAAAGGAAGCA GCAGCAAAGGGAGGAGGAGGAAGCGGAGGAGGAGGAAGCGAAGCAGCAGCAAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAA ACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAA GAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGA AGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCU UCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAU CUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGA GGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGU UCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAU CGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUC GACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACG CAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCU GAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUAC AAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGC CGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGG AAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAA AAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGA GCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGA CAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUAC GUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCA CAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAG CCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUC CUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGC UGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGA CUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCA CAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCA AGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAA GGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAA AACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGAC UGAGCGACUACGACGUCGACCACACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAA CAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACA CAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAA CAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAA GGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCA CACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGG UCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUU CUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC

		AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCA GCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAG CCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUC ACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCA AGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGC ACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUG UUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGG ACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCU GGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUC CACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCA UCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGA AAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGG GCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUA GACAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGU UGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAA GAACCUCGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAU
357	mRNA encoding base editor with linker L066	GGGAGACCCAAGCUGGCUAGCGUUUAAACUUAAGCUUUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGUUGCAGGC CUUAUUCGGAUCCGCCACCAUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAAC GGAAUCGGAAGACACAAGACAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCC UGCACAACCAGGCAAAGAACCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGA CCCGGCACAGAUCUACAGAGUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAG GAAAACACACACGUCAGACUGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACG CAGGAGCACAGGUCAGCAUCAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUG GGACGGACUGGACGAACACAGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAAGCAGCAGCAAAGGAAGCA GCAGCAAAGGAAGCAGCAGCAAAGGGAGGAGGAGGAAGCGGAGGAGGAGGAAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAA ACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAA GAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGA AGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCU UCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAU CUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGA GGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGU UCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAU CGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUC GACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACG CAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCU GAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUAC AAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGC CGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGG

		AAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAA AAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGA GCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGA CAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUAC GUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCA CAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAG CCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUC CUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGC UGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGA CUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCA CAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCA AGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAA GGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAA AACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGAC UGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAA CAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACA CAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAA CAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAA GGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCA CACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGG UCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUU CUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGAC AAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCA GCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAG CCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUC ACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCA AGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGC ACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUG UUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGG ACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCU GGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUC CACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCA UCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGA AAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGG GCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAGUCUA GACAUCACAUUUAAAAGCAUCUCAGCCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUUUUUCUUUUUCGU UGGUGUAAAGCCAACACCCUGUCUAAAAAACAUAAAUUUCUUUAAUCAUUUUGCCUCUUUUCUCUGUGCUUCAAUUAAUAAAAAAUGGAAA GAACCUCGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCGAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAU

358-359		Not Used
360	Open reading frame for Nme2 Cas9	atgGACGGCTCCGGCGGCGGCTCCCCCAAGAAGAAGCGGAAGGTGGAGGACAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAGA AGAAGAAGGGCGGCTCCGGCGGCGGCgccgccttcaagcccaaccccatcaactacatcctgggcctggacatcggcatcgcctccgtggg ctgggccatggtggagatcgacgaggaggagaaccccatccggctgatcgacctgggcgtgcgggtgttcgagcgggccgaggtgcccaag accggcgactccctggccatggcccggcggctggcccggtccgtgcggcggctgacccggcggcgggcccaccggctgctgcgggcccggc ggctgctgaagcgggagggcgtgctgcaggccgccgacttcgacgagaacggcctgatcaagtccctgcccaacacccccctggcagctgcg ggccgccgccctggaccggaagctgacccccctggagtggtccgccgtgctgctgcacctgatcaagcaccggggctacctgtcccagcgg aagaacgagggcgagaccgccgacaaggagctgggcgccctgctgaagggcgtggccaacaacgcccacgccctgcagaccggcgacttcc ggacccccgccgagctggccctgaacaagttcgagaaggagtccggccacatccggaaccagcggggcgactactcccacaccttctcccg gaaggacctgcaggccgagctgatcctgctgttcgagaagcagaaggagttcggcaacccccacgtgtccggcggcctgaaggagggcatc gagaccctgctgatgacccagcggcccgccctgtccggcgacgccgtgcagaagatgctgggccactgcaccttcgagcccgccgagccca aggccgccaagaacacctacaccgccgagcggttcatctggctgaccaagctgaacaacctgcggatcctggagcagggctccgagcggcc cctgaccgacaccgagcgggccaccctgatggacgagccctaccggaagtccaagctgacctacgcccaggcccggaagctgctgggcctg gaggacaccgccttcttcaagggcctgcggtacggcaaggacaacgccgaggcctccaccctgatggagatgaaggcctaccacgccatct cccgggccctggagaaggagggcctgaaggacaagaagtccccccctgaacctgtcctccgagctgcaggacgagatcggcaccgccttctc cctgttcaagaccgacgaggacatcaccggccggctgaaggaccgggtgcagcccgagatcctggaggccctgctgaagcacatctccttc gacaagttcgtgcagatctccctgaaggccctgcggcggatcgtgcccctgatggagcagggcaagcggtacgacgaggcctgcgccgaga tctacggcgaccactacggcaagaagaacaccgaggagaagatctacctgccccccatccccgccgacgagatccggaaccccgtggtgct gcgggccctgtcccaggcccggaaggtgatcaacggcgtggtgcggcggtacggctccccccgcccggatccacatcgagaccgcccgggag gtgggcaagtccttcaaggaccggaaggagatcgagaagcggcaggaggagaaccggaaggaccgggagaaggccgccgccaagttccggg agtacttccccaacttcgtgggcgagcccaagtccaaggacatcctgaagctgcggctgtacgagcagcagcacggcaagtgcctgtactc cggcaaggagatcaacctggtgcggctgaacgagaagggctacgtggagatcgaccacgccctgcccttctcccggacctgggacgactcc ttcaacaacaaggtgctggtgctgggctccgagaaccagaacaagggcaaccagacccccctacgagtacttcaacggcaaggacaactccc gggagtggcaggagttcaaggcccgggtggagacctcccggttcccccggtccaagaagcagcggatcctgctgcagaagttcgacgagga cggcttcaaggagtgcaacctgaacgacacccggtacgtgaaccggttcctgtgccagttcgtggccgaccacatcctgctgaccggcaag ggcaagcggcgggtgttcgcctccaacggccagatcaccaacctgctgcggggcttctggggcctgcggaaggtgcgggccgagaacgacc ggcaccacgccctggacgccgtggtggtggcctgctccaccgtggccatgcagcagaagatcacccggttcgtgcggtacaaggagatgaa cgccttcgacggcaagaccatcgacaaggagaccggcaaggtgctgcaccagaagacccacttcccccagccctgggagttcttcgcccag gaggtgatgatccgggtgttcggcaagcccgacggcaagcccgagttcgaggaggccgacacccccgagaagctgcggaccctgctggccg agaagctgtcctcccggcccgaggccgtgcacgagtacgtgacccccctgttcgtgtcccgggcccccaaccggaagatgtccggcgccca caaggacaccctgcggtccgccaagcggttcgtgaagcacaacgagaagatctccgtgaagcgggtgtggctgaccgagatcaagctggcc gacctggagaacatggtgaactacaagaacggccgggagatcgagctgtacgaggccctgaaggcccggctggaggcctacggcggcaacg ccaagcaggccttcgaccccaaggacaacccccttctacaagaagggcggccagctggtgaaggccgtgcgggtggagaagacccaggagtc cggcgtgctgctgaacaagaagaacgcctacaccatcgccgacaacggcgacatggtgcgggtggacgtgttctgcaaggtggacaagaag ggcaagaaccagtacttcatcgtgcccatctacgcctggcaggtggccgagaacatcctgcccgacatcgactgcaagggctaccggatcg acgactcctacaccttctgcttctccctgcacaagtacgacctgatcgccttccagaaggacgagaagtccaaggtggagttcgcctacta catcaactgcgactcctccaacggccggttctacctggcctggcacgacaagggctccaaggagcagcagttccggatctccacccagaac ctggtgctgatccagaagtaccaggtgaacgagctgggcaaggagatccggccctgccggctgaagaagcggccccccgtgcggtag

361	Open reading frame for base editor with linker L070	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGA CCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAA CCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGG GUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGC UGCGGAUCUUCGCCGCCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAU CAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACggcaccaaggacuccaccaaggacaucccccgagacccccu ccaaggacGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUC CAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCC GAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGA UGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGG CAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACG UGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAU CCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUG AUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACG ACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCU GUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACA UCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCU GAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUG CGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCC CCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAA GGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGA AGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUU CGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGAC UUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGC UGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCU GAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACC UGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGA GAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGC AUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGA ACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAA GGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGU CCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAU GAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGAC

		UUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGA AGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGG CAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCC AGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGC CCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAG GGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCC UGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCG GAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGA UCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCG GGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGG AAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCC AGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGA
362	Open reading frame for base editor with linker L071	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGA CCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAA CCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGG GUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGC UGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAU CAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACggccgggacgugcggcagcccgaggugaaggaggagaaagc ccgaguccGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUC CAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCC GAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGA UGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGG CAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACG UGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAU CCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUG AUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACG ACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCU GUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACA UCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCU GAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUG CGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCC CCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAA GGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGA

		AGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUU CGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGAC UUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGC UGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCU GAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACC UGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGA GAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGC AUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGA ACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAA GGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGU CCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAU GAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGAC UUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGA AGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGG CAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCC AGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGC CCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAG GGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCC UGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCG GAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGA UCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCG GGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGG AAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCC AGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGA
363	Open reading frame for base editor with linker L072	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGA CCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAA CCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGG GUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGC UGCGGAUCUUCGCCGCCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAU CAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACgagggcaaguccuccggcuccggcuccgaguccaagucca ccgccggcGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUC CAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCC GAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGA UGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGG

		CAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACG UGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAU CCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUG AUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACG ACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCU GUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACA UCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCU GAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUG CGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCC CCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAA GGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGA AGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUU CGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGAC UUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGC UGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCU GAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACC UGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGA GAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGC AUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGA ACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAA GGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGU CCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAU GAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGAC UUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGA AGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGG CAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCC AGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGC CCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAG GGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCC UGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCG GAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGA UCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCG

		GGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGG AAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCC AGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGA
364	Open reading frame for base editor with linker L073	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGA CCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAA CCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGG GUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGC UGCGGAUCUUCGCCGCCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAU CAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACaccccggcucccccgccggcucccccaccuccaccgaggg agggcaccGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUC CAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCC GAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGA UGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGG CAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACG UGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAU CCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUG AUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACG ACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCU GUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACA UCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCU GAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUG CGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCC CCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAA GGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGA AGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUU CGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGAC UUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGC UGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCU GAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACC UGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGA GAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGC AUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGA ACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAA GGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG

		AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGU CCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAU GAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGAC UUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGA AGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGG CAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCC AGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGC CCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAG GGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCC UGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCG GAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGA UCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCG GGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGG AAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCC AGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGA
365	Open reading frame for base editor with linker L074	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGA CCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAA CCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGG GUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGC UGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAU CAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACggcuccgagcccgccaccuccggcuccgagacccccggca ccuccaccGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUC CAAGAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCC GAGGCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGA UGGCCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGG CAACAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGAC CUGCGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACG UGGACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAU CCUGUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUG AUCGCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACG ACGACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCU GUCCGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUG ACCCUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACA UCGACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCU GAACCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUG CGGCGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCC

		CCCUGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAA GGGCGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUG UACGAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGA AGAAGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUU CGACUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGAC UUCCUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGC UGAAGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCU GAUCAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUG AUCCACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACC UGGCCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGA GAACAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGC AUCAAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGA ACGGCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAA GGACGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAG AUGAAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGU CCGAGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAU GAACACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGAC UUCCAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGA AGUACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGG CAAGGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGG CCCCUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCC AGGUGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGC CCGGAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAG GGCAAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCC UGGAGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCG GAUGCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUAC GAGAAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGA UCUCCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCG GGAGCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGG AAGCGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCC AGCUGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGA
366	Open reading frame for base editor with linker L019	AUGGAGGCCUCCCCCGCCUCCGGCCCCCGGCACCUGAUGGACCCCCACAUCUUCACCUCCAACUUCAACAACGGCAUCGGCCGGCACAAGA CCUACCUGUGCUACGAGGUGGAGCGGCUGGACAACGGCACCUCCGUGAAGAUGGACCAGCACCGGGGCUUCCUGCACAACCAGGCCAAGAA CCUGCUGUGCGGCUUCUACGGCCGGCACGCCGAGCUGCGGUUCCUGGACCUGGUGCCCUCCCUGCAGCUGGACCCCGCCCAGAUCUACCGG GUGACCUGGUUCAUCUCCUGGUCCCCCUGCUUCUCCUGGGGCUGCGCCGGCGAGGUGCGGGCCUUCCUGCAGGAGAACACCCACGUGCGGC UGCGGAUCUUCGCCGCCCGGAUCUACGACUACGACCCCCUGUACAAGGAGGCCCUGCAGAUGCUGCGGGACGCCGGCGCCCAGGUGUCCAU CAUGACCUACGACGAGUUCAAGCACUGCUGGGACACCUUCGUGGACCACCAGGGCUGCCCCUUCCAGCCCUGGGACGGCCUGGACGAGCAC UCCCAGGCCCUGUCCGGCCGGCUGCGGGCCAUCCUGCAGAACCAGGGCAACgaggccgccgccaaggaggccgccgccaaggaggccgccg ccaagGACAAGAAGUACUCCAUCGGCCUGGCCAUCGGCACCAACUCCGUGGGCUGGGCCGUGAUCACCGACGAGUACAAGGUGCCCUCCAA

		GAAGUUCAAGGUGCUGGGCAACACCGACCGGCACUCCAUCAAGAAGAACCUGAUCGGCGCCCUGCUGUUCGACUCCGGCGAGACCGCCGAG GCCACCCGGCUGAAGCGGACCGCCCGGCGGCGGUACACCCGGCGGAAGAACCGGAUCUGCUACCUGCAGGAGAUCUUCUCCAACGAGAUGG CCAAGGUGGACGACUCCUUCUUCCACCGGCUGGAGGAGUCCUUCCUGGUGGAGGAGGACAAGAAGCACGAGCGGCACCCCAUCUUCGGCAA CAUCGUGGACGAGGUGGCCUACCACGAGAAGUACCCCACCAUCUACCACCUGCGGAAGAAGCUGGUGGACUCCACCGACAAGGCCGACCUG CGGCUGAUCUACCUGGCCCUGGCCCACAUGAUCAAGUUCCGGGGCCACUUCCUGAUCGAGGGCGACCUGAACCCCGACAACUCCGACGUGG ACAAGCUGUUCAUCCAGCUGGUGCAGACCUACAACCAGCUGUUCGAGGAGAACCCCAUCAACGCCUCCGGCGUGGACGCCAAGGCCAUCCU GUCCGCCCGGCUGUCCAAGUCCCGGCGGCUGGAGAACCUGAUCGCCCAGCUGCCCGGCGAGAAGAAGAACGGCCUGUUCGGCAACCUGAUC GCCCUGUCCCUGGGCCUGACCCCCAACUUCAAGUCCAACUUCGACCUGGCCGAGGACGCCAAGCUGCAGCUGUCCAAGGACACCUACGACG ACGACCUGGACAACCUGCUGGCCCAGAUCGGCGACCAGUACGCCGACCUGUUCCUGGCCGCCAAGAACCUGUCCGACGCCAUCCUGCUGUC CGACAUCCUGCGGGUGAACACCGAGAUCACCAAGGCCCCCCUGUCCGCCUCCAUGAUCAAGCGGUACGACGAGCACCACCAGGACCUGACC CUGCUGAAGGCCCUGGUGCGGCAGCAGCUGCCCGAGAAGUACAAGGAGAUCUUCUUCGACCAGUCCAAGAACGGCUACGCCGGCUACAUCG ACGGCGGCGCCUCCCAGGAGGAGUUCUACAAGUUCAUCAAGCCCAUCCUGGAGAAGAUGGACGGCACCGAGGAGCUGCUGGUGAAGCUGAA CCGGGAGGACCUGCUGCGGAAGCAGCGGACCUUCGACAACGGCUCCAUCCCCCACCAGAUCCACCUGGGCGAGCUGCACGCCAUCCUGCGG CGGCAGGAGGACUUCUACCCCUUCCUGAAGGACAACCGGGAGAAGAUCGAGAAGAUCCUGACCUUCCGGAUCCCCUACUACGUGGGCCCCC UGGCCCGGGGCAACUCCCGGUUCGCCUGGAUGACCCGGAAGUCCGAGGAGACCAUCACCCCCUGGAACUUCGAGGAGGUGGUGGACAAGGG CGCCUCCGCCCAGUCCUUCAUCGAGCGGAUGACCAACUUCGACAAGAACCUGCCCAACGAGAAGGUGCUGCCCAAGCACUCCCUGCUGUAC GAGUACUUCACCGUGUACAACGAGCUGACCAAGGUGAAGUACGUGACCGAGGGCAUGCGGAAGCCCGCCUUCCUGUCCGGCGAGCAGAAGA AGGCCAUCGUGGACCUGCUGUUCAAGACCAACCGGAAGGUGACCGUGAAGCAGCUGAAGGAGGACUACUUCAAGAAGAUCGAGUGCUUCGA CUCCGUGGAGAUCUCCGGCGUGGAGGACCGGUUCAACGCCUCCCUGGGCACCUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUC CUGGACAACGAGGAGAACGAGGACAUCCUGGAGGACAUCGUGCUGACCCUGACCCUGUUCGAGGACCGGGAGAUGAUCGAGGAGCGGCUGA AGACCUACGCCCACCUGUUCGACGACAAGGUGAUGAAGCAGCUGAAGCGGCGGCGGUACACCGGCUGGGGCCGGCUGUCCCGGAAGCUGAU CAACGGCAUCCGGGACAAGCAGUCCGGCAAGACCAUCCUGGACUUCCUGAAGUCCGACGGCUUCGCCAACCGGAACUUCAUGCAGCUGAUC CACGACGACUCCCUGACCUUCAAGGAGGACAUCCAGAAGGCCCAGGUGUCCGGCCAGGGCGACUCCCUGCACGAGCACAUCGCCAACCUGG CCGGCUCCCCCGCCAUCAAGAAGGGCAUCCUGCAGACCGUGAAGGUGGUGGACGAGCUGGUGAAGGUGAUGGGCCGGCACAAGCCCGAGAA CAUCGUGAUCGAGAUGGCCCGGGAGAACCAGACCACCCAGAAGGGCCAGAAGAACUCCCGGGAGCGGAUGAAGCGGAUCGAGGAGGGCAUC AAGGAGCUGGGCUCCCAGAUCCUGAAGGAGCACCCCGUGGAGAACACCCAGCUGCAGAACGAGAAGCUGUACCUGUACUACCUGCAGAACG GCCGGGACAUGUACGUGGACCAGGAGCUGGACAUCAACCGGCUGUCCGACUACGACGUGGACCACAUCGUGCCCCAGUCCUUCCUGAAGGA CGACUCCAUCGACAACAAGGUGCUGACCCGGUCCGACAAGAACCGGGGCAAGUCCGACAACGUGCCCUCCGAGGAGGUGGUGAAGAAGAUG AAGAACUACUGGCGGCAGCUGCUGAACGCCAAGCUGAUCACCCAGCGGAAGUUCGACAACCUGACCAAGGCCGAGCGGGGCGGCCUGUCCG AGCUGGACAAGGCCGGCUUCAUCAAGCGGCAGCUGGUGGAGACCCGGCAGAUCACCAAGCACGUGGCCCAGAUCCUGGACUCCCGGAUGAA CACCAAGUACGACGAGAACGACAAGCUGAUCCGGGAGGUGAAGGUGAUCACCCUGAAGUCCAAGCUGGUGUCCGACUUCCGGAAGGACUUC CAGUUCUACAAGGUGCGGGAGAUCAACAACUACCACCACGCCCACGACGCCUACCUGAACGCCGUGGUGGGCACCGCCCUGAUCAAGAAGU ACCCCAAGCUGGAGUCCGAGUUCGUGUACGGCGACUACAAGGUGUACGACGUGCGGAAGAUGAUCGCCAAGUCCGAGCAGGAGAUCGGCAA GGCCACCGCCAAGUACUUCUUCUACUCCAACAUCAUGAACUUCUUCAAGACCGAGAUCACCCUGGCCAACGGCGAGAUCCGGAAGCGGCCC CUGAUCGAGACCAACGGCGAGACCGGCGAGAUCGUGUGGGACAAGGGCCGGGACUUCGCCACCGUGCGGAAGGUGCUGUCCAUGCCCCAGG UGAACAUCGUGAAGAAGACCGAGGUGCAGACCGGCGGCUUCUCCAAGGAGUCCAUCCUGCCCAAGCGGAACUCCGACAAGCUGAUCGCCCG GAAGAAGGACUGGGACCCCAAGAAGUACGGCGGCUUCGACUCCCCCACCGUGGCCUACUCCGUGCUGGUGGUGGCCAAGGUGGAGAAGGGC AAGUCCAAGAAGCUGAAGUCCGUGAAGGAGCUGCUGGGCAUCACCAUCAUGGAGCGGUCCUCCUUCGAGAAGAACCCCAUCGACUUCCUGG AGGCCAAGGGCUACAAGGAGGUGAAGAAGGACCUGAUCAUCAAGCUGCCCAAGUACUCCCUGUUCGAGCUGGAGAACGGCCGGAAGCGGAU

		GCUGGCCUCCGCCGGCGAGCUGCAGAAGGGCAACGAGCUGGCCCUGCCCUCCAAGUACGUGAACUUCCUGUACCUGGCCUCCCACUACGAG AAGCUGAAGGGCUCCCCCGAGGACAACGAGCAGAAGCAGCUGUUCGUGGAGCAGCACAAGCACUACCUGGACGAGAUCAUCGAGCAGAUCU CCGAGUUCUCCAAGCGGGUGAUCCUGGCCGACGCCAACCUGGACAAGGUGCUGUCCGCCUACAACAAGCACCGGGACAAGCCCAUCCGGGA GCAGGCCGAGAACAUCAUCCACCUGUUCACCCUGACCAACCUGGGCGCCCCCGCCGCCUUCAAGUACUUCGACACCACCAUCGACCGGAAG CGGUACACCUCCACCAAGGAGGUGCUGGACGCCACCCUGAUCCACCAGUCCAUCACCGGCCUGUACGAGACCCGGAUCGACCUGUCCCAGC UGGGCGGCGACGGCGGCGGCUCCCCCAAGAAGAAGCGGAAGGUGUGA
367	Open reading frame for base editor with linker L015	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGGAGGAGGAGGAAGCGAAGCAGCAGCAAAGGAAGCAGCAG CAAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAA GAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAA GCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGG CAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAA CAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUG AGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCG ACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCU GAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUC GCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACG ACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAG CGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACA CUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCG ACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAA CAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGA AGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGC UGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUAC GAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGA AGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGA CAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUC CUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGA AGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAU CAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUC CACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGG CAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAA CAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUC

		AAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACG GAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGA CGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUG AAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCG AACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAA CACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGU ACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAA GGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCG CUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGG UCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAG AAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGA AAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGG AAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAU GCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCA GCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGA ACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAG AGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGC UGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAU GCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAAC GUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGA GCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUA
368	Open reading frame for base editor with linker L016	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACgaggccgccgccaagggcggcggcggcagcggcggcggcg gcagcGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAA GAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAA GCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGG CAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAA CAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUG AGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCG ACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCU GAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUC GCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACG

		ACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAG CGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACA CUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCG ACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAA CAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGA AGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGC UGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUAC GAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGA AGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGA CAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUC CUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGA AGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAU CAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUC CACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGG CAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAA CAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUC AAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACG GAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGA CGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUG AAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCG CGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUG AAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCG AACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAA CACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGU ACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAA GGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCG CUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGG UCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAG AAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGA AAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGG AAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAU GCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCA GCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGA ACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAG AGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGC UGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAU GCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAAC GUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGA

		GCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAG
369	Open reading frame for base editor with linker L019	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACgaggccgccgccaaggaggccgccgccaaggaggccgccg ccaagGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUACAAGGUCCCGAGCAA GAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAA GCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCUUCAGCAACGAAAUGG CAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCUUCGGAAA CAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUG AGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGGACAACAGCGACGUCG ACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCAAUCCU GAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUC GCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCAAGGACACAUACGACG ACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGCUGAG CGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACA CUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAUACGCAGGAUACAUCG ACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACUGCUGGUCAAGCUGAA CAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGA AGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGUACUACGUCGGACCGC UGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGACAAGGG AGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUAC GAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGAGCGGAGAACAGAAGA AGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUCGAAUGCUUCGA CAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUC CUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGAUCGAAGAAAGACUGA AGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGAAAGCUGAU CAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUC CACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAACACAUCGCAAACCUGG CAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAAGCCGGAAAA CAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUC AAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGUACUACCUGCAGAACG GAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUUCCUGAAGGA CGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUG AAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGAGAGGAGGACUGAGCG AACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCUGGACAGCAGAAUGAA

		CACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUC CAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAGCACUGAUCAAGAAGU ACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAGGAAAUCGGAAA GGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCG CUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCCUGAGCAUGCCGCAGG UCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGACAAGCUGAUCGCAAG AAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGA AAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACCCGAUCGACUUCCUGG AAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAACGGAAGAAAGAGAAU GCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUGGCAAGCCACUACGAA AAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAAUCAUCGAACAGAUCA GCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGACAAGCCGAUCAGAGA ACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACAACAAUCGACAGAAAG AGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAAUCGAUCUGAGCCAGC UGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCAGGAAAGCAUCCUGAU GCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAAAGCACAGACGAAAAC GUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACAAGAUCAAGAUGCUGA GCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAG
370	Open reading frame for base editor with linker L021	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACggcggcggcggcagcggcggcggcggcagcggcggcggcg gcagcggcggcggcggcagcGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUA CAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGC GGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACA CCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACA GACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGG ACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGA CGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUG UUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCA AGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGA CGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACAC CACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACU GCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUG

		CACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGU ACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGA AGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAG CACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGA GCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAA GAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUC AAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGA UCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACU GAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAAC UUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAAC ACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAG ACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGA AUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGU ACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAA GUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGA GAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCU GGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGAC UUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAG CACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGA ACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAA AUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCC UGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGA CAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCA AAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACC CGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAA CGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUG GCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAA UCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGA CAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACA ACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAA UCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCA GGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAA AGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACA AGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAAUAG
371	Open reading frame for base editor with linker L024	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC

		UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACggcggcggcggcagcggaggaggaggaagcgaagccgccg ccaaagaggccgccgccaagGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUA CAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGC GGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACA CCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACA GACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGG ACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGA CGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUG UUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCA AGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGA CGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACAC CACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACU GCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUG CACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGU ACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGA AGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAG CACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGA GCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAA GAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUC AAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGA UCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACU GAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAAC UUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAAC ACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAG ACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGA AUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGU ACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAA GUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGA GAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCU GGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGAC UUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAG CACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGA ACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAA AUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCC

		UGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGA CAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCA AAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACC CGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAA CGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUG GCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAA UCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGA CAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACA ACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAA UCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCA GGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAA AGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACA AGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
372	Open reading frame for base editor with linker L025	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACggcggcggcggaagcgaggccgcugccaaaggcggaggag gaagcggaggaggcggcagcGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUA CAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGC GGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACA CCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACA GACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGG ACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGA CGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUG UUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCA AGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGA CGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACAC CACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACU GCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUG CACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGU ACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGA AGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAG CACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGA GCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAA

		GAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUC AAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGA UCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACU GAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAAC UUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAAC ACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAG ACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGA AUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGU ACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAA GUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGA GAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCU GGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGAC UUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAG CACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGA ACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAA AUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCC UGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGA CAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCA AAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACC CGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAA CGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUG GCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAA UCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGA CAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACA ACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAA UCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCA GGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAA AGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACA AGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
373	Open reading frame for base editor with linker L036	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACgaggccgccgccaaggaggccgccgccaaggaggccgccg ccaaggaggccgccgccaagGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGUCAUCACAGACGAAUA CAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCACUGCUGUUCGACAGC

		GGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUGCAGGAAAUCU UCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACA CCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAGCUGGUCGACAGCACA GACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAGGAGACCUGAACCCGG ACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGA CGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAAAAGAAGAACGGACUG UUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCAGCUGAGCA AGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGCAAAGAACCUGAGCGA CGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAGAGAUACGACGAACAC CACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAGAACGGAU ACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAACAGAAGAACU GCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUCCACCUGGGAGAACUG CACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGACAUUCAGAAUCCCGU ACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGA AGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAAAAGGUCCUGCCGAAG CACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCGGCAUUCCUGA GCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAA GAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGACCUGCUGAAGAUCAUC AAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGACAGAGAAAUGA UCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACU GAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGAUUCGCAAACAGAAAC UUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACAGCCUGCACGAAC ACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAG ACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAGAAAGAAUGAAGAGA AUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAAAAGCUGUACCUGU ACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCA GAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGUCCCGAGCGAAGAA GUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUGACAAAGGCAGAGA GAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCACGUCGCACAGAUCCU GGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGCAAGCUGGUCAGCGAC UUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCAGUCGUCGGAACAG CACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGA ACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACACUGGCAAACGGAGAA AUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAACAGUCAGAAAGGUCC UGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGA CAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCA AAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCAGCUUCGAAAAGAACC CGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAAA CGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUCAACUUCCUGUACCUG

		GCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGCACUACCUGGACGAAA UCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAGA CAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUCAAGUACUUCGACACA ACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGACUGUACGAAACAAGAA UCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAAGCAGCUGGUCAUCCA GGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCACACAGCAUACGACGAA AGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACAGCAACGGAGAAAACA AGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
374	Open reading frame for base editor with linker L051	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACggaggaggaggaagcgaagccgccgccaaagaggccgccg ccaagggaggcggcggcagcgaggccgccgccaagGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGU CAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCA CUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCU ACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAA GAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAG CUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAG GAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAA CGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAA AAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAA AGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGC AAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAG AGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACC AGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGA CGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUC CACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGA CAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACC GUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAA AAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAA AGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGA AGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGAC CUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCG AAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACAC AGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGA UUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAG

		ACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGU CAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGA GAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACG AAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGA CCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAAC GUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACC UGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCA CGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGC AAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAU GAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACA CUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAA CAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCC GAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGC GUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCA GCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCU GUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUC AACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGC ACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUA CAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUC AAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGAC UGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAA GCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCAC ACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACA GCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
375	Open reading frame for base editor with linker L062	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACgaggccgccgccaaggaggccgccgccaagggcggcggcg gcagcggcggcggcggcagcggcggcggcggcagcGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGU CAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCA CUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCU ACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAA GAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAG CUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAG GAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAA

		CGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAA AAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAA AGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGC AAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAG AGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACC AGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGA CGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUC CACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGA CAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACC GUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAA AAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAA AGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGA AGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGAC CUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCG AAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACAC AGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGA UUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAG ACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGU CAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGA GAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACG AAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGA CCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAAC GUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACC UGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCA CGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGC AAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAU GAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACA CUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAA CAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCC GAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGC GUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCA GCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCU GUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUC AACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGC ACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUA CAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUC AAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGAC UGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAA

		GCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCAC ACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACA GCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
376	Open reading frame for base editor with linker L063	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAAGCAGCAGCAAAGGAAGCAGCAGCAAAGGGAGGAGGAG GAAGCGGAGGAGGAGGAAGCGAAGCAGCAGCAAAGGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGU CAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCA CUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCU ACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAA GAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAG CUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAG GAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAA CGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAA AAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAA AGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGC AAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAG AGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACC AGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGA CGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUC CACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGA CAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACC GUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAA AAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAA AGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGA AGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGAC CUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCG AAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACAC AGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGA UUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAG ACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGU CAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGA GAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACG AAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGA CCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAAC

		GUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACC UGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCA CGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGC AAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAU GAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACA CUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAA CAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCC GAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGC GUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCA GCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCU GUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUC AACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGC ACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUA CAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUC AAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGAC UGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAA GCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCAC ACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACA GCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
377	Open reading frame for base editor with linker L066	AUGGAAGCAAGCCCGGCAAGCGGACCGAGACACCUGAUGGACCCGCACAUCUUCACAAGCAACUUCAACAACGGAAUCGGAAGACACAAGA CAUACCUGUGCUACGAAGUCGAAAGACUGGACAACGGAACAAGCGUCAAGAUGGACCAGCACAGAGGAUUCCUGCACAACCAGGCAAAGAA CCUGCUGUGCGGAUUCUACGGAAGACACGCAGAACUGAGAUUCCUGGACCUGGUCCCGAGCCUGCAGCUGGACCCGGCACAGAUCUACAGA GUCACAUGGUUCAUCAGCUGGAGCCCGUGCUUCAGCUGGGGAUGCGCAGGAGAAGUCAGAGCAUUUCUGCAGGAAAACACACACGUCAGAC UGAGAAUCUUCGCAGCAAGAAUCUACGACUACGACCCGCUGUACAAGGAAGCACUGCAGAUGCUGAGAGACGCAGGAGCACAGGUCAGCAU CAUGACAUACGACGAAUUCAAGCACUGCUGGGACACAUUCGUCGACCACCAGGGAUGCCCGUUCCAGCCGUGGGACGGACUGGACGAACAC AGCCAGGCACUGAGCGGAAGACUGAGAGCAAUCCUGCAGAACCAGGGAAACGAAGCAGCAGCAAAGGAAGCAGCAGCAAAGGAAGCAGCAG CAAAGGGAGGAGGAGGAAGCGGAGGAGGAGGAAGCGACAAGAAGUACAGCAUCGGACUGGCCAUCGGAACAAACAGCGUCGGAUGGGCAGU CAUCACAGACGAAUACAAGGUCCCGAGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAACCUGAUCGGAGCA CUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUGAAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCU ACCUGCAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUUCCACAGACUGGAAGAAAGCUUCCUGGUCGAAGAAGACAA GAAGCACGAAAGACACCCGAUCUUCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCACCUGAGAAAGAAG CUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUACCUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUUCCUGAUCGAAG GAGACCUGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUACAACCAGCUGUUCGAAGAAAACCCGAUCAA CGCAAGCGGAGUCGACGCAAAGGCAAUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGCUGCCGGGAGAA AAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGGGACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAA AGCUGCAGCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGAGACCAGUACGCAGACCUGUUCCUGGCAGC AAAGAACCUGAGCGACGCAAUCCUGCUGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAGCAUGAUCAAG AGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUGGUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACC

		AGAGCAAGAACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAGUUCAUCAAGCCGAUCCUGGAAAAGAUGGA CGGAACAGAAGAACUGCUGGUCAAGCUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUCCCGCACCAGAUC CACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACUUCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCCUGA CAUUCAGAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAUGACAAGAAAGAGCGAAGAAACAAUCACACC GUGGAACUUCGAAGAAGUCGUCGACAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAACCUGCCGAACGAA AAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACAGUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAA AGCCGGCAUUCCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACAAACAGAAAGGUCACAGUCAAGCAGCUGAAGGA AGACUACUUCAAGAAGAUCGAAUGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGGGAACAUACCACGAC CUGCUGAAGAUCAUCAAGGACAAGGACUUCCUGGACAACGAAGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCG AAGACAGAGAAAUGAUCGAAGAAAGACUGAAGACAUACGCACACCUGUUCGACGACAAGGUCAUGAAGCAGCUGAAGAGAAGAAGAUACAC AGGAUGGGGAAGACUGAGCAGAAAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUCCUGAAGAGCGACGGA UUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACAGCCUGACAUUCAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAG ACAGCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAUCCUGCAGACAGUCAAGGUCGUCGACGAACUGGU CAAGGUCAUGGGAAGACACAAGCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGA GAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACUGGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACG AAAAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAACUGGACAUCAACAGACUGAGCGACUACGACGUCGA CCACAUCGUCCCGCAGAGCUUCCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAAC GUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUACUGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACC UGACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAGAGACAGCUGGUCGAAACAAGACAGAUCACAAAGCA CGUCGCACAGAUCCUGGACAGCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGGUCAUCACACUGAAGAGC AAGCUGGUCAGCGACUUCAGAAAGGACUUCCAGUUCUACAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACG CAGUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUCUACGGAGACUACAAGGUCUACGACGUCAGAAAGAU GAUCGCAAAGAGCGAACAGGAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUCUUCAAGACAGAAAUCACA CUGGCAAACGGAGAAAUCAGAAAGAGACCGCUGAUCGAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCAA CAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAAGUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCC GAAGAGAAACAGCGACAAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGACAGCCCGACAGUCGCAUACAGC GUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGAGCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAAGCA GCUUCGAAAAGAACCCGAUCGACUUCCUGGAAGCAAAGGGAUACAAGGAAGUCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCU GUUCGAACUGGAAAACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAACUGGCACUGCCGAGCAAGUACGUC AACUUCCUGUACCUGGCAAGCCACUACGAAAAGCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCACAAGC ACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUCAUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUA CAACAAGCACAGAGACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAACCUGGGAGCACCGGCAGCAUUC AAGUACUUCGACACAACAAUCGACAGAAAGAGAUACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAGGAC UGUACGAAACAAGAAUCGAUCUGAGCCAGCUGGGAGGAGACAGCGGAGGAAGCACAAACCUGAGCGACAUCAUCGAAAAGGAAACAGGAAA GCAGCUGGUCAUCCAGGAAAGCAUCCUGAUGCUGCCGGAAGAAGUCGAAGAAGUCAUCGGAAACAAGCCGGAAAGCGACAUCCUGGUCCAC ACAGCAUACGACGAAAGCACAGACGAAAACGUCAUGCUGCUGACAAGCGACGCACCGGAAUACAAGCCGUGGGCACUGGUCAUCCAGGACA GCAACGGAGAAAACAAGAUCAAGAUGCUGAGCGGAGGAAGCCCGAAGAAGAAGAGAAAGGUCUAA
378-386 Not Used		
387	amino acid	MAAFKPNPINYILGLAIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGV

	sequence for D16A Nme2Cas9 nickase	LQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVANNAHALQTGDFRTPAELAL NKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYT AERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEG LKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGK KNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVG EPKSKDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKA RVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAV VVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPE AVHEYVTPLFVSRAPNRKMSGAHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIELYEALKARLEAYGGNAKQAFDPK DNPFYKKGGQLVKAVRVEKTQESGVLLNKKNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKGYRIDDSYTFCF SLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGSKEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVR
388	coding sequence for D16A Nme2Cas9 nickase	GCTGCTTTTAAGCCTAATCCTATTAATTATATTCTTGGTCTTGCTATTGGTATTGCTTCTGTTGGTTGGGCTATGGTTGAGATTGATGAGG AGGAGAATCCTATTCGTCTTATTGATCTTGGTGTTCGTGTTTTTGAGCGTGCTGAGGTTCCTAAGACTGGTGATTCTCTTGCTATGGCTCG TCGTCTTGCTCGTTCTGTTCGTCGTCTTACTCGTCGTCGTGCTCATCGTCTTCTTCGTGCTCGTCGTCTTCTTAAGCGTGAGGGTGTTCTT CAGGCTGCTGATTTTGATGAGAATGGTCTTATTAAGTCTCTTCCTAATACTCCTTGGCAGCTTCGTGCTGCTGCTCTTGATCGTAAGCTTA CTCCTCTTGAGTGGTCTGCTGTTCTTCTTCATCTTATTAAGCATCGTGGTTATCTTTCTCAGCGTAAGAATGAGGGTGAGACTGCTGATAA GGAGCTTGGTGCTCTTCTTAAGGGTGTTGCTAATAATGCTCATGCTCTTCAGACTGGTGATTTTCGTACTCCTGCTGAGCTTGCTCTTAAT AAGTTTGAGAAGGAGTCTGGTCATATTCGTAATCAGCGTGGTGATTATTCTCATACTTTTTCTCGTAAGGATCTTCAGGCTGAGCTTATTC TTCTTTTTGAGAAGCAGAAGGAGTTTGGTAATCCTCATGTTTCTGGTGGTCTTAAGGAGGGTATTGAGACTCTTCTTATGACTCAGCGTCC TGCTCTTTCTGGTGATGCTGTTCAGAAGATGCTTGGTCATTGTACTTTTGAGCCTGCTGAGCCTAAGGCTGCTAAGAATACTTATACTGCT GAGCGTTTTATTTGGCTTACTAAGCTTAATAATCTTCGTATTCTTGAGCAGGGTTCTGAGCGTCCTCTTACTGATACTGAGCGTGCTACTC TTATGGATGAGCCTTATCGTAAGTCTAAGCTTACTTATGCTCAGGCTCGTAAGCTTCTTGGTCTTGAGGATACTGCTTTTTTTAAGGGTCT TCGTTATGGTAAGGATAATGCTGAGGCTTCTACTCTTATGGAGATGAAGGCTTATCATGCTATTTCTCGTGCTCTTGAGAAGGAGGGTCTT AAGGATAAGAAGTCTCCTCTTAATCTTTCTTCTGAGCTTCAGGATGAGATTGGTACTGCTTTTTCTCTTTTTAAGACTGATGAGGATATTA CTGGTCGTCTTAAGGATCGTGTTCAGCCTGAGATTCTTGAGGCTCTTCTTAAGCATATTTCTTTTGATAAGTTTGTTCAGATTTCTCTTAA GGCTCTTCGTCGTATTGTTCCTCTTATGGAGCAGGGTAAGCGTTATGATGAGGCTTGTGCTGAGATTTATGGTGATCATTATGGTAAGAAG AATACTGAGGAGAAGATTTATCTTCCTCCTATTCCTGCTGATGAGATTCGTAATCCTGTTGTTCTTCGTGCTCTTTCTCAGGCTCGTAAGG TTATTAATGGTGTTGTTCGTCGTTATGGTTCTCCTGCTCGTATTCATATTGAGACTGCTCGTGAGGTTGGTAAGTCTTTTAAGGATCGTAA GGAGATTGAGAAGCGTCAGGAGGAGAATCGTAAGGATCGTGAGAAGGCTGCTGCTAAGTTTCGTGAGTATTTTCCTAATTTTGTTGGTGAG CCTAAGTCTAAGGATATTCTTAAGCTTCGTCTTTATGAGCAGCAGCATGGTAAGTGTCTTTATTCTGGTAAGGAGATTAATCTTGTTCGTC TTAATGAGAAGGGTTATGTTGAGATTGATCATGCTCTTCCTTTTTCTCGTACTTGGGATGATTCTTTTAATAATAAGGTTCTTGTTCTTGG TTCTGAGAATCAGAATAAGGGTAATCAGACTCCTTATGAGTATTTTAATGGTAAGGATAATTCTCGTGAGTGGCAGGAGTTTAAGGCTCGT GTTGAGACTTCTCGTTTTCCTCGTTCTAAGAAGCAGCGTATTCTTCTTCAGAAGTTTGATGAGGATGGTTTTAAGGAGTGTAATCTTAATG ATACTCGTTATGTTAATCGTTTTCTTTGTCAGTTTGTTGCTGATCATATTCTTCTTACTGGTAAGGGTAAGCGTCGTGTTTTTGCTTCTAA TGGTCAGATTACTAATCTTCTTCGTGGTTTTTGGGGTCTTCGTAAGGTTCGTGCTGAGAATGATCGTCATCATGCTCTTGATGCTGTTGTT GTTGCTTGTTCTACTGTTGCTATGCAGCAGAAGATTACTCGTTTTGTTCGTTATAAGGAGATGAATGCTTTTGATGGTAAGACTATTGATA AGGAGACTGGTAAGGTTCTTCATCAGAAGACTCATTTTCCTCAGCCTTGGGAGTTTTTTGCTCAGGAGGTTATGATTCGTGTTTTTGGTAA GCCTGATGGTAAGCCTGAGTTTGAGGAGGCTGATACTCCTGAGAAGCTTCGTACTCTTCTTGCTGAGAAGCTTTCTTCTCGTCCTGAGGCT GTTCATGAGTATGTTACTCCTCTTTTTGTTTCTCGTGCTCCTAATCGTAAGATGTCTGGTGCTCATAAGGATACTCTTCGTTCTGCTAAGC

		GTTTTGTTAAGCATAATGAGAAGATTTCTGTTAAGCGTGTTTGGCTTACTGAGATTAAGCTTGCTGATCTTGAGAATATGGTTAATTATAA GAATGGTCGTGAGATTGAGCTTTATGAGGCTCTTAAGGCTCGTCTTGAGGCTTATGGTGGTAATGCTAAGCAGGCTTTTGATCCTAAGGAT AATCCTTTTTATAAGAAGGGTGGTCAGCTTGTTAAGGCTGTTCGTGTTGAGAAGACTCAGGAGTCTGGTGTTCTTCTTAATAAGAAGAATG CTTATACTATTGCTGATAATGGTGATATGGTTCGTGTTGATGTTTTTTGTAAGGTTGATAAGAAGGGTAAGAATCAGTATTTTATTGTTCC TATTTATGCTTGGCAGGTTGCTGAGAATATTCTTCCTGATATTGATTGTAAGGGTTATCGTATTGATGATTCTTATACTTTTTGTTTTTCT CTTCATAAGTATGATCTTATTGCTTTTCAGAAGGATGAGAAGTCTAAGGTTGAGTTTGCTTATTATATTAATTGTGATTCTTCTAATGGTC GTTTTTATCTTGCTTGGCATGATAAGGGTTCTAAGGAGCAGCAGTTTCGTATTTCTACTCAGAATCTTGTTCTTATTCAGAAGTATCAGGT TAATGAGCTTGGTAAGGAGATTCGTCCTTGTCGTCTTAAGAAGCGTCCTCCTGTTCGT
389	coding sequence for D16A Nme2Cas9 nickase	GCCGCCTTCAAGCCCAACCCCATCAACTACATCCTGGGCCTGGCCATCGGCATCGCCTCCGTGGGCTGGGCCATGGTGGAGATCGACGAGG AGGAGAACCCCATCCGGCTGATCGACCTGGGCGTGCGGGTGTTCGAGCGGGCCGAGGTGCCCAAGACCGGCGACTCCCTGGCCATGGCCCG GCGGCTGGCCCGGTCCGTGCGGCGGCTGACCCGGCGGCGGGCCCACCGGCTGCTGCGGGCCCGGCGGCTGCTGAAGCGGGAGGGCGTGCTG CAGGCCGCCGACTTCGACGAGAACGGCCTGATCAAGTCCCTGCCCAACACCCCCTGGCAGCTGCGGGCCGCCGCCCTGGACCGGAAGCTGA CCCCCCTGGAGTGGTCCGCCGTGCTGCTGCACCTGATCAAGCACCGGGGCTACCTGTCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAA GGAGCTGGGCGCCCTGCTGAAGGGCGTGGCCAACAACGCCCACGCCCTGCAGACCGGCGACTTCCGGACCCCCGCCGAGCTGGCCCTGAAC AAGTTCGAGAAGGAGTCCGGCCACATCCGGAACCAGCGGGGCGACTACTCCCACACCTTCTCCCGGAAGGACCTGCAGGCCGAGCTGATCC TGCTGTTCGAGAAGCAGAAGGAGTTCGGCAACCCCCACGTGTCCGGCGGCCTGAAGGAGGGCATCGAGACCCTGCTGATGACCCAGCGGCC CGCCCTGTCCGGCGACGCCGTGCAGAAGATGCTGGGCCACTGCACCTTCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCTACACCGCC GAGCGGTTCATCTGGCTGACCAAGCTGAACAACCTGCGGATCCTGGAGCAGGGCTCCGAGCGGCCCCTGACCGACACCGAGCGGGCCACCC TGATGGACGAGCCCTACCGGAAGTCCAAGCTGACCTACGCCCAGGCCCGGAAGCTGCTGGGCCTGGAGGACACCGCCTTCTTCAAGGGCCT GCGGTACGGCAAGGACAACGCCGAGGCCTCCACCCTGATGGAGATGAAGGCCTACCACGCCATCTCCCGGGCCCTGGAGAAGGAGGGCCTG AAGGACAAGAAGTCCCCCCTGAACCTGTCCTCCGAGCTGCAGGACGAGATCGGCACCGCCTTCTCCCTGTTCAAGACCGACGAGGACATCA CCGGCCGGCTGAAGGACCGGGTGCAGCCCGAGATCCTGGAGGCCCTGCTGAAGCACATCTCCTTCGACAAGTTCGTGCAGATCTCCCTGAA GGCCCTGCGGCGGATCGTGCCCCTGATGGAGCAGGGCAAGCGGTACGACGAGGCCTGCGCCGAGATCTACGGCGACCACTACGGCAAGAAG AACACCGAGGAGAAGATCTACCTGCCCCCCATCCCCGCCGACGAGATCCGGAACCCCGTGGTGCTGCGGGCCCTGTCCCAGGCCCGGAAGG TGATCAACGGCGTGGTGCGGCGGTACGGCTCCCCCGCCCGGATCCACATCGAGACCGCCCGGGAGGTGGGCAAGTCCTTCAAGGACCGGAA GGAGATCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGTTCCGGGAGTACTTCCCCAACTTCGTGGGCGAG CCCAAGTCCAAGGACATCCTGAAGCTGCGGCTGTACGAGCAGCAGCACGGCAAGTGCCTGTACTCCGGCAAGGAGATCAACCTGGTGCGGC TGAACGAGAAGGGCTACGTGGAGATCGACCACGCCCTGCCCTTCTCCCGGACCTGGGACGACTCCTTCAACAACAAGGTGCTGGTGCTGGG CTCCGAGAACCAGAACAAGGGCAACCAGACCCCCTACGAGTACTTCAACGGCAAGGACAACTCCCGGGAGTGGCAGGAGTTCAAGGCCCGG GTGGAGACCTCCCGGTTCCCCCGGTCCAAGAAGCAGCGGATCCTGCTGCAGAAGTTCGACGAGGACGGCTTCAAGGAGTGCAACCTGAACG ACACCCGGTACGTGAACCGGTTCCTGTGCCAGTTCGTGGCCGACCACATCCTGCTGACCGGCAAGGGCAAGCGGCGGGTGTTCGCCTCCAA CGGCCAGATCACCAACCTGCTGCGGGGCTTCTGGGGCCTGCGGAAGGTGCGGGCCGAGAACGACCGGCACCACGCCCTGGACGCCGTGGTG GTGGCCTGCTCCACCGTGGCCATGCAGCAGAAGATCACCCGGTTCGTGCGGTACAAGGAGATGAACGCCTTCGACGGCAAGACCATCGACA AGGAGACCGGCAAGGTGCTGCACCAGAAGACCCACTTCCCCCAGCCCTGGGAGTTCTTCGCCCAGGAGGTGATGATCCGGGTGTTCGGCAA GCCCGACGGCAAGCCCGAGTTCGAGGAGGCCGACACCCCCGAGAAGCTGCGGACCCTGCTGGCCGAGAAGCTGTCCTCCCGGCCCGAGGCC GTGCACGAGTACGTGACCCCCCTGTTCGTGTCCCGGGCCCCCAACCGGAAGATGTCCGGCGCCCACAAGGACACCCTGCGGTCCGCCAAGC GGTTCGTGAAGCACAACGAGAAGATCTCCGTGAAGCGGGTGTGGCTGACCGAGATCAAGCTGGCCGACCTGGAGAACATGGTGAACTACAA GAACGGCCGGGAGATCGAGCTGTACGAGGCCCTGAAGGCCCGGCTGGAGGCCTACGGCGGCAACGCCAAGCAGGCCTTCGACCCCAAGGAC AACCCCTTCTACAAGAAGGGCGGCCAGCTGGTGAAGGCCGTGCGGGTGGAGAAGACCCAGGAGTCCGGCGTGCTGCTGAACAAGAAGAACG

		CCTACACCATCGCCGACAACGGCGACATGGTGCGGGTGGACGTGTTCTGCAAGGTGGACAAGAAGGGCAAGAACCAGTACTTCATCGTGCC CATCTACGCCTGGCAGGTGGCCGAGAACATCCTGCCCGACATCGACTGCAAGGGCTACCGGATCGACGACTCCTACACCTTCTGCTTCTCC CTGCACAAGTACGACCTGATCGCCTTCCAGAAGGACGAGAAGTCCAAGGTGGAGTTCGCCTACTACATCAACTGCGACTCCTCCAACGGCC GGTTCTACCTGGCCTGGCACGACAAGGGCTCCAAGGAGCAGCAGTTCCGGATCTCCACCCAGAACCTGGTGCTGATCCAGAAGTACCAGGT GAACGAGCTGGGCAAGGAGATCCGGCCCTGCCGGCTGAAGAAGCGGCCCCCCGTGCGG
390	coding sequence for D16A Nme2Cas9 nickase	GCCGCCTTCAAGCCCAACCCCATCAACTACATCCTGGGCCTGGCCATCGGCATCGCCTCCGTGGGCTGGGCCATGGTGGAGATCGACGAGG AGGAGAACCCCATCCGGCTGATCGACCTGGGCGTGCGGGTGTTCGAGCGGGCCGAGGTGCCCAAGACCGGCGACTCCCTGGCCATGGCCCG GCGGCTGGCCCGGTCCGTGCGGCGGCTGACCCGGCGGCGGGCCCACCGGCTGCTGCGGGCCCGGCGGCTGCTGAAGCGGGAGGGCGTGCTG CAGGCCGCCGACTTCGACGAGAACGGCCTGATCAAGTCCCTGCCCAACACCCCCTGGCAGCTGCGGGCCGCCGCCCTGGACCGGAAGCTGA CCCCCCTGGAGTGGTCCGCCGTGCTGCTGCACCTGATCAAGCACCGGGGCTACCTGTCCCAGCGGAAGAACGAGGGCGAGACCGCCGACAA GGAGCTGGGCGCCCTGCTGAAGGGCGTGGCCAACAACGCCCACGCCCTGCAGACCGGCGACTTCCGGACCCCCGCCGAGCTGGCCCTGAAC AAGTTCGAGAAGGAGTCCGGCCACATCCGGAACCAGCGGGGCGACTACTCCCACACCTTCTCCCGGAAGGACCTGCAGGCCGAGCTGATCC TGCTGTTCGAGAAGCAGAAGGAGTTCGGCAACCCCCACGTGTCCGGCGGCCTGAAGGAGGGCATCGAGACCCTGCTGATGACCCAGCGGCC CGCCCTGTCCGGCGACGCCGTGCAGAAGATGCTGGGCCACTGCACCTTCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCTACACCGCC GAGCGGTTCATCTGGCTGACCAAGCTGAACAACCTGCGGATCCTGGAGCAGGGCTCCGAGCGGCCCCTGACCGACACCGAGCGGGCCACCC TGATGGACGAGCCCTACCGGAAGTCCAAGCTGACCTACGCCCAGGCCCGGAAGCTGCTGGGCCTGGAGGACACCGCCTTCTTCAAGGGCCT GCGGTACGGCAAGGACAACGCCGAGGCCTCCACCCTGATGGAGATGAAGGCCTACCACGCCATCTCCCGGGCCCTGGAGAAGGAGGGCCTG AAGGACAAGAAGTCCCCCCTGAACCTGTCCTCCGAGCTGCAGGACGAGATCGGCACCGCCTTCTCCCTGTTCAAGACCGACGAGGACATCA CCGGCCGGCTGAAGGACCGGGTGCAGCCCGAGATCCTGGAGGCCCTGCTGAAGCACATCTCCTTCGACAAGTTCGTGCAGATCTCCCTGAA GGCCCTGCGGCGGATCGTGCCCCTGATGGAGCAGGGCAAGCGGTACGACGAGGCCTGCGCCGAGATCTACGGCGACCACTACGGCAAGAAG AACACCGAGGAGAAGATCTACCTGCCCCCCATCCCCGCCGACGAGATCCGGAACCCCGTGGTGCTGCGGGCCCTGTCCCAGGCCCGGAAGG TGATCAACGGCGTGGTGCGGCGGTACGGCTCCCCCGCCCGGATCCACATCGAGACCGCCCGGGAGGTGGGCAAGTCCTTCAAGGACCGGAA GGAGATCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGTTCCGGGAGTACTTCCCCAACTTCGTGGGCGAG CCCAAGTCCAAGGACATCCTGAAGCTGCGGCTGTACGAGCAGCAGCACGGCAAGTGCCTGTACTCCGGCAAGGAGATCAACCTGGTGCGGC TGAACGAGAAGGGCTACGTGGAGATCGACCACGCCCTGCCCTTCTCCCGGACCTGGGACGACTCCTTCAACAACAAGGTGCTGGTGCTGGG CTCCGAGAACCAGAACAAGGGCAACCAGACCCCCTACGAGTACTTCAACGGCAAGGACAACTCCCGGGAGTGGCAGGAGTTCAAGGCCCGG GTGGAGACCTCCCGGTTCCCCCGGTCCAAGAAGCAGCGGATCCTGCTGCAGAAGTTCGACGAGGACGGCTTCAAGGAGTGCAACCTGAACG ACACCCGGTACGTGAACCGGTTCCTGTGCCAGTTCGTGGCCGACCACATCCTGCTGACCGGCAAGGGCAAGCGGCGGGTGTTCGCCTCCAA CGGCCAGATCACCAACCTGCTGCGGGGCTTCTGGGGCCTGCGGAAGGTGCGGGCCGAGAACGACCGGCACCACGCCCTGGACGCCGTGGTG GTGGCCTGCTCCACCGTGGCCATGCAGCAGAAGATCACCCGGTTCGTGCGGTACAAGGAGATGAACGCCTTCGACGGCAAGACCATCGACA AGGAGACCGGCAAGGTGCTGCACCAGAAGACCCACTTCCCCCAGCCCTGGGAGTTCTTCGCCCAGGAGGTGATGATCCGGGTGTTCGGCAA GCCCGACGGCAAGCCCGAGTTCGAGGAGGCCGACACCCCCGAGAAGCTGCGGACCCTGCTGGCCGAGAAGCTGTCCTCCCGGCCCGAGGCC GTGCACGAGTACGTGACCCCCCTGTTCGTGTCCCGGGCCCCCAACCGGAAGATGTCCGGCGCCCACAAGGACACCCTGCGGTCCGCCAAGC GGTTCGTGAAGCACAACGAGAAGATCTCCGTGAAGCGGGTGTGGCTGACCGAGATCAAGCTGGCCGACCTGGAGAACATGGTGAACTACAA GAACGGCCGGGAGATCGAGCTGTACGAGGCCCTGAAGGCCCGGCTGGAGGCCTACGGCGGCAACGCCAAGCAGGCCTTCGACCCCAAGGAC AACCCCTTCTACAAGAAGGGCGGCCAGCTGGTGAAGGCCGTGCGGGTGGAGAAGACCCAGGAGTCCGGCGTGCTGCTGAACAAGAAGAACG CCTACACCATCGCCGACAACGGCGACATGGTGCGGGTGGACGTGTTCTGCAAGGTGGACAAGAAGGGCAAGAACCAGTACTTCATCGTGCC CATCTACGCCTGGCAGGTGGCCGAGAACATCCTGCCCGACATCGACTGCAAGGGCTACCGGATCGACGACTCCTACACCTTCTGCTTCTCC CTGCACAAGTACGACCTGATCGCCTTCCAGAAGGACGAGAAGTCCAAGGTGGAGTTCGCCTACTACATCAACTGCGACTCCTCCAACGGCC

		GGTTCTACCTGGCCTGGCACGACAAGGGCTCCAAGGAGCAGCAGTTCCGGATCTCCACCCAGAACCTGGTGCTGATCCAGAAGTACCAGGT GAACGAGCTGGGCAAGGAGATCCGGCCCTGCCGGCTGAAGAAGCGGCCCCCCGTGCGG
391	open reading frame for D16A Nme2Cas9 nickase	ATGGCTGCTTTTAAGCCTAATCCTATTAATTATATTCTTGGTCTTGCTATTGGTATTGCTTCTGTTGGTTGGGCTATGGTTGAGATTGATG AGGAGGAGAATCCTATTCGTCTTATTGATCTTGGTGTTCGTGTTTTTGAGCGTGCTGAGGTTCCTAAGACTGGTGATTCTCTTGCTATGGC TCGTCGTCTTGCTCGTTCTGTTCGTCGTCTTACTCGTCGTCGTGCTCATCGTCTTCTTCGTGCTCGTCGTCTTCTTAAGCGTGAGGGTGTT CTTCAGGCTGCTGATTTTGATGAGAATGGTCTTATTAAGTCTCTTCCTAATACTCCTTGGCAGCTTCGTGCTGCTGCTCTTGATCGTAAGC TTACTCCTCTTGAGTGGTCTGCTGTTCTTCTTCATCTTATTAAGCATCGTGGTTATCTTTCTCAGCGTAAGAATGAGGGTGAGACTGCTGA TAAGGAGCTTGGTGCTCTTCTTAAGGGTGTTGCTAATAATGCTCATGCTCTTCAGACTGGTGATTTTCGTACTCCTGCTGAGCTTGCTCTT AATAAGTTTGAGAAGGAGTCTGGTCATATTCGTAATCAGCGTGGTGATTATTCTCATACTTTTTCTCGTAAGGATCTTCAGGCTGAGCTTA TTCTTCTTTTTGAGAAGCAGAAGGAGTTTGGTAATCCTCATGTTTCTGGTGGTCTTAAGGAGGGTATTGAGACTCTTCTTATGACTCAGCG TCCTGCTCTTTCTGGTGATGCTGTTCAGAAGATGCTTGGTCATTGTACTTTTGAGCCTGCTGAGCCTAAGGCTGCTAAGAATACTTATACT GCTGAGCGTTTTATTTGGCTTACTAAGCTTAATAATCTTCGTATTCTTGAGCAGGGTTCTGAGCGTCCTCTTACTGATACTGAGCGTGCTA CTCTTATGGATGAGCCTTATCGTAAGTCTAAGCTTACTTATGCTCAGGCTCGTAAGCTTCTTGGTCTTGAGGATACTGCTTTTTTTAAGGG TCTTCGTTATGGTAAGGATAATGCTGAGGCTTCTACTCTTATGGAGATGAAGGCTTATCATGCTATTTCTCGTGCTCTTGAGAAGGAGGGT CTTAAGGATAAGAAGTCTCCTCTTAATCTTTCTTCTGAGCTTCAGGATGAGATTGGTACTGCTTTTTCTCTTTTTAAGACTGATGAGGATA TTACTGGTCGTCTTAAGGATCGTGTTCAGCCTGAGATTCTTGAGGCTCTTCTTAAGCATATTTCTTTTGATAAGTTTGTTCAGATTTCTCT TAAGGCTCTTCGTCGTATTGTTCCTCTTATGGAGCAGGGTAAGCGTTATGATGAGGCTTGTGCTGAGATTTATGGTGATCATTATGGTAAG AAGAATACTGAGGAGAAGATTTATCTTCCTCCTATTCCTGCTGATGAGATTCGTAATCCTGTTGTTCTTCGTGCTCTTTCTCAGGCTCGTA AGGTTATTAATGGTGTTGTTCGTCGTTATGGTTCTCCTGCTCGTATTCATATTGAGACTGCTCGTGAGGTTGGTAAGTCTTTTAAGGATCG TAAGGAGATTGAGAAGCGTCAGGAGGAGAATCGTAAGGATCGTGAGAAGGCTGCTGCTAAGTTTCGTGAGTATTTTCCTAATTTTGTTGGT GAGCCTAAGTCTAAGGATATTCTTAAGCTTCGTCTTTATGAGCAGCAGCATGGTAAGTGTCTTTATTCTGGTAAGGAGATTAATCTTGTTC GTCTTAATGAGAAGGGTTATGTTGAGATTGATCATGCTCTTCCTTTTTCTCGTACTTGGGATGATTCTTTTAATAATAAGGTTCTTGTTCT TGGTTCTGAGAATCAGAATAAGGGTAATCAGACTCCTTATGAGTATTTTAATGGTAAGGATAATTCTCGTGAGTGGCAGGAGTTTAAGGCT CGTGTTGAGACTTCTCGTTTTCCTCGTTCTAAGAAGCAGCGTATTCTTCTTCAGAAGTTTGATGAGGATGGTTTTAAGGAGTGTAATCTTA ATGATACTCGTTATGTTAATCGTTTTCTTTGTCAGTTTGTTGCTGATCATATTCTTCTTACTGGTAAGGGTAAGCGTCGTGTTTTTGCTTC TAATGGTCAGATTACTAATCTTCTTCGTGGTTTTTGGGGTCTTCGTAAGGTTCGTGCTGAGAATGATCGTCATCATGCTCTTGATGCTGTT GTTGTTGCTTGTTCTACTGTTGCTATGCAGCAGAAGATTACTCGTTTTGTTCGTTATAAGGAGATGAATGCTTTTGATGGTAAGACTATTG ATAAGGAGACTGGTAAGGTTCTTCATCAGAAGACTCATTTTCCTCAGCCTTGGGAGTTTTTTGCTCAGGAGGTTATGATTCGTGTTTTTGG TAAGCCTGATGGTAAGCCTGAGTTTGAGGAGGCTGATACTCCTGAGAAGCTTCGTACTCTTCTTGCTGAGAAGCTTTCTTCTCGTCCTGAG GCTGTTCATGAGTATGTTACTCCTCTTTTTGTTTCTCGTGCTCCTAATCGTAAGATGTCTGGTGCTCATAAGGATACTCTTCGTTCTGCTA AGCGTTTTGTTAAGCATAATGAGAAGATTTCTGTTAAGCGTGTTTGGCTTACTGAGATTAAGCTTGCTGATCTTGAGAATATGGTTAATTA TAAGAATGGTCGTGAGATTGAGCTTTATGAGGCTCTTAAGGCTCGTCTTGAGGCTTATGGTGGTAATGCTAAGCAGGCTTTTGATCCTAAG GATAATCCTTTTTATAAGAAGGGTGGTCAGCTTGTTAAGGCTGTTCGTGTTGAGAAGACTCAGGAGTCTGGTGTTCTTCTTAATAAGAAGA ATGCTTATACTATTGCTGATAATGGTGATATGGTTCGTGTTGATGTTTTTTGTAAGGTTGATAAGAAGGGTAAGAATCAGTATTTTATTGT TCCTATTTATGCTTGGCAGGTTGCTGAGAATATTCTTCCTGATATTGATTGTAAGGGTTATCGTATTGATGATTCTTATACTTTTTGTTTT TCTCTTCATAAGTATGATCTTATTGCTTTTCAGAAGGATGAGAAGTCTAAGGTTGAGTTTGCTTATTATATTAATTGTGATTCTTCTAATG GTCGTTTTTATCTTGCTTGGCATGATAAGGGTTCTAAGGAGCAGCAGTTTCGTATTTCTACTCAGAATCTTGTTCTTATTCAGAAGTATCA GGTTAATGAGCTTGGTAAGGAGATTCGTCCTTGTCGTCTTAAGAAGCGTCCTCCTGTTCGTUGA
392	open reading	ATGGCCGCCTTCAAGCCCAACCCCATCAACTACATCCTGGGCCTGGCCATCGGCATCGCCTCCGTGGGCTGGGCCATGGTGGAGATCGACG

	frame for D16A Nme2Cas9 nickase	AGGAGGAGAACCCCATCCGGCTGATCGACCTGGGCGTGCGGGTGTTCGAGCGGGCCGAGGTGCCCAAGACCGGCGACTCCCTGGCCATGGC CCGGCGGCTGGCCCGGTCCGTGCGGCGGCTGACCCGGCGGCGGGCCCACCGGCTGCTGCGGGCCCGGCGGCTGCTGAAGCGGGAGGGCGTG CTGCAGGCCGCCGACTTCGACGAGAACGGCCTGATCAAGTCCCTGCCCAACACCCCCTGGCAGCTGCGGGCCGCCGCCCTGGACCGGAAGC TGACCCCCCTGGAGTGGTCCGCCGTGCTGCTGCACCTGATCAAGCACCGGGGCTACCTGTCCCAGCGGAAGAACGAGGGCGAGACCGCCGA CAAGGAGCTGGGCGCCCTGCTGAAGGGCGTGGCCAACAACGCCCACGCCCTGCAGACCGGCGACTTCCGGACCCCCGCCGAGCTGGCCCTG AACAAGTTCGAGAAGGAGTCCGGCCACATCCGGAACCAGCGGGGCGACTACTCCCACACCTTCTCCCGGAAGGACCTGCAGGCCGAGCTGA TCCTGCTGTTCGAGAAGCAGAAGGAGTTCGGCAACCCCCACGTGTCCGGCGGCCTGAAGGAGGGCATCGAGACCCTGCTGATGACCCAGCG GCCCGCCCTGTCCGGCGACGCCGTGCAGAAGATGCTGGGCCACTGCACCTTCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCTACACC GCCGAGCGGTTCATCTGGCTGACCAAGCTGAACAACCTGCGGATCCTGGAGCAGGGCTCCGAGCGGCCCCTGACCGACACCGAGCGGGCCA CCCTGATGGACGAGCCCTACCGGAAGTCCAAGCTGACCTACGCCCAGGCCCGGAAGCTGCTGGGCCTGGAGGACACCGCCTTCTTCAAGGG CCTGCGGTACGGCAAGGACAACGCCGAGGCCTCCACCCTGATGGAGATGAAGGCCTACCACGCCATCTCCCGGGCCCTGGAGAAGGAGGGC CTGAAGGACAAGAAGTCCCCCCTGAACCTGTCCTCCGAGCTGCAGGACGAGATCGGCACCGCCTTCTCCCTGTTCAAGACCGACGAGGACA TCACCGGCCGGCTGAAGGACCGGGTGCAGCCCGAGATCCTGGAGGCCCTGCTGAAGCACATCTCCTTCGACAAGTTCGTGCAGATCTCCCT GAAGGCCCTGCGGCGGATCGTGCCCCTGATGGAGCAGGGCAAGCGGTACGACGAGGCCTGCGCCGAGATCTACGGCGACCACTACGGCAAG AAGAACACCGAGGAGAAGATCTACCTGCCCCCCATCCCCGCCGACGAGATCCGGAACCCCGTGGTGCTGCGGGCCCTGTCCCAGGCCCGGA AGGTGATCAACGGCGTGGTGCGGCGGTACGGCTCCCCCGCCCGGATCCACATCGAGACCGCCCGGGAGGTGGGCAAGTCCTTCAAGGACCG GAAGGAGATCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGTTCCGGGAGTACTTCCCCAACTTCGTGGGC GAGCCCAAGTCCAAGGACATCCTGAAGCTGCGGCTGTACGAGCAGCAGCACGGCAAGTGCCTGTACTCCGGCAAGGAGATCAACCTGGTGC GGCTGAACGAGAAGGGCTACGTGGAGATCGACCACGCCCTGCCCTTCTCCCGGACCTGGGACGACTCCTTCAACAACAAGGTGCTGGTGCT GGGCTCCGAGAACCAGAACAAGGGCAACCAGACCCCCTACGAGTACTTCAACGGCAAGGACAACTCCCGGGAGTGGCAGGAGTTCAAGGCC CGGGTGGAGACCTCCCGGTTCCCCCGGTCCAAGAAGCAGCGGATCCTGCTGCAGAAGTTCGACGAGGACGGCTTCAAGGAGTGCAACCTGA ACGACACCCGGTACGTGAACCGGTTCCTGTGCCAGTTCGTGGCCGACCACATCCTGCTGACCGGCAAGGGCAAGCGGCGGGTGTTCGCCTC CAACGGCCAGATCACCAACCTGCTGCGGGGCTTCTGGGGCCTGCGGAAGGTGCGGGCCGAGAACGACCGGCACCACGCCCTGGACGCCGTG GTGGTGGCCTGCTCCACCGTGGCCATGCAGCAGAAGATCACCCGGTTCGTGCGGTACAAGGAGATGAACGCCTTCGACGGCAAGACCATCG ACAAGGAGACCGGCAAGGTGCTGCACCAGAAGACCCACTTCCCCCAGCCCTGGGAGTTCTTCGCCCAGGAGGTGATGATCCGGGTGTTCGG CAAGCCCGACGGCAAGCCCGAGTTCGAGGAGGCCGACACCCCCGAGAAGCTGCGGACCCTGCTGGCCGAGAAGCTGTCCTCCCGGCCCGAG GCCGTGCACGAGTACGTGACCCCCCTGTTCGTGTCCCGGGCCCCCAACCGGAAGATGTCCGGCGCCCACAAGGACACCCTGCGGTCCGCCA AGCGGTTCGTGAAGCACAACGAGAAGATCTCCGTGAAGCGGGTGTGGCTGACCGAGATCAAGCTGGCCGACCTGGAGAACATGGTGAACTA CAAGAACGGCCGGGAGATCGAGCTGTACGAGGCCCTGAAGGCCCGGCTGGAGGCCTACGGCGGCAACGCCAAGCAGGCCTTCGACCCCAAG GACAACCCCTTCTACAAGAAGGGCGGCCAGCTGGTGAAGGCCGTGCGGGTGGAGAAGACCCAGGAGTCCGGCGTGCTGCTGAACAAGAAGA ACGCCTACACCATCGCCGACAACGGCGACATGGTGCGGGTGGACGTGTTCTGCAAGGTGGACAAGAAGGGCAAGAACCAGTACTTCATCGT GCCCATCTACGCCTGGCAGGTGGCCGAGAACATCCTGCCCGACATCGACTGCAAGGGCTACCGGATCGACGACTCCTACACCTTCTGCTTC TCCCTGCACAAGTACGACCTGATCGCCTTCCAGAAGGACGAGAAGTCCAAGGTGGAGTTCGCCTACTACATCAACTGCGACTCCTCCAACG GCCGGTTCTACCTGGCCTGGCACGACAAGGGCTCCAAGGAGCAGCAGTTCCGGATCTCCACCCAGAACCTGGTGCTGATCCAGAAGTACCA GGTGAACGAGCTGGGCAAGGAGATCCGGCCCTGCCGGCTGAAGAAGCGGCCCCCCGTGCGGUGA
393	open reading frame for D16A Nme2Cas9 nickase	ATGGCCGCCTTCAAGCCCAACCCCATCAACTACATCCTGGGCCTGGCCATCGGCATCGCCTCCGTGGGCTGGGCCATGGTGGAGATCGACG AGGAGGAGAACCCCATCCGGCTGATCGACCTGGGCGTGCGGGTGTTCGAGCGGGCCGAGGTGCCCAAGACCGGCGACTCCCTGGCCATGGC CCGGCGGCTGGCCCGGTCCGTGCGGCGGCTGACCCGGCGGCGGGCCCACCGGCTGCTGCGGGCCCGGCGGCTGCTGAAGCGGGAGGGCGTG CTGCAGGCCGCCGACTTCGACGAGAACGGCCTGATCAAGTCCCTGCCCAACACCCCCTGGCAGCTGCGGGCCGCCGCCCTGGACCGGAAGC

		TGACCCCCCTGGAGTGGTCCGCCGTGCTGCTGCACCTGATCAAGCACCGGGGCTACCTGTCCCAGCGGAAGAACGAGGGCGAGACCGCCGA CAAGGAGCTGGGCGCCCTGCTGAAGGGCGTGGCCAACAACGCCCACGCCCTGCAGACCGGCGACTTCCGGACCCCCGCCGAGCTGGCCCTG AACAAGTTCGAGAAGGAGTCCGGCCACATCCGGAACCAGCGGGGCGACTACTCCCACACCTTCTCCCGGAAGGACCTGCAGGCCGAGCTGA TCCTGCTGTTCGAGAAGCAGAAGGAGTTCGGCAACCCCCACGTGTCCGGCGGCCTGAAGGAGGGCATCGAGACCCTGCTGATGACCCAGCG GCCCGCCCTGTCCGGCGACGCCGTGCAGAAGATGCTGGGCCACTGCACCTTCGAGCCCGCCGAGCCCAAGGCCGCCAAGAACACCTACACC GCCGAGCGGTTCATCTGGCTGACCAAGCTGAACAACCTGCGGATCCTGGAGCAGGGCTCCGAGCGGCCCCTGACCGACACCGAGCGGGCCA CCCTGATGGACGAGCCCTACCGGAAGTCCAAGCTGACCTACGCCCAGGCCCGGAAGCTGCTGGGCCTGGAGGACACCGCCTTCTTCAAGGG CCTGCGGTACGGCAAGGACAACGCCGAGGCCTCCACCCTGATGGAGATGAAGGCCTACCACGCCATCTCCCGGGCCCTGGAGAAGGAGGGC CTGAAGGACAAGAAGTCCCCCCTGAACCTGTCCTCCGAGCTGCAGGACGAGATCGGCACCGCCTTCTCCCTGTTCAAGACCGACGAGGACA TCACCGGCCGGCTGAAGGACCGGGTGCAGCCCGAGATCCTGGAGGCCCTGCTGAAGCACATCTCCTTCGACAAGTTCGTGCAGATCTCCCT GAAGGCCCTGCGGCGGATCGTGCCCCTGATGGAGCAGGGCAAGCGGTACGACGAGGCCTGCGCCGAGATCTACGGCGACCACTACGGCAAG AAGAACACCGAGGAGAAGATCTACCTGCCCCCCATCCCCGCCGACGAGATCCGGAACCCCGTGGTGCTGCGGGCCCTGTCCCAGGCCCGGA AGGTGATCAACGGCGTGGTGCGGCGGTACGGCTCCCCCGCCCGGATCCACATCGAGACCGCCCGGGAGGTGGGCAAGTCCTTCAAGGACCG GAAGGAGATCGAGAAGCGGCAGGAGGAGAACCGGAAGGACCGGGAGAAGGCCGCCGCCAAGTTCCGGGAGTACTTCCCCAACTTCGTGGGC GAGCCCAAGTCCAAGGACATCCTGAAGCTGCGGCTGTACGAGCAGCAGCACGGCAAGTGCCTGTACTCCGGCAAGGAGATCAACCTGGTGC GGCTGAACGAGAAGGGCTACGTGGAGATCGACCACGCCCTGCCCTTCTCCCGGACCTGGGACGACTCCTTCAACAACAAGGTGCTGGTGCT GGGCTCCGAGAACCAGAACAAGGGCAACCAGACCCCCTACGAGTACTTCAACGGCAAGGACAACTCCCGGGAGTGGCAGGAGTTCAAGGCC CGGGTGGAGACCTCCCGGTTCCCCCGGTCCAAGAAGCAGCGGATCCTGCTGCAGAAGTTCGACGAGGACGGCTTCAAGGAGTGCAACCTGA ACGACACCCGGTACGTGAACCGGTTCCTGTGCCAGTTCGTGGCCGACCACATCCTGCTGACCGGCAAGGGCAAGCGGCGGGTGTTCGCCTC CAACGGCCAGATCACCAACCTGCTGCGGGGCTTCTGGGGCCTGCGGAAGGTGCGGGCCGAGAACGACCGGCACCACGCCCTGGACGCCGTG GTGGTGGCCTGCTCCACCGTGGCCATGCAGCAGAAGATCACCCGGTTCGTGCGGTACAAGGAGATGAACGCCTTCGACGGCAAGACCATCG ACAAGGAGACCGGCAAGGTGCTGCACCAGAAGACCCACTTCCCCCAGCCCTGGGAGTTCTTCGCCCAGGAGGTGATGATCCGGGTGTTCGG CAAGCCCGACGGCAAGCCCGAGTTCGAGGAGGCCGACACCCCCGAGAAGCTGCGGACCCTGCTGGCCGAGAAGCTGTCCTCCCGGCCCGAG GCCGTGCACGAGTACGTGACCCCCCTGTTCGTGTCCCGGGCCCCCAACCGGAAGATGTCCGGCGCCCACAAGGACACCCTGCGGTCCGCCA AGCGGTTCGTGAAGCACAACGAGAAGATCTCCGTGAAGCGGGTGTGGCTGACCGAGATCAAGCTGGCCGACCTGGAGAACATGGTGAACTA CAAGAACGGCCGGGAGATCGAGCTGTACGAGGCCCTGAAGGCCCGGCTGGAGGCCTACGGCGGCAACGCCAAGCAGGCCTTCGACCCCAAG GACAACCCCTTCTACAAGAAGGGCGGCCAGCTGGTGAAGGCCGTGCGGGTGGAGAAGACCCAGGAGTCCGGCGTGCTGCTGAACAAGAAGA ACGCCTACACCATCGCCGACAACGGCGACATGGTGCGGGTGGACGTGTTCTGCAAGGTGGACAAGAAGGGCAAGAACCAGTACTTCATCGT GCCCATCTACGCCTGGCAGGTGGCCGAGAACATCCTGCCCGACATCGACTGCAAGGGCTACCGGATCGACGACTCCTACACCTTCTGCTTC TCCCTGCACAAGTACGACCTGATCGCCTTCCAGAAGGACGAGAAGTCCAAGGTGGAGTTCGCCTACTACATCAACTGCGACTCCTCCAACG GCCGGTTCTACCTGGCCTGGCACGACAAGGGCTCCAAGGAGCAGCAGTTCCGGATCTCCACCCAGAACCTGGTGCTGATCCAGAAGTACCA GGTGAACGAGCTGGGCAAGGAGATCCGGCCCTGCCGGCTGAAGAAGCGGCCCCCCGTGCGGUAA
394-400		Not used
401	G018932	mC*mG*mC*ACGGGUACCAGGGGCCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
402	G018933	mG*mC*mA*CGGGUACCAGGGGCCACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
403	G018934	mC*mA*mC*GGGUACCAGGGGCCACGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
404	G018935	mG*mG*mG*AGGCGCCCCGUGGCCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
405	G018936	mG*mC*mG*AUCAGGGAGGCGCCCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
406	G018937	mC*mG*mG*GCUGGCCUCCCACAAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
407	G018938	mC*mU*mC*CUUGUGGGAGGCCAGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
408	G018939	mG*mG*mC*UGGCCUCCCACAAGGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
409	G018940	mU*mU*mG*UCUCCCCUCCUUGUGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
410	G018941	mC*mC*mA*CAAGGAGGGGAGACAAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
411	G018942	mC*mA*mC*AAGGAGGGGAGACAAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
412	G018943	mC*mA*mA*UUGUCUCCCCUCCUUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
413	G018944	mC*mC*mA*AUUGUCUCCCCUCCUUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
414	G018945	mA*mU*mC*CCUCGAAUACUGAUGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
415	G018946	mA*mA*mC*CACUCAUCAGUAUUCGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
416	G018947	mG*mA*mA*CCACUCAUCAGUAUUCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
417		Not Used
418	G018949	mG*mG*mC*CCGUGACUUUUCCUCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
419	G018950	mU*mG*mC*UUCACACUCAAUGUGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
420	G018951	mG*mC*mU*UCACACUCAAUGUGUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
421	G018952	mC*mU*mU*CACACUCAAUGUGUGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
422	G018953	mU*mU*mC*ACACUCAAUGUGUGUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
423		Not Used
424	G018955	mA*mG*mG*GGCCCUGACCCUGCUAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
425		Not Used
426		Not Used
427	G018958	mU*mG*mA*GAUUUCUUGUCUCACUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
428	G018959	mG*mA*mG*AUUUCUUGUCUCACUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
429	G018960	mU*mA*mA*AGCACCUGUUAAAAUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
430	G018961	mA*mA*mU*CUGUCCUUCAUUUUAACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
431	G018962	mG*mU*mC*ACAGGGGAAGGUCCCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
432	G018963	mA*mA*mC*AUGAAGAAAGCAGGUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
433	G018964	mA*mA*mA*CAUGAAGAAAGCAGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
434	G018965	mU*mC*mA*GGAAACAUGAAGAAAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
435	G018966	mU*mG*mU*CCUGUGAGAUACCAGAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
436		Not Used
437	G018968	mA*mG*mG*CUGAUGCCUGAGGUCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
438	G018969	mG*mG*mC*UGAUGCCUGAGGUCCUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
439	G018970	mC*mA*mC*AAUAUCCCAAGGACCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
440	G018971	mG*mG*mU*CCUUGGGAUAUUGUGUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
441	G018972	mG*mU*mC*CUUGGGAUAUUGUGUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
442	G018973	mC*mU*mC*CCAAACACAAUAUCCCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
443	G018974	mU*mC*mC*UCUAGCCACAUCUUCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
444		Not Used
445	G018976	mC*mC*mU*CUAGCCACAUCUUCUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
446	G018977	mC*mC*mC*ACAGAAGAUGUGGCUAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
447	G018978	mG*mU*mC*AGAUCCCACAGAAGAUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
448	G018979	mA*mU*mC*UUCUGUGGGAUCUGACCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
449	G018980	mC*mC*mC*AGGCAGUGACAGUGCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
450	G018981	mC*mU*mG*GGCACUGUCACUGCCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
451	G018982	mC*mC*mU*GGGCACUGUCACUGCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
452	G018983	mU*mG*mG*AGGGCCUGAUGUGUGUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
453	G018984	mG*mC*mC*UGAUGUGUGUUGGGUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
454		Not Used
455	G018986	mC*mC*mC*AACACCCAACACACAUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
456	G018987	mU*mU*mG*GGUGUUGGGCGGAACAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
457	G018988	mU*mG*mG*AUGUAUUGAGCAUGCGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
458	G018989	mG*mG*mA*UGUAUUGAGCAUGCGAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
459	G018990	mA*mC*mG*GCCAUCCUCGGCGUCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
460	G018991	mG*mA*mC*GGCCAUCCUCGGCGUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
461	G018992	mG*mA*mC*GCCGAGGAUGGCCGUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
462	G018993	mU*mG*mA*CGGCCAUCCUCGGCGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
463	G018994	mG*mG*mC*GCCAUGACGGCCAUCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
464	G018995	mA*mC*mA*GCGACGCCGCGAGCCAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
465	G018996	mC*mG*mA*CGCCGCGAGCCAGAGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
466	G018997	mC*mG*mA*GCCAGAGGAUGGAGCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
467	G018998	mC*mG*mG*CUCCAUCCUCUGGCUCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
468	G018999	mG*mA*mG*CCAGAGGAUGGAGCCGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
469	G019000	mG*mC*mG*CCCGCGGCUCCAUCCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
470	G019001	mG*mC*mC*CGUCCGUGGGGGAUGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
471	G019002	mU*mC*mA*UCCCCCACGGACGGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
472	G019003	mA*mU*mC*UCGGACCCGGAGACUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
473	G019004	mG*mG*mG*GUCCCGCGGCUUCGGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
474	G019005	mC*mU*mC*GGGGUCCCGCGGCUUCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
475	G019006	mU*mC*mU*CGGGGUCCCGCGGCUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
476	G019007	mG*mU*mC*UCGGGGUCCCGCGGCUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
477	G019008	mG*mC*mA*AGGGUCUCGGGGUCCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
478	G019009	mG*mG*mA*CCCCGAGACCCUUGCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
479	G019010	mG*mA*mC*CCCGAGACCCUUGCCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
480	G019011	mC*mG*mA*GACCCUUGCCCCGGGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
481	G019012	mC*mU*mC*CCGGGGCAAGGGUCUCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
482	G019013	mU*mC*mU*CCCGGGGCAAGGGUCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
483	G019014	mC*mU*mC*UCCCGGGGCAAGGGUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
484	G019015	mC*mC*mU*UGCCCCGGGAGAGGCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
485	G019016	mC*mU*mG*GGCCUCUCCCGGGGCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
486	G019017	mC*mC*mU*GGGCCUCUCCCGGGGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
487	G019018	mA*mG*mG*CGCCUGGGCCUCUCCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
488	G019019	mU*mU*mU*AGGCCAAAAAUCCCCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
489	G021205	mA*mG*mG*CAUUUUGCAUCUGUCAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
490	G021206	mC*mA*mG*GCAUUUUGCAUCUGUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
491	G021207	mC*mC*mC*UGGGCACUGUCACUGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
492	G021208	mC*mG*mC*UGCAGCGCACGGGUACCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
493	G021209	mG*mC*mU*GCAGCGCACGGGUACCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
494	G021210	mC*mU*mG*CAGCGCACGGGUACCAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
495	G021211	mU*mC*mC*UUGUGGGAGGCCAGCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
496	G023522 Guide RNA Targeting CD38	mC*mU*mU*GACGCAUCGCGCCAGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUA UCACGAAAGGGCACCGAGUCGGmUmGmC*mU
497	G019427 Guide RNA	mA*mA*mU*CCUAAUUGCCUCCUGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	Targeting ANAPC5	
498	G023519 Guide RNA Targeting B2M	mA*mC*mU*CACGCUGGAUAGCCUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGmUmGmC*mU
499	G023520 Guide RNA Targeting TRAC	mU*mU*mC*AAAACCUGUCAGUGAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGmUmGmC*mU
500	G023521 Guide RNA Targeting CIITA	mC*mG*mC*CCAGGUCCUCACGUCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGmUmGmC*mU
501	G023523 Guide RNA Targeting HLA-A	mG*mC*mU*GCAGCGCACGGGUACCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGmUmGmC*mU
502	G023524 Guide RNA Targeting TRBC1/2	mC*mC*mA*CACCCAAAAGGCCACACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGmUmGmC*mU
503	G020073	mG*mUGUGUCCCmUCUmCCCCACCCGUCCmGUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmGmAmAmAmGmAmUGUGCmCGCAACGCUCUmGmCCmUmUmCmUGGCAUCGmU*mU
504	G020927	mA*mU*mC*UGCCUUUAUAGGGCACCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
505	G020928	mA*mG*mU*CACGAUGCCUUUAUAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
506	G020929	mG*mG*mU*UGCAAGGAAUGAGAACCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
507	G021237	mG*mGGGCCAGCmUUCmAGACACAAAUACmGUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmGmAmAmAmGmAmUGUGCmCGCmAmAmCmGCUCUmGmCCmUmUmCmUGmGCmAmUC*mG*mU*mU
508	G021249	mC*mAGAAUUGAmGAGmACUCAGCCCAGGmGUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmGmAmAmAmGmAmUGUGCmCGCmAmAmCmGCUCUmGmCCmUmUmCmUGmGCmAmUC*mG*mU*mU
509	G021321	mG*mAGAAGCCGmUCAmCACAGAUCCACAmGUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmGmAmAmAmGmAmUGUGCmCGCmAmAmCmGCUCUmGmCCmUmUmCmUGmGCmAmUC*mG*mU*mU
510	G021844	mC*mC*mA*mAmGUGmUmCUmUCmCAGUAmCGAUmUUGmGUUGmUmAmGmCUCCCmUmUmC(L1)mGmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmC(L1)mGmAmUGUGCmCGmCAAmCGCUCUmGmCC(L1)GGCAUCG*mU*mU
511	G000502	mA*mC*mA*CAAAUACCAGUCCAGCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
512	Wild-type	NNNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCCCUUUCUCAUUUCGGAAACGAAAUGAGAACCGUUGCUACAAUAAGGCCGUCUGAAAA

	NmeCas9 guide RNA	GAUGUGCCGCAACGCUCUGCCCCUUAAAGCUUCUGCUUUAAGGGGCAUCGUUUA
513	Shortened/unmodified NmeCas9 guide RNA	$(N)_{20-25}$ GUUGUAGCUCCCUGAAACCGUUGCUACAAUAAGGCCGUCGAAAGAUGUGCCGCAACGCUCUGCCUUCUGGCAUCGUU
514	Shortened/unmodified NmeCas9 guide RNA	$(N)_{20-25}$ GUUGUAGCUCCCUGAAACCGUUGCUACAAUAAGGCCGUCGAAAGAUGU GCCGCAACGCUCUGCCUUCUGGCAUCGUUUAUU
515	Shortened/unmodified NmeCas9 guide RNA	$(N)_{20-25}$ GUUGUAGCUCCCUGGAAACCCGUUGCUACAAUAAGGCCGUCGAAAGA UGUGCCGCAACGCUCUGCCUUCUGGCAUCGUUUAUU
516	Mod-N77 conserved portion only	GUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmGmAmAmAmGmAmUGUGCmCGCmAmAmCmGCU CUmGmCCmUmUmCmUGmGCmAmUC*mG*mU*mU
517	Mod-N78 conserved portion only	GUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmGmAmAmAmGmAmUGUGCmCGmCAAmCGCUCU mGmCCmUmUmCmUGGCAUCG*mU*mU
518	Shortened/unmodified NmeCas9 guide RNA comprising linkers	$(N)_{20-25}$ GUUGUAGCUCCCUUC(L1)GACCGUUGCUACAAUAAGGCCGUC(L1)GAUGUGCCGCAACGCUCUGCC(L1)GGCAUCGUU
519	Shortened/modified NmeCas9 guide RNA comprising linkers	mN*mN*mN*mNmNNNmNmNNmNNmNNNNNmNNNNmNNNmGUUGmUmAmGmCUCCCmUmUmC(L1)mGmAmCmCGUUmGmCUAmCAAU*AAG mGmCCmGmUmC(L1)mGmAmUGUGCmCGmCAAmCGCUCUmGmCC(L1)GGCAUCG*mU*mU
520	Shortened SpyCas9 guide RNA	mN*mN*mN*NNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAA GGGCACCGAGUCGGmUmGmC*mU
521	Shortened SpyCas9 guide RNA (UNMODIFIED)	NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCGGUGCU
522	Shortened NmeCas9 guide RNA	mN*mNNNNNNNNmNNNmNNNNNNNNNNNNmGUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmG mAmAmAmGmAmUGUGCmCGCmAmAmCmGCUCUmGmCCmUmUmCmUGmGCmAmUC*mG*mU*mU
523	Shortened/modified spyCas9 guide RNA comprising	mA*mC*mG*CAAAUAUCAGUCCAGCGGUUUUAGAmGmCmUmA(L1)mUmAmGmCAAGUUAAAAUAAGGC(L2)GUCCGUUAUCAC(L1)GG GCACCGAGUCGG*mU*mG*mC

	internal linkers	
524	Shortened spyCas9 guide RNA comprising internal linkers	NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUA(L1)UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU(L1)AAGUGGCACCGAGUCGGUGCUUUUU
525-528		Not used
529	G016788	mU*mU*mC*CCGGCCUUUUUACCUUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
530	G016030	mU*mC*mA*ACUGCGACCAGUUCAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
531	G016031	mA*mG*mC*GCAGGCAGUGGCAGGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
532	G016032	mA*mC*mC*UGCAACAACAGGAUUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
533	G016033	mC*mC*mA*GGAACACCUGCAACAACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
534	G016034	mC*mA*mG*CAGCAAGAGCCUGGAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
535	G016035	mG*mC*mA*GCAGCAAGAGCCUGGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
536	G016036	mC*mA*mA*UCUCUUCUUCUCCAGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
537	G016037	mA*mG*mC*AGCUCGCUGCCAGCCUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
538	G016038	mA*mG*mC*UCGCUGCCAGCCUUCGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
539	G016039	mC*mU*mC*UGGACCAGGCGGCCCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
540	G016040	mG*mC*mA*GCCCUCGACAGCCCCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
541	G016041	mC*mA*mG*CCCUCGACAGCCCCCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
542	G016042	mA*mG*mC*CAAGUACCCCCUCCCAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
543	G016043	mC*mA*mG*CCAACAGCACCUCAGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
544	G016044	mU*mG*mC*CGCGCCCGCAGUGUCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
545	G016045	mC*mG*mA*CAGCCCCCCCGGGGCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
546	G016046	mG*mG*mC*AGCGAGCUGCUGGGCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
547	G016047	mG*mC*mC*AGCUCUGCCAGGGCCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
548	G016048	mA*mG*mC*GAGCGAAGGCAGGGCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
549	G016049	mA*mA*mG*GCUGGCAGCGAGCUGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
550	G016050	mA*mG*mC*CCUCGACAGCCCCCCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
551	G016051	mG*mG*mA*UGCAGCGAGCGAAGGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
552	G016052	mU*mC*mC*ACCGAGGCAGCCGCCGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
553	G016053	mU*mC*mU*CCAACAAGCUUCCAAAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
554	G016054	mA*mG*mC*AGCCCCCGGAGGGAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
555	G016055	mG*mC*mC*ACAGCCCUACUUUGUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
556	G016056	mC*mU*mG*CGCCCACGAGGCCGAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
557	G016057	mC*mA*mG*CCGCCGAUGGCCCGAGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
558	G016058	mC*mG*mA*GGCUCCCCAAUCCAGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
559	G016059	mC*mU*mC*AACGAGGAACUGGAGAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
560	G016060	mC*mC*mA*CAGCCCUACUUUGUGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
561	G016061	mC*mA*mA*GAGCCUGGAGCGGGAACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
562	G016062	mG*mA*mC*UGCCAGUCACCACAGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
563	G016063	mG*mC*mC*CACGAGGCCGAGGAGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
564	G016064	mA*mU*mC*AGCCCAGCCAGAAAGCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
565	G016065	mC*mA*mG*AGAAGACAAAGUCGUACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
566	G016066	mU*mG*mG*GAGUCCCUGCAGCAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
567	G016067	mC*mA*mG*CAGGCUGUUGUGUGACAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
568	G016068	mC*mU*mG*GUCAGGGCAAGAGCUAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
569	G016069	mC*mA*mG*GUUCAGGCAUGCUGGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
570	G016070	mU*mU*mU*CCAAGGACUUCAGCUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
571	G016071	mA*mG*mG*CCGAGGAGGCUGGAAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
572	G016072	mC*mA*mG*UGGCUGAUGGAGCGAAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
573	G016073	mG*mG*mG*AUGCAGCGAGCGAAGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
574	G016074	mA*mA*mG*CUGCCCUCCACGCUCACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
575	G016075	mU*mU*mC*CUCCUGCAAUGCUUCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
576	G016076	mU*mC*mC*UCCUGCAAUGCUUCCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
577	G016077	mC*mA*mA*GCUGCCCUCCACGCUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
578	G016078	mU*mC*mC*CUGGACCUCCGCAGCACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
579	G016079	mU*mC*mC*AGCCUCCUCGGCCUCGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
580	G016080	mC*mU*mU*CUCCAGCCAGGUCCAUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
581	G016081	mC*mU*mU*CCUCCUGCAAUGCUUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
582	G016082	mC*mG*mU*CCUCCCCAAGCUCCAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
583	G016083	mC*mC*mA*GCUCCUCGAAGCCGUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
584	G016084	mC*mU*mU*UUCCUCCCUGCAGCAUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
585	G016085	mC*mG*mG*CCGCCACGAGUGGCUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
586	G016086	mC*mG*mC*CCAGGUCCUCACGUCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
587	G016087	mC*mA*mC*GUGCGGACCGGCACCGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
588	G016088	mC*mA*mG*CUUGGCCAGCUCUGCCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
589	G016089	mU*mC*mG*GACUCUGCGGCCCGCGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
590	G016090	mC*mU*mU*CCCCCAGCUGAAGUCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
591	G016091	mG*mC*mC*CAGCUCCCAGGCCAGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
592	G016092	mC*mC*mC*UCCAGCCAGUUGUCAUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
593	G016093	mG*mC*mC*CUCCAGCCAGUUGUCAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
594	G016094	mC*mC*mC*UGACGCUCCUCCGGGACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
595	G016095	mC*mA*mC*ACUGCCCGGCACAAAGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
596	G016096	mC*mA*mG*CUCACAGUGUGCCACCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
597	G016097	mA*mG*mC*UCGGACUCUGCGGCCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
598	G016098	mG*mA*mC*GCCCUAUUUGAGCUGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
599	G016099	mG*mC*mC*AGUGCUGCGGAGGUCCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
600	G016100	mA*mG*mG*GCUCCCAGGCAGCGGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
601	G016101	mG*mC*mC*GAGCCCGCAGGCCCGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

		UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
602	G016102	mG*mC*mU*CCCAGGCAGCGGGCGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
603	G016103	mC*mC*mU*CUCCAGCUGCCGGGCAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
604	G016104	mA*mG*mG*CUUUCCCCAAACUGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
605	G016105	mC*mC*mA*UCUCCACUCUGCCCCAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
606	G016106	mU*mC*mC*UCCUCACAGCCCGGCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
607	G016107	mC*mC*mA*CUCUGCCCCAUGGGCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
608	G016108	mC*mA*mG*CCUCCCGCCCGCUGCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
609	G016109	mC*mC*mC*GGCCGCCUCUCUUUUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
610	G016110	mC*mC*mU*GGGCCCACAGCCACUCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
611	G016111	mU*mC*mC*CCGGCUGCAGCCGCUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
612	G016112	mC*mG*mC*GUUCUGCUCAUCCUAGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
613	G016113	mU*mU*mC*CACAUCCUUCAGGGACUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
614	G016114	mC*mU*mC*UCCAGCUGCCGGGCAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
615	G016115	mG*mG*mG*CCCACAGCCACUCGUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
616	G016116	mC*mC*mC*CGGCCGCCUCUCUUUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
617	G016117	mU*mC*mC*AGCUGCCGGGCAUUGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
618	Exemplary guide sequence for TRBC gene (G016200 guide RNA targeting TRBC with guide	CCACACCCAAAAGGCCACAC

	sequence SEQ ID NO:618)	
619	Exemplary guide sequence for TRBC gene (G016174 guide RNA targeting TRBC with guide sequence SEQ ID NO:619)	CCCACCAGCUCAGCUCCACG
620	Exemplary guide sequence for TRBC gene (G016263 guide RNA targeting TRBC with guide sequence SEQ ID NO:620)	UCCCUAGCAGGAUCUCAUAG
621	Exemplary guide sequence for TRBC gene (G016270 guide RNA targeting TRBC with guide sequence SEQ ID NO:621)	GGCUCAAACACAGCGACCUC
622	Exemplary guide sequence for TRBC gene (G016212 guide RNA targeting TRBC with guide sequence SEQ ID NO:622)	GGCACACCAGUGUGGCCUUU
623	Exemplary guide sequence for TRBC gene (G016207 guide RNA targeting	AGGUGGCCGAGACCCUCAGG

	TRBC with guide sequence SEQ ID NO:623)	
624	Exemplary guide sequence for TRBC gene (G016205 guide RNA targeting TRBC with guide sequence SEQ ID NO:624)	ACCUGCUCUACCCCAGGCCU
625	Exemplary guide sequence for TRBC gene (G016245 guide RNA targeting TRBC with guide sequence SEQ ID NO:625)	CAUAGAGGAUGGUGGCAGAC
626	Exemplary guide sequence for TRBC gene (G016232 guide RNA targeting TRBC with guide sequence SEQ ID NO:626)	CCACGUGGAGCUGAGCUGGU
627	Exemplary guide sequence for TRBC gene (G016227 guide RNA targeting TRBC with guide sequence SEQ ID NO:627)	GGAGAAUGACGAGUGGACCC
628	Exemplary guide sequence for TRBC gene (G016273 guide	CCAGUGUGGCCUUUUGGGUG

	RNA targeting TRBC with guide sequence SEQ ID NO:628)	
629	Exemplary guide sequence for TRBC gene (G016251 guide RNA targeting TRBC with guide sequence SEQ ID NO:629)	CAAACACAGCGACCUCGGGU
630	Exemplary guide sequence for TRBC gene (G016276 guide RNA targeting TRBC with guide sequence SEQ ID NO:630)	UACCAUGGCCAUCAACACAA
631	Exemplary guide sequence for TRBC gene (G016167 guide RNA targeting TRBC with guide sequence SEQ ID NO:631)	GCGCUGACGAUCUGGGUGAC
632	Exemplary guide sequence for TRBC gene (G016253 guide RNA targeting TRBC with guide sequence SEQ ID NO:632)	CACGGACCCGCAGCCCCUCA
633	Exemplary guide sequence for TRBC gene	UCAAACACAGCGACCUCGGG

	(G016272 guide RNA targeting TRBC with guide sequence SEQ ID NO:633)	
634	Exemplary guide sequence for TRBC gene (G016258 guide RNA targeting TRBC with guide sequence SEQ ID NO:634)	CAGGGAAGAAGCCUGUGGCC
635	Exemplary guide sequence for TRBC gene (G016183 guide RNA targeting TRBC with guide sequence SEQ ID NO:635)	CAGUGUGGCCUUUUGGGUGU
636	Exemplary guide sequence for TRBC gene (G016222 guide RNA targeting TRBC with guide sequence SEQ ID NO:636)	ACCACGUGGAGCUGAGCUGG
637	Exemplary guide sequence for TRBC gene (G016233 guide RNA targeting TRBC with guide sequence SEQ ID NO:637)	GAGGGCGGGCUGCUCCUUGA
638	Exemplary guide sequence for	AGCUCAGCUCCACGUGGUCA

	TRBC gene (G016264 guide RNA targeting TRBC with guide sequence SEQ ID NO:638)	
639	Exemplary guide sequence for TRBC gene (G016255 guide RNA targeting TRBC with guide sequence SEQ ID NO:639)	GAACAAGGUGUUCCCACCCG
640	Exemplary guide sequence for TRBC gene (G016177 guide RNA targeting TRBC with guide sequence SEQ ID NO:640)	GCACACCAGUGUGGCCUUUU
641	Exemplary guide sequence for TRBC gene (G016283 guide RNA targeting TRBC with guide sequence SEQ ID NO:641)	GAGCUGGUGGGUGAAUGGGA
642	Exemplary guide sequence for TRBC gene (G016194 guide RNA targeting TRBC with guide sequence SEQ ID NO:642)	GGCUGCUCCUUGAGGGGCUG
643	Exemplary guide	CGGGUGGGAACACCUUGUUC

	sequence for TRBC gene (G016266 guide RNA targeting TRBC with guide sequence SEQ ID NO:643)	
644	Exemplary guide sequence for TRBC gene (G016196 guide RNA targeting TRBC with guide sequence SEQ ID NO:644)	CAGCUCAGCUCCACGUGGUC
645	Exemplary guide sequence for TRBC gene (G016274 guide RNA targeting TRBC with guide sequence SEQ ID NO:645)	GACGAUCUGGGUGACGGGUU
646	Exemplary guide sequence for TRBC gene (G016282 guide RNA targeting TRBC with guide sequence SEQ ID NO:646)	UAGCAGGAUCUCAUAGAGGA
647	Exemplary guide sequence for TRBC gene (G016267 guide RNA targeting TRBC with guide sequence SEQ ID NO:647)	GACCAGCACAGCAUACAGGG

648	Exemplary guide sequence for TRBC gene (G016261 guide RNA targeting TRBC with guide sequence SEQ ID NO:648)	CUGACCACGUGGAGCUGAGC
649	Exemplary guide sequence for TRBC gene (G016277 guide RNA targeting TRBC with guide sequence SEQ ID NO:649)	CCAACAGUGUCCUACCAGCA
650	Exemplary guide sequence for TRBC gene (G016285 guide RNA targeting TRBC with guide sequence SEQ ID NO:650)	CUGGUGGGUGAAUGGGAAGG
651	Exemplary guide sequence for TRBC gene (G016275 guide RNA targeting TRBC with guide sequence SEQ ID NO:651)	CUAUGAGAUCCUGCUAGGGA
652	Exemplary guide sequence for TRBC gene (G016281 guide RNA targeting TRBC with guide sequence SEQ ID	CAGGAUCUCAUAGAGGAUGG

	NO:652)	
653	Exemplary guide sequence for TRBC gene (G016280 guide RNA targeting TRBC with guide sequence SEQ ID NO:653)	GGCCACCCUGUAUGCUGUGC
654	Exemplary guide sequence for TRBC gene (G016279 guide RNA targeting TRBC with guide sequence SEQ ID NO:654)	CAACAGUGUCCUACCAGCAA
655	Exemplary guide sequence for TRBC gene (G016250 guide RNA targeting TRBC with guide sequence SEQ ID NO:655)	AGCUGAGCUGGUGGGUGAAU
656	Exemplary guide sequence for TRBC gene (G016278 guide RNA targeting TRBC with guide sequence SEQ ID NO:656)	AACAGUGUCCUACCAGCAAG
657	Exemplary guide sequence for TRBC gene (G016259 guide RNA targeting TRBC with guide	AGGCUUCUUCCCUGACCACG

	sequence SEQ ID NO:657)	
658	Exemplary guide sequence for TRBC gene (G016284 guide RNA targeting TRBC with guide sequence SEQ ID NO:658)	UCUUCUGCAGGUCAAGAGAA
659	Exemplary guide sequence for TRBC gene (G016249 guide RNA targeting TRBC with guide sequence SEQ ID NO:659)	GAGCUGAGCUGGUGGGUGAA
660	Exemplary guide sequence for TRBC gene (G016256 guide RNA targeting TRBC with guide sequence SEQ ID NO:660)	GGUCAGCGCCCUUGUGUUGA
661	Exemplary guide sequence for TRBC gene (G016190 guide RNA targeting TRBC with guide sequence SEQ ID NO:661)	AGAUCGUCAGCGCCGAGGCC
662	Exemplary guide sequence for TRBC gene (G016248 guide RNA targeting	GCUGCUCCUUGAGGGGCUGC

	TRBC with guide sequence SEQ ID NO:662)	
663	Exemplary guide sequence for TRBC gene (G016257 guide RNA targeting TRBC with guide sequence SEQ ID NO:663)	GUAUCUGGAGUCAUUGAGGG
664	Exemplary guide sequence for TRBC gene (G016260 guide RNA targeting TRBC with guide sequence SEQ ID NO:664)	AUCCUCUAUGAGAUCCUGCU
665	Exemplary guide sequence for TRBC gene (G016262 guide RNA targeting TRBC with guide sequence SEQ ID NO:665)	UCCUCUAUGAGAUCCUGCUA
666	Exemplary guide sequence for TRBC gene (G016242 guide RNA targeting TRBC with guide sequence SEQ ID NO:666)	GCAGUAUCUGGAGUCAUUGA
667	Exemplary guide sequence for TRBC gene (G016268 guide	GCUGACCAGCACAGCAUACA

	RNA targeting TRBC with guide sequence SEQ ID NO:667)	
668	Exemplary guide sequence for TRBC gene (G016265 guide RNA targeting TRBC with guide sequence SEQ ID NO:668)	ACAGGGUGGCCUUCCCUAGC
669	Exemplary guide sequence for TRBC gene (G016254 guide RNA targeting TRBC with guide sequence SEQ ID NO:669)	CGCUGACCAGCACAGCAUAC
670	Exemplary guide sequence for TRBC gene (G013015 guide RNA targeting TRBC with guide sequence SEQ ID NO:670)	AGAGAUCUCCCACACCCAAA
671	Exemplary guide sequence for TRBC gene (G014832 guide RNA targeting TRBC with guide sequence SEQ ID NO:671)	GGCUCUCGGAGAAUGACGAG
672	Exemplary guide sequence for TRBC gene	UGUGGCCAGGCACACCAGUG

	(G016168 guide RNA targeting TRBC with guide sequence SEQ ID NO:672)	
673	Exemplary guide sequence for TRBC gene (G016171 guide RNA targeting TRBC with guide sequence SEQ ID NO:673)	GCGGCUGCUCAGGCAGUAUC
674	Exemplary guide sequence for TRBC gene (G016172 guide RNA targeting TRBC with guide sequence SEQ ID NO:674)	ACACUGGUGUGCCUGGCCAC
675	Exemplary guide sequence for TRBC gene (G016179 guide RNA targeting TRBC with guide sequence SEQ ID NO:675)	CGUAGAACUGGACUUGACAG
676	Exemplary guide sequence for TRBC gene (G016180 guide RNA targeting TRBC with guide sequence SEQ ID NO:676)	CUUGACAGCGGAAGUGGUUG
677	Exemplary guide sequence for	ACUGGACUUGACAGCGGAAG

	TRBC gene (G016181 guide RNA targeting TRBC with guide sequence SEQ ID NO:677)	
678	Exemplary guide sequence for TRBC gene (G016187 guide RNA targeting TRBC with guide sequence SEQ ID NO:678)	UGAGGGUCUCGGCCACCUUC
679	Exemplary guide sequence for TRBC gene (G016188 guide RNA targeting TRBC with guide sequence SEQ ID NO:679)	AGUCCAGUUCUACGGGCUCU
680	Exemplary guide sequence for TRBC gene (G016189 guide RNA targeting TRBC with guide sequence SEQ ID NO:680)	CACCCAGAUCGUCAGCGCCG
681	Exemplary guide sequence for TRBC gene (G016191 guide RNA targeting TRBC with guide sequence SEQ ID NO:681)	GAUCGUCAGCGCCGAGGCCU

682	Exemplary guide sequence for TRBC gene (G016192 guide RNA targeting TRBC with guide sequence SEQ ID NO:682)	AUCGUCAGCGCCGAGGCCUG
683	Exemplary guide sequence for TRBC gene (G016193 guide RNA targeting TRBC with guide sequence SEQ ID NO:683)	ACUGCCUGAGCAGCCGCCUG
684	Exemplary guide sequence for TRBC gene (G016195 guide RNA targeting TRBC with guide sequence SEQ ID NO:684)	GAGCAGCCGCCUGAGGGUCU
685	Exemplary guide sequence for TRBC gene (G016197 guide RNA targeting TRBC with guide sequence SEQ ID NO:685)	AGGGCGGGCUGCUCCUUGAG
686	Exemplary guide sequence for TRBC gene (G016199 guide RNA targeting TRBC with guide sequence SEQ ID	GCUGUCAAGUCCAGUUCUAC

	NO:686)	
687	Exemplary guide sequence for TRBC gene (G016201 guide RNA targeting TRBC with guide sequence SEQ ID NO:687)	AGGCCUCGGCGCUGACGAUC
688	Exemplary guide sequence for TRBC gene (G016203 guide RNA targeting TRBC with guide sequence SEQ ID NO:688)	CUGCCUGAGCAGCCGCCUGA
689	Exemplary guide sequence for TRBC gene (G016204 guide RNA targeting TRBC with guide sequence SEQ ID NO:689)	AGACCCUCAGGCGGCUGCUC
690	Exemplary guide sequence for TRBC gene (G016206 guide RNA targeting TRBC with guide sequence SEQ ID NO:690)	CGCUGUCAAGUCCAGUUCUA

691	Exemplary guide sequence for TRBC gene (G016211 guide RNA targeting TRBC with guide sequence SEQ ID NO:691)	AAAGGCCACACUGGUGUGCC
692	Exemplary guide sequence for TRBC gene (G016214 guide RNA targeting TRBC with guide sequence SEQ ID NO:692)	UUGACAGCGGAAGUGGUUGC
693	Exemplary guide sequence for TRBC gene (G016216 guide RNA targeting TRBC with guide sequence SEQ ID NO:693)	CCACUCACCUGCUCUACCCC
694	Exemplary guide sequence for TRBC gene (G016217 guide RNA targeting TRBC with guide sequence SEQ ID NO:694)	UGGCUCAAACACAGCGACCU
695	Exemplary guide sequence for TRBC gene (G016219 guide RNA targeting TRBC with guide sequence SEQ ID	UGGCAGACAGGACCCCUUGC

	NO:695)	
696	Exemplary guide sequence for TRBC gene (G016228 guide RNA targeting TRBC with guide sequence SEQ ID NO:696)	GUUGCGGGGGUUCUGCCAGA
697	Exemplary guide sequence for TRBC gene (G016230 guide RNA targeting TRBC with guide sequence SEQ ID NO:697)	UGAGGGCGGGCUGCUCCUUG
698	Exemplary guide sequence for TRBC gene (G016238 guide RNA targeting TRBC with guide sequence SEQ ID NO:698)	UCUCCGAGAGCCCGUAGAAC
699	Exemplary guide sequence for TRBC gene (G016239 guide RNA targeting TRBC with guide sequence SEQ ID NO:699)	GGCCUCGGCGCUGACGAUCU

700	Exemplary guide sequence for TRBC gene (G016240 guide RNA targeting TRBC with guide sequence SEQ ID NO:700)	AUGACGAGUGGACCCAGGAU
701	Exemplary guide sequence for TRBC gene (G016241 guide RNA targeting TRBC with guide sequence SEQ ID NO:701)	UGACGAGUGGACCCAGGAUA
702	Exemplary guide sequence for TRBC gene (G016243 guide RNA targeting TRBC with guide sequence SEQ ID NO:702)	GCGGGGGUUCUGCCAGAAGG
703	Exemplary guide sequence for TRBC gene (G016252 guide RNA targeting TRBC with guide sequence SEQ ID NO:703)	GGUGCACAGUGGGGUCAGCA
704	Exemplary guide sequence for TRBC gene (G016269 guide RNA targeting TRBC with guide sequence SEQ ID	GGCGCUGACGAUCUGGGUGA

	NO:704)	
705	Exemplary guide sequence for TRBC gene (G016271 guide RNA targeting TRBC with guide sequence SEQ ID NO:705)	AGACAGGACCCCUUGCUGGU
706	Exemplary guide sequence for TRAC gene (G016019 guide RNA targeting TRAC with guide sequence SEQ ID NO:706)	AACAAAUGUGUCACAAAGUA
707	Exemplary guide sequence for TRAC gene (G016016 guide RNA targeting TRAC with guide sequence SEQ ID NO:707)	UUUCAAAACCUGUCAGUGAU
708	Exemplary guide sequence for TRAC gene (G016026 guide RNA targeting TRAC with guide sequence SEQ ID NO:708)	CUUACCUGGGCUGGGGAAGA

709	Exemplary guide sequence for TRAC gene (G016013 guide RNA targeting TRAC with guide sequence SEQ ID NO:709)	CCGAAUCCUCCUCCUGAAAG
710	Exemplary guide sequence for TRAC gene (G016022 guide RNA targeting TRAC with guide sequence SEQ ID NO:710)	CUGACAGGUUUUGAAAGUUU
711	Exemplary guide sequence for TRAC gene (G016023 guide RNA targeting TRAC with guide sequence SEQ ID NO:711)	CUGGGGAAGAAGGUGUCUUC
712	Exemplary guide sequence for TRAC gene (G016010 guide RNA targeting TRAC with guide sequence SEQ ID NO:712)	UCCUCCUCCUGAAAGUGGCC
713	Exemplary guide sequence for TRAC gene (G016014 guide RNA targeting TRAC with guide sequence SEQ ID	CCACUUUCAGGAGGAGGAUU

	NO:713)	
714	Exemplary guide sequence for TRAC gene (G016020 guide RNA targeting TRAC with guide sequence SEQ ID NO:714)	AUUUGUUUGAGAAUCAAAAU
715	Exemplary guide sequence for TRAC gene (G015998 guide RNA targeting TRAC with guide sequence SEQ ID NO:715)	CUUCAAGAGCAACAGUGCUG
716	Exemplary guide sequence for TRAC gene (G016025 guide RNA targeting TRAC with guide sequence SEQ ID NO:716)	AGCUGCCCUUACCUGGGCUG
717	Exemplary guide sequence for TRAC gene (G015997 guide RNA targeting TRAC with guide sequence SEQ ID NO:717)	AGAGCAACAGUGCUGUGGCC

718	Exemplary guide sequence for TRAC gene (G016029 guide RNA targeting TRAC with guide sequence SEQ ID NO:718)	AAAGCUGCCCUUACCUGGGC
719	Exemplary guide sequence for TRAC gene (G016028 guide RNA targeting TRAC with guide sequence SEQ ID NO:719)	AAGCUGCCCUUACCUGGGCU
720	Exemplary guide sequence for TRAC gene (G016021 guide RNA targeting TRAC with guide sequence SEQ ID NO:720)	UGGAAUAAUGCUGUUGUUGA
721	Exemplary guide sequence for TRAC gene (G016015 guide RNA targeting TRAC with guide sequence SEQ ID NO:721)	CACCAAAGCUGCCCUUACCU
722	Exemplary guide sequence for TRAC gene (G012086 guide RNA targeting TRAC with guide sequence SEQ ID	AGAGUCUCUCAGCUGGUACA

	NO:722)	
723	Exemplary guide sequence for TRAC gene (G013004 guide RNA targeting TRAC with guide sequence SEQ ID NO:723)	ACACGGCAGGGUCAGGGUUC
724	Exemplary guide sequence for TRAC gene (G013005 guide RNA targeting TRAC with guide sequence SEQ ID NO:724)	AGCUGGUACACGGCAGGGUC
725	Exemplary guide sequence for TRAC gene (G013006 guide RNA targeting TRAC with guide sequence SEQ ID NO:725)	CUCUCAGCUGGUACACGGCA
726	Exemplary guide sequence for TRAC gene (G013007 guide RNA targeting TRAC with guide sequence SEQ ID NO:726)	UCUCUCAGCUGGUACACGGC

727	Exemplary guide sequence for TRAC gene (G013008 guide RNA targeting TRAC with guide sequence SEQ ID NO:727)	UGGAUUUAGAGUCUCUCAGC
728	Exemplary guide sequence for TRAC gene (G013009 guide RNA targeting TRAC with guide sequence SEQ ID NO:728)	UAGGCAGACAGACUUGUCAC
729	Exemplary guide sequence for TRAC gene (G014831 guide RNA targeting TRAC with guide sequence SEQ ID NO:729)	UUCGGAACCCAAUCACUGAC
730	Exemplary guide sequence for TRAC gene (G015999 guide RNA targeting TRAC with guide sequence SEQ ID NO:730)	GCUGGUACACGGCAGGGUCA
731	Exemplary guide sequence for TRAC gene (G016000 guide RNA targeting TRAC with guide sequence SEQ ID	ACAAAACUGUGCUAGACAUG

	NO:731)	
732	Exemplary guide sequence for TRAC gene (G016001 guide RNA targeting TRAC with guide sequence SEQ ID NO:732)	GAGAAUCAAAAUCGGUGAAU
733	Exemplary guide sequence for TRAC gene (G016002 guide RNA targeting TRAC with guide sequence SEQ ID NO:733)	AAGUUCCUGUGAUGUCAAGC
734	Exemplary guide sequence for TRAC gene (G016003 guide RNA targeting TRAC with guide sequence SEQ ID NO:734)	UGCUCAUGACGCUGCGGCUG
735	Exemplary guide sequence for TRAC gene (G016004 guide RNA targeting TRAC with guide sequence SEQ ID NO:735)	UUAAUCUGCUCAUGACGCUG

736	Exemplary guide sequence for TRAC gene (G016005 guide RNA targeting TRAC with guide sequence SEQ ID NO:736)	GAUUAAACCCGGCCACUUUC
737	Exemplary guide sequence for TRAC gene (G016006 guide RNA targeting TRAC with guide sequence SEQ ID NO:737)	CUCGACCAGCUUGACAUCAC
738	Exemplary guide sequence for TRAC gene (G016007 guide RNA targeting TRAC with guide sequence SEQ ID NO:738)	UAAACCCGGCCACUUUCAGG
739	Exemplary guide sequence for TRAC gene (G016008 guide RNA targeting TRAC with guide sequence SEQ ID NO:739)	ACCCGGCCACUUUCAGGAGG
740	Exemplary guide sequence for TRAC gene (G016009 guide RNA targeting TRAC with guide sequence SEQ ID	AAAGUCAGAUUUGUUGCUCC

	NO:740)	
741	Exemplary guide sequence for TRAC gene (G016011 guide RNA targeting TRAC with guide sequence SEQ ID NO:741)	GCACCAAAGCUGCCCUUACC
742	Exemplary guide sequence for TRAC gene (G016012 guide RNA targeting TRAC with guide sequence SEQ ID NO:742)	AUCCUCCUCCUGAAAGUGGC
743	Exemplary guide sequence for TRAC gene (G016017 guide RNA targeting TRAC with guide sequence SEQ ID NO:743)	UUCAAAACCUGUCAGUGAUU
744	Exemplary guide sequence for TRAC gene (G016018 guide RNA targeting TRAC with guide sequence SEQ ID NO:744)	UGUGCUAGACAUGAGGUCUA

745	Exemplary guide sequence for TRAC gene (G016024 guide RNA targeting TRAC with guide sequence SEQ ID NO:745)	CGUCAUGAGCAGAUUAAACC
746	Exemplary guide sequence for TRAC gene (G016027 guide RNA targeting TRAC with guide sequence SEQ ID NO:746)	UCAAGGCCCCUCACCUCAGC
763-800		Not Used
801	G016200 guide RNA targeting TRBC with guide sequence SEQ ID NO:618	mC*mC*mA*CACCCAAAAGGCCACACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
802	G016174 guide RNA targeting TRBC with guide sequence SEQ ID NO:619	mC*mC*mC*ACCAGCUCAGCUCCACGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
803	G016263 guide RNA targeting TRBC with guide sequence SEQ ID NO:620	mU*mC*mC*CUAGCAGGAUCUCAUAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
804	G016270 guide RNA targeting TRBC with guide sequence SEQ ID NO:621	mG*mG*mC*UCAAACACAGCGACCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
805	G016212 guide	mG*mG*mC*ACACCAGUGUGGCCUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

	RNA targeting TRBC with guide sequence SEQ ID NO:622	UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
806	G016207 guide RNA targeting TRBC with guide sequence SEQ ID NO:623	mA*mG*mG*UGGCCGAGACCCUCAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
807	G016205 guide RNA targeting TRBC with guide sequence SEQ ID NO:624	mA*mC*mC*UGCUCUACCCCAGGCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
808	G016245 guide RNA targeting TRBC with guide sequence SEQ ID NO:625	mC*mA*mU*AGAGGAUGGUGGCAGACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
809	G016232 guide RNA targeting TRBC with guide sequence SEQ ID NO:626	mC*mC*mA*CGUGGAGCUGAGCUGGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
810	G016227 guide RNA targeting TRBC with guide sequence SEQ ID NO:627	mG*mG*mA*GAAUGACGAGUGGACCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
811	G016273 guide RNA targeting TRBC with guide sequence SEQ ID NO:628	mC*mC*mA*GUGUGGCCUUUUGGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
812	G016251 guide RNA targeting TRBC with guide sequence SEQ ID NO:629	mC*mA*mA*ACACAGCGACCUCGGGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

813	G016276 guide RNA targeting TRBC with guide sequence SEQ ID NO:630	mU*mA*mC*CAUGGCCAUCAACACAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
814	G016167 guide RNA targeting TRBC with guide sequence SEQ ID NO:631	mG*mC*mG*CUGACGAUCUGGGUGACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
815	G016253 guide RNA targeting TRBC with guide sequence SEQ ID NO:632	mC*mA*mC*GGACCCGCAGCCCCUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
816	G016272 guide RNA targeting TRBC with guide sequence SEQ ID NO:633	mU*mC*mA*AACACAGCGACCUCGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
817	G016258 guide RNA targeting TRBC with guide sequence SEQ ID NO:634	mC*mA*mG*GGAAGAAGCCUGUGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
818	G016183 guide RNA targeting TRBC with guide sequence SEQ ID NO:635	mC*mA*mG*UGUGGCCUUUUGGGUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
819	G016222 guide RNA targeting TRBC with guide sequence SEQ ID NO:636	mA*mC*mC*ACGUGGAGCUGAGCUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
820	G016233 guide RNA targeting TRBC with guide sequence SEQ ID	mG*mA*mG*GGCGGGCUGCUCCUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	NO:637	
821	G016264 guide RNA targeting TRBC with guide sequence SEQ ID NO:638	mA*mG*mC*UCAGCUCCACGUGGUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
822	G016255 guide RNA targeting TRBC with guide sequence SEQ ID NO:639	mG*mA*mA*CAAGGUGUUCCCACCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
823	G016177 guide RNA targeting TRBC with guide sequence SEQ ID NO:640	mG*mC*mA*CACCAGUGUGGCCUUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
824	G016283 guide RNA targeting TRBC with guide sequence SEQ ID NO:641	mG*mA*mG*CUGGUGGGUGAAUGGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
825	G016194 guide RNA targeting TRBC with guide sequence SEQ ID NO:642	mG*mG*mC*UGCUCCUUGAGGGGCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
826	G016266 guide RNA targeting TRBC with guide sequence SEQ ID NO:643	mC*mG*mG*GUGGGAACACCUUGUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
827	G016196 guide RNA targeting TRBC with guide sequence SEQ ID NO:644	mC*mA*mG*CUCAGCUCCACGUGGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
828	G016274 guide RNA targeting TRBC with guide	mG*mA*mC*GAUCUGGGUGACGGGUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	sequence SEQ ID NO:645	
829	G016282 guide RNA targeting TRBC with guide sequence SEQ ID NO:646	mU*mA*mG*CAGGAUCUCAUAGAGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
830	G016267 guide RNA targeting TRBC with guide sequence SEQ ID NO:647	mG*mA*mC*CAGCACAGCAUACAGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
831	G016261 guide RNA targeting TRBC with guide sequence SEQ ID NO:648	mC*mU*mG*ACCACGUGGAGCUGAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
832	G016277 guide RNA targeting TRBC with guide sequence SEQ ID NO:649	mC*mC*mA*ACAGUGUCCUACCAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
833	G016285 guide RNA targeting TRBC with guide sequence SEQ ID NO:650	mC*mU*mG*GUGGGUGAAUGGGAAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
834	G016275 guide RNA targeting TRBC with guide sequence SEQ ID NO:651	mC*mU*mA*UGAGAUCCUGCUAGGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
835	G016281 guide RNA targeting TRBC with guide sequence SEQ ID NO:652	mC*mA*mG*GAUCUCAUAGAGGAUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
836	G016280 guide RNA targeting	mG*mG*mC*CACCCUGUAUGCUGUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	TRBC with guide sequence SEQ ID NO:653	
837	G016279 guide RNA targeting TRBC with guide sequence SEQ ID NO:654	mC*mA*mA*CAGUGUCCUACCAGCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
838	G016250 guide RNA targeting TRBC with guide sequence SEQ ID NO:655	mA*mG*mC*UGAGCUGGUGGGUGAAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
839	G016278 guide RNA targeting TRBC with guide sequence SEQ ID NO:656	mA*mA*mC*AGUGUCCUACCAGCAAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
840	G016259 guide RNA targeting TRBC with guide sequence SEQ ID NO:657	mA*mG*mG*CUUCUUCCCUGACCACGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
841	G016284 guide RNA targeting TRBC with guide sequence SEQ ID NO:658	mU*mC*mU*UCUGCAGGUCAAGAGAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
842	G016249 guide RNA targeting TRBC with guide sequence SEQ ID NO:659	mG*mA*mG*CUGAGCUGGUGGGUGAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
843	G016256 guide RNA targeting TRBC with guide sequence SEQ ID NO:660	mG*mG*mU*CAGCGCCCUUGUGUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
844	G016190 guide	mA*mG*mA*UCGUCAGCGCCGAGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm

	RNA targeting TRBC with guide sequence SEQ ID NO:661	UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
845	G016248 guide RNA targeting TRBC with guide sequence SEQ ID NO:662	mG*mC*mU*GCUCCUUGAGGGGCUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
846	G016257 guide RNA targeting TRBC with guide sequence SEQ ID NO:663	mG*mU*mA*UCUGGAGUCAUUGAGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
847	G016260 guide RNA targeting TRBC with guide sequence SEQ ID NO:664	mA*mU*mC*CUCUAUGAGAUCCUGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
848	G016262 guide RNA targeting TRBC with guide sequence SEQ ID NO:665	mU*mC*mC*UCUAUGAGAUCCUGCUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
849	G016242 guide RNA targeting TRBC with guide sequence SEQ ID NO:666	mG*mC*mA*GUAUCUGGAGUCAUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
850	G016268 guide RNA targeting TRBC with guide sequence SEQ ID NO:667	mG*mC*mU*GACCAGCACAGCAUACAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
851	G016265 guide RNA targeting TRBC with guide sequence SEQ ID NO:668	mA*mC*mA*GGGUGGCCUUCCCUAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

852	G016254 guide RNA targeting TRBC with guide sequence SEQ ID NO:669	mC*mG*mC*UGACCAGCACAGCAUACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
853	G016019 guide RNA targeting TRAC with guide sequence SEQ ID NO:706	mA*mA*mC*AAAUGUGUCACAAAGUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
854	G016016 guide RNA targeting TRAC with guide sequence SEQ ID NO:707	mU*mU*mU*CAAAACCUGUCAGUGAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
855	G016026 guide RNA targeting TRAC with guide sequence SEQ ID NO:708	mC*mU*mU*ACCUGGGCUGGGGAAGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
856	G016013 guide RNA targeting TRAC with guide sequence SEQ ID NO:709	mC*mC*mG*AAUCCUCCUCCUGAAAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
857	G016022 guide RNA targeting TRAC with guide sequence SEQ ID NO:710	mC*mU*mG*ACAGGUUUUGAAAGUUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
858	G016023 guide RNA targeting TRAC with guide sequence SEQ ID NO:711	mC*mU*mG*GGGAAGAAGGUGUCUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
859	G016010 guide RNA targeting TRAC with guide sequence SEQ ID	mU*mC*mC*UCCUCCUGAAAGUGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	NO:712	
860	G016014 guide RNA targeting TRAC with guide sequence SEQ ID NO:713	mC*mC*mA*CUUUCAGGAGGAGGAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
861	G016020 guide RNA targeting TRAC with guide sequence SEQ ID NO:714	mA*mU*mU*UGUUUGAGAAUCAAAAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
862	G015998 guide RNA targeting TRAC with guide sequence SEQ ID NO:715	mC*mU*mU*CAAGAGCAACAGUGCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
863	G016025 guide RNA targeting TRAC with guide sequence SEQ ID NO:716	mA*mG*mC*UGCCCUUACCUGGGCUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
864	G015997 guide RNA targeting TRAC with guide sequence SEQ ID NO:717	mA*mG*mA*GCAACAGUGCUGUGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
865	G016029 guide RNA targeting TRAC with guide sequence SEQ ID NO:718	mA*mA*mA*GCUGCCCUUACCUGGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
866	G016028 guide RNA targeting TRAC with guide sequence SEQ ID NO:719	mA*mA*mG*CUGCCCUUACCUGGGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
867	G016021 guide RNA targeting TRAC with guide	mU*mG*mG*AAUAAUGCUGUUGUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	sequence SEQ ID NO:720	
868	G016015 guide RNA targeting TRAC with guide sequence SEQ ID NO:721	mC*mA*mC*CAAAGCUGCCCUUACCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
869	G019761 Guide RNA Targeting CD38	mG*mA*mU*CCUCGUCGUGGUGCUCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
870	G019762 Guide RNA Targeting CD38	mC*mC*mU*CGUCGUGGUGCUCGCGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
871	G019763 Guide RNA Targeting CD38	mG*mC*mA*UCGCGCCAGGACGGUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
872	G019764 Guide RNA Targeting CD38	mU*mG*mU*ACUUGACGCAUCGCGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
873	G019765 Guide RNA Targeting CD38	mC*mA*mC*CGGGCUGAACUCGCAGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
874	G019766 Guide RNA Targeting CD38	mU*mC*mG*CGGUGGUCGUCCCGAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
875	G019767 Guide RNA Targeting CD38	mC*mC*mA*CCGCGAGCACCACGACGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
876	G019768 Guide RNA Targeting CD38	mC*mA*mU*CGCGCCAGGACGGUCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
877	G019769 Guide RNA Targeting CD38	mG*mA*mC*GGUCUCGGGAAAGCGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
878	G019770 Guide RNA Targeting CD38	mG*mC*mG*CUUUCCCGAGACCGUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
879	G019771 Guide RNA Targeting	mC*mU*mU*GACGCAUCGCGCCAGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	CD38	
880	G019772 Guide RNA Targeting CD38	mU*mG*mC*GAGUUCAGCCCGGUGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
881	G019773 Guide RNA Targeting CD38	mU*mG*mC*UCGCGGUGGUCGUCCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
882	G019774 Guide RNA Targeting CD38	mA*mG*mG*GUUUGUCCCCGGACACCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
883	G019775 Guide RNA Targeting CD38	mG*mC*mG*AGUUCAGCCCGGUGUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
884	G019776 Guide RNA Targeting CD38	mG*mG*mU*CUCGGGAAAGCGCUUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
885	G019777 Guide RNA Targeting CD38	mC*mG*mA*GUUCAGCCCGGUGUCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
886	G019778 Guide RNA Targeting CD38	mC*mC*mG*GCAGCAGGGUUUGUCCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
887	G019779 Guide RNA Targeting CD38	mG*mU*mU*GGGCUCUCCUAGAGAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
888	G019780 Guide RNA Targeting CD38	mC*mG*mA*GCACCACGACGAGGAUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
889	G019781 Guide RNA Targeting CD38	mG*mG*mC*CAACUGCGAGUUCAGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
890	G019782 Guide RNA Targeting CD38	mU*mA*mC*UGACGCCAAGACAGAGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
891	G019783 Guide RNA Targeting CD38	mU*mC*mG*CCAACCCACCUCAUCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
892	G019784 Guide RNA Targeting	mC*mC*mG*GGGACAAACCCUGCUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	CD38	
893	G019785 Guide RNA Targeting CD38	mG*mG*mA*AAGCGCUUGGUGGUGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
894	G019786 Guide RNA Targeting CD38	mU*mC*mU*CCUAGAGAGCCGGCAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
895	G019787 Guide RNA Targeting CD38	mC*mG*mC*CAGCAGUGGAGCGGUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
896	G019788 Guide RNA Targeting CD38	mC*mU*mC*CACUGCUGGCGCCACCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
897	G019789 Guide RNA Targeting CD38	mC*mA*mG*GGUUUGUCCCCGGACACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
898	G019790 Guide RNA Targeting CD38	mC*mU*mG*UCUUGGCGUCAGUAUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
899	G019791 Guide RNA Targeting CD38	mU*mG*mC*CCGGACCGCUCCACUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
900	G019792 Guide RNA Targeting CD38	mC*mU*mC*CUAGAGAGCCGGCAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
901	G019793 Guide RNA Targeting CD38	mC*mC*mU*GGUCCUGAUCCUCGUCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
902	G019794 Guide RNA Targeting CD38	mU*mC*mC*CGAGGUGGCGCCAGCAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
903	G019795 Guide RNA Targeting CD38	mG*mC*mG*CCAGCAGUGGAGCGGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
904	G019796 Guide RNA Targeting CD38	mU*mC*mC*ACUGCUGGCGCCACCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
905	G019797 Guide RNA Targeting	mA*mG*mG*AGAGCCCAACUCUGUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	CD38	
906	G019798 Guide RNA Targeting CD38	mA*mC*mU*GACGCCAAGACAGAGUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
907	G019799 Guide RNA Targeting CD38	mA*mA*mC*CCUGCUGCCGGCUCUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
908	G019800 Guide RNA Targeting CD38	mU*mA*mU*CAGCCACUAAUGAAGUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
909	G019801 Guide RNA Targeting CD38	mU*mG*mA*AAGCAUCCCAUACACUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
910	G019802 Guide RNA Targeting CD38	mC*mC*mC*CCAAUUACCUUGUUGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
911	G019803 Guide RNA Targeting CD38	mA*mU*mG*UAGACUGCCAAAGUGUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
912	G019804 Guide RNA Targeting CD38	mA*mA*mG*UGUAUGGGAUGCUUUCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
913	G019805 Guide RNA Targeting CD38	mA*mA*mU*UACCUUGUUGCAAGGUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
914	G019806 Guide RNA Targeting CD38	mU*mG*mA*GUUCCCAACUUCAUUAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
915	G019807 Guide RNA Targeting CD38	mU*mG*mU*AGACUGCCAAAGUGUAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
916	G019808 Guide RNA Targeting CD38	mA*mG*mU*GUAUGGGAUGCUUUCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
917	G019809 Guide RNA Targeting CD38	mC*mU*mA*UCAGCCACUAAUGAAGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
918	G019810 Guide RNA Targeting	mU*mC*mU*UCUUCAGUAAUGUUGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	CD38	
919	G019811 Guide RNA Targeting CD38	mU*mC*mU*GGCCCAUCAGUUCACACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
920	G019812 Guide RNA Targeting CD38	mG*mC*mG*GGACAUGUUCACCCUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
921	G019813 Guide RNA Targeting CD38	mC*mC*mA*GCGGGACAUGUUCACCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
922	G019814 Guide RNA Targeting CD38	mA*mG*mC*CUAGCAGCGUGUCCUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
923	G019815 Guide RNA Targeting CD38	mU*mG*mA*AUUCACCACACCAUGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
924	G019816 Guide RNA Targeting CD38	mA*mU*mC*AGUUCACACAGGUCCAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
925	G019817 Guide RNA Targeting CD38	mG*mC*mU*GAUGACCUCACAUGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
926	G019818 Guide RNA Targeting CD38	mG*mC*mC*UAGCAGCGUGUCCUCCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
927	G019819 Guide RNA Targeting CD38	mA*mC*mC*AUGUGAGGUCAUCAGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
928	G019820 Guide RNA Targeting CD38	mU*mC*mA*GUUCACACAGGUCCAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
929	G019821 Guide RNA Targeting CD38	mC*mC*mA*GGGUGAACAUGUCCCGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
930	G019822 Guide RNA Targeting CD38	mG*mC*mU*GGACCUGUGUGAACUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
931	G019823 Guide RNA Targeting	mA*mC*mC*CUGGAGGACACGCUGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	CD38	
932	G019824 Guide RNA Targeting CD38	mC*mU*mG*GACCUGUGUGAACUGAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
933	G019825 Guide RNA Targeting CD38	mU*mC*mU*GGAAAACGGUUUCCCGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
937	G019829 Guide RNA Targeting CD38	mG*mU*mU*UUCCAGAAUACUGAAACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
938	G019830 Guide RNA Targeting CD38	mG*mG*mG*AUCCAUUGAGCAUCACAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
939	G019831 Guide RNA Targeting CD38	mC*mA*mU*CACAUGGACCACAUCACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
940	G019832 Guide RNA Targeting CD38	mG*mU*mG*GUCCAUGUGAUGCUCAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
941	G019833 Guide RNA Targeting CD38	mA*mG*mA*GAAGGUUCAGACACUAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
942	G019834 Guide RNA Targeting CD38	mU*mC*mU*AGUGUCUGAACCUUCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
943	G019835 Guide RNA Targeting CD38	mG*mG*mU*UCAGACACUAGAGGCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
944	G019836 Guide RNA Targeting CD38	mC*mC*mA*UAAUUUGCAACCAGAGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
945	G019837 Guide RNA Targeting CD38	mC*mC*mU*UCUCUGGUUGCAAAUUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
946	G019838 Guide RNA Targeting CD38	mA*mG*mG*UUCAGACACUAGAGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
947	G019839 Guide RNA Targeting	mC*mU*mA*GAGGCCUGGGUGAUACAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

	CD38	
948	G019840 Guide RNA Targeting CD38	mU*mU*mU*UAGCACUUUUGGGAGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
949	G019841 Guide RNA Targeting CD38	mU*mC*mU*UCCACCAUGUAUCACCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
950	G019842 Guide RNA Targeting CD38	mC*mU*mU*CCCCAGAGACUUAUGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
951	G019843 Guide RNA Targeting CD38	mU*mU*mA*CCUGUAGAUAUUCUUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
952	G019844 Guide RNA Targeting CD38	mG*mC*mU*CUUUUAUGGUGGGAUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
953	G019845 Guide RNA Targeting CD38	mG*mG*mA*UCCCACCAUAAAAGAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
954	G019846 Guide RNA Targeting CD38	mC*mG*mA*UUCCAGCUCUUUUAUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
955	G019847 Guide RNA Targeting CD38	mG*mA*mU*UCCAGCUCUUUUAUGGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
956	G019848 Guide RNA Targeting CD38	mA*mA*mU*CGAUUCCAGCUCUUUUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
957-959		Not Used
960	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC01	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSVWRHTSQNTSNHVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLF
961	Amino acid sequence of exemplary	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSVWRHTSQNTSNHVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLFIYIARLYH

	cytidine deaminase; see TABLE 58, BC02	
962	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC03	MNWNALRSKAIEVSRHAYAPYSGFPVGAAALVDDGRTVTGCNVENVSYGLGLCAECAVVCALHSGGGGRLVALSCVGPDGGVLMPCGRCRQ VLLEHGGPELLIDHAHGPRPLRELLPDAFGPDDLGRR
963	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC04	MTHHALIEAAKAAREKAYAPYSNFKVGAALVTNDGKVFHGCNVENASYGLCNCAERTALFSALAAGYRPGEFAAIAVVGETHGPIAPCGAC RQVMIELGKPTLEVVLTNMQGDVRVTSAGDLLPDAFYLA
964	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC05	MPDIDWKQLRDKATQVAAGAYAPYSRFPVGAAALVDDGRVVTGCNVENVSYGLALCAECGVVCALHATGGGRLVALACVDGRGAPLMPCGR CRQLLFEHGGPELLVDHLAGPRRLGDLLPEPFHADLTGEPT
965	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC06	MNSKQLIQEAIEARKQAYVPYSKFQVGAALLTQDGKVYRGCNVENASYGLCNCAERTALFKAVSEGDKEFVAIAIVADTKRPVPPCGACRQ VMVELCKQDTKVYLSNLHGDVQETTVGELLPGAFLAEDLHE
966	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC07	MPDVDWNMLRGNATQAAAGAYVPYSRFAVGAAALVDDGRVVTGCNVENVSYGLTLCAECAVVCALHSTGGGRLLALACVDGHGSVLMPCGR CRQVLLEHGGSELLIDHPVRPRRLGDLLPDAFGLDDLPRERR
967	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC08	MGDVNWDTLQKAAVAARANSYAPYSNFPVGVAGFVNDGRLITGVNVENASYGLALCAECSMISALYATGGGRLVAVYCVDGNGDSLMPCGR CRQLLYEHGGPELKIMTPKGVQTMAQLLPQAFNPQERIFGNDE

968	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC09	MNRQELITEALKARDMAYAPYSKFQVGAALLTKDGKVYRGCNIENAAYSMCNCAERTALFKAVSEGDTEFQMLAVAADTPGPVSPCGACRQ VISELCTKDVIVVLTNLQGQIKEMTVEELLPGAFSSEDLHDERKL
969	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC10	MKVGGIEDRQLEALKRAALKACELSYSPYSHFRVGCSILTNNDVIFTGANVENASYSNCICAERSAMIQVLMAGHRSGWKCMVICGDSEDQ CVSPCGVCRQFINEFVVKDFPIVMLNSTGSRSKVMTMGELLPMAFGPSHLN
970	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC11	MNIEQQLYDVVKQLIEQRYPNDWGGAAAIRVEDGTIYTSVAPDVINASTELCMETGAILEAHKFQKKVTHSICLARENEHSELKVLSPCGV CQERLFYWGPEVQCAITNAKQDIIFKPLKELQPYHWTEAYHDEMVKEWSTR
971	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC12	MAQERPSCAVEPEHVQRLLLSSREAKKSAYCPYSRFPVGAALLTGDGRIFSGCNIENACYPLGVCAERTAIQKAISEGYKDFRAIAISSDL QEEFISPCGACRQVMREFGTDWAVYMTKPDGTFVVRTVQELLPASFGPEDLQKIQ
972	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC13	MAQKRPACTLKPECVQQLLVCSQEAKKSAYCPYSHFPVGAALLTQEGRIFKGCNIENACYPLGICAERTAIQKAVSEGYKDFRAIAIASDM QDDFISPCGACRQVMREFGTNWPVYMTKPDGTYIVMTVQELLPSSFGPEDLQKTQ
973	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC14	MVTGGMASKWDQKGMDIAYEEAALGYKEGGVPIGGCLINNKDGSVLGRGHNMRFQKGSATLHGEISTLENCGRLEGKVYKDTTLYTTLSPC DMCTGAIIMYGIPRCVVGENVNFKSKGEKYLQTRGHEVVVVDDERCKKIMKQFIDERPQDWFEDIGE
974	Amino acid sequence of exemplary	MTTTKANLTEFEQQLVDKAIGAMENAYCKYSNFKVGAALVCDDGEIIIGANHENASYGATICAERSAIVTALTKGHRKFKYIVVATELEAP CSPCGVCRQVLIEFGDYKVILGSSTSDQIIETTTYELLPYAFTPKSLDDHEKETEERKHHNDHNNKE

	cytidine deaminase; see TABLE 58, BC15	
975	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC16	MAANSLPQDISDVELVHLARAAMKRAHCPYSKFPVGAALLTESGEIVQGCNVENASYGGTICAERSAIVSAVSQGYTKFRAIAVVTELSEP ASPCGLCRQFLVEFGDYKVVVGTASNNKILITSTRALLPFAFTPESLDTFEQEKASEAKGLKQDDATEHNVTVVS
976	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC17	MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRNQVDSETHCHAERCFLSWFCDDILSPNTKYQVTWY TSWSPCPDCAGEVAEFLARHSNVNLTIFTARLYYFQYPCYQEGLRSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPFKPWKGLKTNFRLLK RRLRESLQ
977	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC18	MNPQIRNPMEAMYPHIFYFHFKNLLKACGRNESWLCFTMEVTKHHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWY TSWSPCPECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDFVSCWKNFVYSDDEPFKPWKGLQTNFRLLK RRLREILQ
978	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC19	MNDALHIGLPPFLVQANNEPRVLAAPEARMGYVLELVRANIAADGGPFAAAVFERDSGLLIAAGTNRVVPGRCSAAHAEILALSLAQAKLD THDLSADGLPACELVTSAEPCVMCFGAVIWSGVRSLVCAARSDDVEAIGFDEGPRPENWMGGLEARGITVTTGLLRDAACALLREYNACNG VIYNARCGVHK
979	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC20	MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCA RHVADFLRGNPNLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLSRQLRRILLP LYEVDDLRDAFRTLGL
980	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC21	MDSLLMKQKKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSCSLDFGHLRNKSGCHVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCA RHVAEFLRWNPNLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGLHENSVRLTRQLRRILLP LYEVDDLRDAFRMLGF

981	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC23	MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFINEIKSMGLDETQCYQVTCYLTWSPCSSCAWE LVDFIKAHDHLNLGIFASRLYYHWCKPQQKGLRLLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLERIKI PGVRAQGRYMDILCDAEV
982	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC24	MAQKEEAAVATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERLPANFFKFQFRNVEYSSGRNKTFLCYVVEAQGKGGQVQASRGYLE DEHAAAHAEEAFFNTILPAFDPALRYNVTWYVSSSPCAACADRIIKTLSKTKNLRLLILVGRLFMWEEPEIQAALKKLKEAGCKLRIMKPQ DFEYVWQNFVEQEEGESKAFQPWEDIQENFLYYEEKLADILK
983	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC25	MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRLPVNFFKFQFRNVEYSSGRNKTFLCYVVEVQSKGGQAQATQGYLE DEHAGAHAEEAFFNTILPAFDPALKYNVTWYVSSSPCAACADRILKTLSKTKNLRLLILVSRLFMWEEPEVQAALKKLKEAGCKLRIMKPQ DFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK
984	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC26	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSVWRHTSQNTSNHVEVNFLEKFTTERYFRPNTRCSITWFLSWS PCGECSRAITEFLSRHPYVTLFIYIARLYHHTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPPSNEAYWPRYPHLWVKLYVLEL YCIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK
985	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC28	MNSKTGPSVGDATLRRRIKPWEFVAFFNPQELRKETCLLYEIKWGNQNIWRHSNQNTSQHAEINFMEKFTAERHFNSSVRCSITWFLSWSP CWECSKAIRKFLDHYPNVTLAIFISRLYWHMDQQHRQGLKELVHSGVTIQIMSYSEYHYCWRNFVDYPQGEEDYWPKYPYLWIMLYVLELH CIILGLPPCLKISGSHSNQLALFSLDLQDCHYQKIPYNVLVATGLVQPFVTWR
986	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC29	MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWGMSRKIWRSSGKNTTNHVEVNFIKKFTSERDFHPSMSCSITWFLSWS PCWECSQAIREFLSRHPGVTLVIYVARLFWHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWMMLYALEL HCIILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR
987	Amino acid sequence of exemplary	MASEKGPSNKDYTLRRRIEPWEFEVFFDPQELRKEACLLYEIKWGASSKTWRSSGKNTTNHVEVNFLEKLTSEGRLGPSTCCSITWFLSWS PCWECSMAIREFLSQHPGVTLIIFVARLFQHMDRRNRQGLKDLVTSGVTVRVMSVSEYCYCWENFVNYPPGKAAQWPRYPPRWMLMYALEL YCIILGLPPCLKISRRHQKQLTFFSLTPQYCHYKMIPPYILLATGLLQPSVPWR

	cytidine deaminase; see TABLE 58, BC30	
988	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC31	MKVSLAGQTVDVKKILNEIPKRTVTAALLEGGEIVAVEEADDEHAERKLVRRHDVEGKVVFVTARPCLYCARELAEAGVAGVVYLGRGRGL GPYYLARSGVEVVEVHPDEPLGYDPVDRLDVLLTFGGNPYLTEEDVAARVYCLLTGRGFDADIAPAPENLSGRVEIMVTRGDPDEAVELLK EELPVFRIRRFLISGEFDRDELRERILEDIEPRILDPFAVRARIARAGAFSSSREAEVFIGDVLTSVGREVNLNDPRTVVTVDVLGPRVSV GVEKR
989	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC32	MHPRFQTAFAQLADNLQSALEPILADKYFPALLTGEQVSSLKSATGLDEDALAFALLPLAAACARTPLSNFNVGAIARGVSGTWYFGANME FIGATMQQTVHAEQSAISHAWLSGEKALAAITVNYTPCGHCRQFMNELNSGLDLRIHLPGREAHALRDYLPDAFGPKDLEIKTLLMDEQDH GYALTGDALSQAAIAAANRSHMPYSKSPSGVALECKDGRIFSGSYAENAAFNPTLPPLQGALILLNLKGYDYPDIQRAVLAEKADAPLIQW DATSATLKALGCHSIDRVLLA
990	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC33	MRNRIEQALQQMPASFAPYLRELVLAKDFDATFSAEQYQQLLTLSGLEDADLRVALLPIAAAYSYAPISEFYVGAIVRGISGRLYLGANME FTGAQLGQTVHAEQCAISHAWMKGEKGVADITINFSPCGHCRQFMNELTTASSLKIQLPKRAAKTLQEYLPESFGPADLGIDSGLMSPVNH GKTSDDDEELIQQALRAMNISHSPYTQNFSGVALKMRSGAIYLGAYAENAAFNPSLPPLQVALAQAMMMGESFEDIEAAALVESATGKISH LADTQATLEVINPDIPLSYLSL
991	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC34	MSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKRLLTNFRYQD SKLQEILRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITCYLTWS PCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRR IKESWGLQDLVNDFGNLQLGPPMS
992	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC35	MDKPSFVIQSKEAESAAKQLGVSVIQLLPSLVKPAQSYARTPISKFNVAVVGLGSSGRIFLGVNVEFPNLPLHHSIHAEQFLVTNLTLNGE RHLNFFAVSAAPCGHCRQFLQEIRDAPEIKILITDPNNSADSDSAADSDGFLRLGSFLPHRFGPDDLLGKDHPLLLESHDNHLKISDLDSI CNGNTDSSADLKQTALAAANRSYAPYSLCPSGVSLVDCDGKVYRGWYMESAAYNPSMGPVQAALVDYVANGGGGGYERIVGAVLVEKEDAV VRQEHTARLLLETISPKCEFKVFHCYEA
993	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC36	MVEPMDPRTFVSNFNNRPILSGLNTVWLCCEVKTKDPSGPPLDAKIFQGKVYSKAKYHPEMRFLRWFHKWRQLHHDQEYKVTWYVSWSPCT RCANSVATFLAKDPKVTLTIFVARLYYFWKPDYQQALRILCQKRGGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPRNNLPKHYTLLQATL GELLRHLMDPGTFTSNFNNKPWVSGQHETYLCYKVERLHNDTWVPLNQHRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQYRVTC FTSWSPCFSCAQEMAKFISNNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYSEFEYCWDTFVDRQGRPFQPWDGLDEHSQALS GRLRAI

994	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC37	MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPRLDAKIFRGQVYSQPEHHAEMCFLSWFCGNQLPAYKCFQITWFV SWTPCPDCVAKLAEFLAEHPNVTLTISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFDDNYAFLHR TLKEILRNPMEAMYPHIFYFHFKNLRKAYGRNESWLCFTMEVVKHHSPVSWKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTW YTSWSPCPECAGEVAEFLARHSNVNLTIFTARLYYFWDTDYQEGLRSLSQEGASVEIMGYKDFKYCWENFVYNDDEPFKPWKGLKYNFLFL DSKLQEILE
995	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC38	MNPQIRNMVEQMEPDIFVYYFNNRPILSGRNTVWLCYEVKTKDPSGPPLDANIFQGKLYPEAKDHPEMKFLHWFRKWRQLHRDQEYEVTWY VSWSPCTRCANSVATFLAEDPKVTLTIFVARLYYFWKPDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDGQGKPFKPRKNLPKHY TLLHATLGELLRHVMDPGTFTSNFNNKPWVSGQRETYLCYKVERSHNDTWVLLNQHRGFLRNQAPDRHGFPKGRHAELCFLDLIPFWKLDD QQYRVTCFTSWSPCFSCAQKMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLHRDGAKIAVMNYSEFEYCWDTFVDRQGRPFQPWDGLD EHSQALSGRLRAI
996	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC39	MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQVYFKPQYHAEMCFLSWFCGNQLPAYKCFQITWF VSWTPCPDCVAKLAEFLSEHPNVTLTISAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQQFMPWYKFDENYAFLH RTLKEILRYLMDPDTFTFNFNNDPLVLRRRQTYLCYEVERLDNGTWVLMDQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIY RVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVYRQGCPFQPWDGLEE HSQALSGRLRAILQNQGN
997	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC40	MKPQIRNMVEPMDPRTFVSNFNNRPILSGLDTVWLCCEVKTKDPSGPPLDAKIFQGKVYPKAKYHPEMRFLRWFHKWRQLHHDQEYKVTWY VSWSPCTRCANSVATFLAKDPKVTLTIFVARLYYFWKPDYQQALRILCQKRDGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPWNNLPKHY TLLQATLGELLRHLMDPGTFTSNFNNKPWVSGQHETYLCYKVERLHNDTWVPLNQHRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDG QQYRVTCFTSWSPCFSCAQEMAKFISNNEHVSLCIFAARIYDDQGRYQEGLRTLHRDGAKIAMMNYSEFEYCWDTFVDRQGRPFQPWDGLD EHSQALSGRLRAILQNQGN
998	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC41	MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQVYSELKYHPEMRFFHWFSKWRKLHRDQEYEVTWY ISWSPCTKCTRDMATFLAEDPKVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYY ILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDL DQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGL DEHSQDLSGRLRAILQNQEN
999	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC42	MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPSRPPLDAKIFRGQVYSKLKYHPEMRFFHWFSKWRKLHRDQEYEVTWY ISWSPCTKCTRDVATFLAEDPKVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYY ILLHIMLGEILRHSMDPPTFTSNFNNELWVRGRHETYLCYEVERLHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDL HQDYRVTCFTSWSPCFSCAQEMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLAKAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGL EEHSQALSGRLRAILQNQGN
1000	Amino acid sequence of exemplary	MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGPVLPKRQSNHRQEVYFRFENHAEMCFLSWFCGNR LPANRRFQITWFVSWNPCLPCVVKVTKFLAEHPNVTLTISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVCNEGQPFMP WYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFKNLLKACGRNESWLCFTMEVTKHHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCD

	cytidine deaminase; see TABLE 58, BC43	DILSPNTNYEVTWYTSWSPCPECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDFVSCWKNFVYSDDEPF KPWKGLQTNFRLLKRRLREILQ
1001	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC44	MQPQRLGPRAGMGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWF HDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDN GGRRFRPWKRLLTNFRYQDSKLQEILRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFLD KIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPF WPWKGLEIISRRTQRRLRRIKESRSHAS	
1002	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC45	MQPQRLGPRAGMGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWF HDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDN GGRRFRPWKRLLTNFRYQDSKLQEILRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFLD KIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPF WPWKGLEIISRRTQRRLRRIKESWGLQDLVNDFGNLQLGPPMS
1003	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC46	MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPR EEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKRL LTNFRYQDSKLQEILRPCYISVPSSSSSTLSNICLTKGLPETRFWVEGRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQ APLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILV DVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGLQDLVNDFGNLQLGPPMS
1004	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC47	MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPR EEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKRL LTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVEGRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQ APLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILV DVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGLQDLVNDFGNLQLGPPMS
1005	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC48	MQPQRLGPRAGMGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWF HDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDN GGRRFRPWKRLLTNFRYQDSKLQEILRPCYISVPSSSSSTLSNICLTKGLPETRFWVEGRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPY LCYQLEQFNGQAPLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKG LCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKEVRTTLLQGPAS
1006	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC49	MQPQRLGPRAGMGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWF HDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDN GGRRFRPWKRLLTNFRYQDSKLQEILRPCYISVPSSSSSTLSNICLTKGLPETRFWVEGRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPY LCYQLEQFNGQAPLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKG LCSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGLQDLVNDFGNLQLGPPMS

1007	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC50	MKETDQMQSLEGSGAERSVGTQTGSMTGQIPRLSKVNLFTLLSLWMELFPGVEAQGQKSQKTEEESRGPLGDNEELTRVSTEKKQVKKTGL VVVKNMKIIGLHCSSEDLHTGQIALIKHGSRLKNCDLYFSRKPCSACLKMIVNAGVNRISYWPSDPEISLLTEASSSEDAKLDAKAAERLK SNSRAHVCVLLQPLVCYMVQFVEETSYKCDFIQKTAKALPGADTDFYSECKQERIKEYEMLFLVSNEERHKQILMTIGLESLCEDPYFSNL RQNMKDLILLLATVASSVPNLKHFGFYCSSPEQINEIHNQSLPQEVARHCMVQARLLAYRTEDHKTGVGAVIWAEAKSRSCDGTGAMYFIG CGYNAFPVGSEYADFPHMDDKHKDREIRKFRYIIHAEQNALTFRCQDIKPEERSMIFVTKCPCDECVPLIKGAGIKQIYAGDVDVGKKKAD ISYMKFGELEGVRKFTWQLNPSEAYSLDPNEPERRENGVLRRRSAKDEQRSSKRPRLETRSAGSATTACF
1008	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC51	MSNNALQTIINARLPGEEGLWQIHLQDGKISAIDAQSGVMPITENSLDAEQGLVIPPFVEPHIHLDTTQTAGQPNWNQSGTLFEGIERWAE RKALLTHDDVKQRAWQTLKWQIANGIQHVRTHVDVSDATLTALKAMLEVKQEVAPWIDLQIVAFPQEGILSYPNGEALLEEALRLGADVVG AIPHFEFTREYGVESLHKTFALAQKYDRLIDVHCDEIDDEQSRFVETVAALAHHEGMGARVTASHTTAMHSYNGAYTSRLFRLLKMSGINF VANPLVNIHLQGRFDTYPKRRGITRVKEMLESGINVCFGHDDVFDPWYPLGTANMLQVLHMGLHVCQLMGYGQINDGLNLITHHSARTLNL QDYGIAAGNSANLIILPAENGFDALRRQVPVRYSVRGGKVIASTQPAQTTVYLEQPEAIDYKR
1009	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC52	MALLTAKTFSLQFNNKRRVNKPYYPRKALLCYQLTPQNGSTPTRGHLKNKKKDHAEIRFINKIKSMGLDETQCYQVTCYLTWSPCPSCAGE LVDFIKAHRHLNLRIFASRLYYHWRPNYQEGLLLLCGSQVPVEVMGLPEFTDCWENFVDHKEPPSFNPSEKLEELDKNSQAIKRRLERIKS RSVDVLENGLRSLQLGPVTPSSSIRNSR
1010	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC53	MEKDINLKIFKGNLIFTKTSDKFTIMKDSYIVVIDGKIASVSSNLPDKYKGNPIIDFRNNIIIPGMNDLHAHASQYKNLGIGMDKELLPWL NNYTFPEEAKFLNVDYAKKTYGRLIKDLIKNGTTRVALFATLHKDSTIELFNMLIKSGIGAYVGKVNMDYNCPDYLTENYITSLNDTEEII LKYKDKSNIVKPIITPRFVPSCSNELMDGLGKLSYKYRLPVQSHLSENLDEIAVVKSLHKKSNFYGEVYDKFGLFGNTPTLMAHCIHSSKE EINLIKRNNVTIVHCPTSNFNLGSGMMPVRKYLNLGINVVLGSDISAGHTCSLFKVIAYAIQNSKIKWQESGKKDMFLSTSEAFYMATKKG GSFFGKVGSFEEGYDFDALVINDSNLYPEDYDLTERLERFIYLGDDRNIMKRYVCGNEIFGPKF
1011	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC54	MKIINARLRRQEALFTLDLQDGIIHRITAQAAMQTADAGAIDAQGRLAIPPFVEPHIHLDATLTAGEPEWNRSGTLFEGITRWSQRKASIT PEDTRQRALKTIGMLRDFGVQHVRTHVDVTDPSLAALQALLAVKQEAADLIDLQIVAFPQEGIESYPNGRELMTRAIEMGADVVGGIPHYE NTRDKGVSSVMFLMDLAQRYGRLVDVHCDEIDDPQSRFLEVLAEEARVRGMGAQVTASHTCAMGSYDNAYCSKLFRLLKASGINFISCPTE SIHLQGRFDSWPKRRGVTRVAELDRAGINVCFAQDSIQDPWYPLGNGNILRILDAGLHICHMLGYDDLQRCLDFVTDNSARALCLGDNYGL AEGRPANLLILDAENDYEAVRRQARVLTSIRHGKVILQREVEHIRYPA
1012	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC55	MGRKLDPTKEKRGPGRKARKQKGAETELVRFLPAVSDENSKRLSSRARKRAAKRRLGSVEAPKTNKSPEAKPLPGKLPKGISAGAVQTAGK KGPQSLFNAPRGKKRPAPGSDEEEEEEDSEEDGMVNHGDLWGSEDDADTVDDYGADSNSEDEEEGEALLPIERAARKQKAREAAAGIQWSE EETEDEEEEKEVTPESGPPKVEEADGGLQINVDEEPFVLPPAGEMEQDAQAPDLQRVHKRIQDIVGILRDFGAQREEGRSRSEYLNRLKKD LAIYYSYGDFLLGKLMDLFPLSELVEFLEANEVPRPVTLRTNTLKTRRRDLAQALINRGVNLDPLGKWSKTGLVVYDSSVPIGATPEYLAG HYMLQGASSMLPVMALAPQEHERILDMCCAPGGKTSYMAQLMKNTGVILANDANAERLKSVVGNLHRLGVTNTIISHYDGRQFPKVVGGFD RVLLDAPCSGTGVISKDPAVKTNKDEKDILRCAHLQKELLLSAIDSVNATSKTGGYLVYCTCSITVEENEWVVDYALKKRNVRLVPTGLDF GQEGFTRFRERRFHPSLRSTRRFYPHTHNMDGFFIAKFKKFSNSIPQSQTGNSETATPTNVDLPQVIPKSENSSQPAKKAKGAAKTKQQLQ KQQHPKKASFQKLNGISKGADSELSTVPSVTKTQASSSFQDSSQPAGKAEGIREPKVTGKLKQRSPKLQSSKKVAFLRQNAPPKGTDTQTP AVLSPSKTQATLKPKDHHQPLGRAKGVEKQQLPEQPFEKAAFQKQNDTPKGPQPPTVSPIRSSRPPPAKRKKSQSRGNSQLLLS

1013	Amino acid sequence of exemplary cytidine deaminase; see TABLE 58, BC56	MGRRSRGRRLQQQQRPEDAEDGAEGGGKRGEAGWEGGYPEIVKENKLFEHYYQELKIVPEGEWGQFMDALREPLPATLRITGYKSHAKEIL HCLKNKYFKELEDLEVDGQKVEVPQPLSWYPEELAWHTNLSRKILRKSPHLEKFHQFLVSETESGNISRQEAVSMIPPLLLNVRPHHKILD MCAAPGSKTTQLIEMLHADMNVPFPEGFVIANDVDNKRCYLLVHQAKRLSSPCIMVVNHDASSIPRLQIDVDGRKEILFYDRILCDVPCSG DGTMRKNIDVWKKWTTLNSLQLHGLQLRIATRGAEQLAEGGRMVYSTCSLNPIEDEAVIASLLEKSEGALELADVSNELPGLKWMPGITQW KVMTKDGQWFTDWDAVPHSRHTQIRPTMFPPKDPEKLQAMHLERCLRILPHHQNTGGFFVAVLVKKSSMPWNKRQPKLQGKSAETRESTQL SPADLTEGKPTDPSKLESPSFTGTGDTEIAHATEDLENNGSKKDGVCGPPPSKKMKLFGFKEDPFVFIPEDDPLFPPIEKFYALDPSFPRM NLLTRTTEGKKRQLYMVSKELRNVLLNNSEKMKVINTGIKVWCRNNSGEEFDCAFRLAQEGIYTLYPFINSRIITVSMEDVKILLTQENPF FRKLSSETYSQAKDLAKGSIVLKYEPDSANPDALQCPIVLCGWRGKASIRTFVPKNERLHYLRMMGLEVLGEKKKEGVILTNESAASTGQP DNDVTEGQRAGEPNSPDAEEANSPDVTAGCDPAGVHPPR
1014	Amino acid sequence of exemplary APOBEC family deaminase	MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPR EEFKITWYMSWSPCFECAEQIVRFLATHHNLSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKRL LTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVEGRRMDPLSEEEFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQ APLKGCLLSEKGKQHAEILFLDKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILV DVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGLQDLVNDFGNLQLGPPMS
1015	Amino acid sequence of exemplary APOBEC family deaminase	MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLRYAIDRKDTFLCYEVTRKDCDSPVSLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPR EEFKITWYMSWSPCFECAEQVLRFLATHHNLSLDIFSSRLYNIRDPENQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRPWKKL LTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVERRRVHLLSEEEFYSQFYNQRVKHLCYYHGVKPYLCYQLEQFNGQ APLKGCLLSEKGKQHAEILFLDKIRSMELSQVIITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQSGILV DVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLHRIKESWGLQDLVNDFGNLQLGPPMS
1016	Amino acid sequence of exemplary APOBEC family deaminase	MVEPMDPRTFVSNFNNRPILSGLNTVWLCCEVKTKDPSGPPLDAKIFQGKVYSKAKYHPEMRFLRWFHKWRQLHHDQEYKVTWYVSWSPCT RCANSVATFLAKDPKVTLTIFVARLYYFWKPDYQQALRILCQKRGGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPRNNLPKHYTLLQATL GELLRHLMDPGTFTSNFNNKPWVSGQHETYLCYKVERLHNDTWVPLNQHRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQYRVTC FTSWSPCFSCAQEMAKFISNNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYSEFEYCWDTFVDRQGRPFQPWDGLDEHSQALS GRLRAI
1017	Amino acid sequence of exemplary APOBEC family deaminase	MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPSRPPLDAKIFRGQVYSKLKYHPEMRFFHWFSKWRKLHRDQEYEVTWY ISWSPCTKCTRDVATFLAEDPKVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYY ILLHIMLGEILRHSMDPPTFTSNFNNELWVRGRHETYLCYEVERLHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDL HQDYRVTCFTSWSPCFSCAQEMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLAKAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGL EEHSQALSGRLRAILQNQGN
1018	Amino acid sequence of exemplary APOBEC family deaminase	MNPQIRNMVEQMEPDIFVYYFNNRPILSGRNTVWLCYEVKTKDPSGPPLDANIFQGKLYPEAKDHPEMKFLHWFRKWRQLHRDQEYEVTWY VSWSPCTRCANSVATFLAEDPKVTLTIFVARLYYFWKPDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDGQGKPFKPRKNLPKHY TLLHATLGELLRHVMDPGTFTSNFNNKPWVSGQRETYLCYKVERSHNDTWVLLNQHRGFLRNQAPDRHGFPKGRHAELCFLDLIPFWKLDD QQYRVTCFTSWSPCFSCAQKMAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLHRDGAKIAVMNYSEFEYCWDTFVDRQGRPFQPWDGLD EHSQALSGRLRAI
1019	Amino acid sequence of exemplary APOBEC family deaminase	MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQVYSELKYHPEMRFFHWFSKWRKLHRDQEYEVTWY ISWSPCTKCTRDMATFLAEDPKVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYY ILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDL DQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGL DEHSQDLSGRLRAILQNQEN

1020	Amino acid sequence of exemplary APOBEC family deaminase	MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPRLDAKIFRGQVYSQPEHHAEMCFLSWFCGNQLPAYKCFQITWFV SWTPCPDCVAKLAEFLAEHPNVTLTISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFDDNYAFLHR TLKEILRNPMEAMYPHIFYFHFKNLRKAYGRNESWLCFTMEVVKHHSPVSWKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTW YTSWSPCPECAGEVAEFLARHSNVNLTIFTARLYYFWDTDYQEGLRSLSQEGASVEIMGYKDFKYCWENFVYNDDEPFKPWKGLKYNFLFL DSKLQEILE
1021	Amino acid sequence of exemplary APOBEC family deaminase	MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQVYFKPQYHAEMCFLSWFCGNQLPAYKCFQITWF VSWTPCPDCVAKLAEFLSEHPNVTLTISAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQQFMPWYKFDENYAFLH RTLKEILRYLMDPDTFTFNFNNDPLVLRRRQTYLCYEVERLDNGTWVLMDQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIY RVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVYRQGCPFQPWDGLEE HSQALSGRLRAILQNQGN
1022	Amino acid sequence of exemplary APOBEC family deaminase	MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGPVLPKRQSNHRQEVYFRFENHAEMCFLSWFCGNR LPANRRFQITWFVSWNPCLPCVVKVTKFLAEHPNVTLTISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVCNEGQPFMP WYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFKNLLKACGRNESWLCFTMEVTKHHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCD DILSPNTNYEVTWYTSWSPCPECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDFVSCWKNFVYSDDEPF KPWKGLQTNFRLLKRRLREILQ
1023	Amino acid sequence of exemplary APOBEC family deaminase	MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWS PCGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYVL ELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK
1076	G018021 guide RNA targeting CIITA gene	mU*mC*mC*UACCUGUCAGAGCCCCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1077	G018022 guide RNA targeting CIITA gene	mG*mA*mG*CCCCAAGGUAAAAAGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1078	G018023 guide RNA targeting CIITA gene	mA*mG*mC*CCCAAGGUAAAAAGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1079	G018024 guide RNA targeting CIITA gene	mU*mU*mU*CCCGGCCUUUUUACCUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1080	G018025 guide RNA targeting CIITA gene	mC*mU*mU*UCCCGGCCUUUUUACCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1081	G018026 guide RNA targeting CIITA gene	mA*mC*mA*CUGUGAGCUGCCUGGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

1082	G018027 guide RNA targeting CIITA gene	mC*mA*mC*ACUGUGAGCUGCCUGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1083	G018028 guide RNA targeting CIITA gene	mG*mU*mG*GCACACUGUGAGCUGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1084	G018029 guide RNA targeting CIITA gene	mG*mA*mG*AUUGAGCUCUACUCAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1085	G018030 guide RNA targeting CIITA gene	mA*mG*mA*UUGAGCUCUACUCAGGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1086	G018031 guide RNA targeting CIITA gene	mC*mA*mG*GUGGGCCCUCCUCCCCUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1087	G018032 guide RNA targeting CIITA gene	mC*mG*mG*AAGAGACCAGAGGGAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1088	G018033 guide RNA targeting CIITA gene	mU*mA*mC*CGGAAGAGACCAGAGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1089	G018034 guide RNA targeting CIITA gene	mA*mG*mG*GAGGCUUAUGCCAAUAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1090	G018035 guide RNA targeting CIITA gene	mG*mG*mC*UUAUGCCAAUAUCGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1091	G018036 guide RNA targeting CIITA gene	mU*mG*mA*CUGCGCUUUUCCUUGUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1092	G018037 guide RNA targeting CIITA gene	mG*mA*mC*UGCGCUUUUCCUUGUCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1093	G018038 guide RNA targeting CIITA gene	mG*mC*mU*UUUCCUUGUCUGGGCAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1094	G018039 guide RNA targeting CIITA gene	mC*mC*mA*GUUCCGCUGCCCAGACAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

1095	G018040 guide RNA targeting CIITA gene	mG*mA*mC*CUGAAGCACUGGAAGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1096	G018041 guide RNA targeting CIITA gene	mG*mC*mA*CUGGAAGCCAGGUGUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1097	G018042 guide RNA targeting CIITA gene	mC*mA*mC*UGGAAGCCAGGUGUGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1098	G018043 guide RNA targeting CIITA gene	mG*mG*mA*AGCCAGGUGUGCAGGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1099	G018044 guide RNA targeting CIITA gene	mA*mG*mC*CAGGUGUGCAGGGCAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1100	G018045 guide RNA targeting CIITA gene	mG*mC*mC*CACCUGCCCUGCACACCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1101	G018046 guide RNA targeting CIITA gene	mG*mG*mG*CUCAGCUGUGAGGAAGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1102	G018047 guide RNA targeting CIITA gene	mG*mG*mG*GCUCAGCUGUGAGGAAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1103	G018048 guide RNA targeting CIITA gene	mG*mA*mC*CAGAUUCCCAGUAUGUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1104	G018049 guide RNA targeting CIITA gene	mC*mC*mA*GAUUCCCAGUAUGUUAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1105	G018050 guide RNA targeting CIITA gene	mC*mA*mG*AUUCCCAGUAUGUUAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1106	G018051 guide RNA targeting CIITA gene	mU*mC*mC*CAGUAUGUUAGGGGGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1107	G018052 guide RNA targeting CIITA gene	mU*mC*mC*AAGCCCCCUAACAUACUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

1108	G018053 guide RNA targeting CIITA gene	mC*mU*mC*CAAGCCCCCUAACAUACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1109	G018054 guide RNA targeting CIITA gene	mA*mA*mA*GGCACUGCAAGAGACAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1110	G018055 guide RNA targeting CIITA gene	mU*mC*mC*AGUAUAUUCAUCUACCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1111	G018056 guide RNA targeting CIITA gene	mU*mU*mC*AUCUACCAUGGUGAGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1112	G018057 guide RNA targeting CIITA gene	mU*mC*mA*UCUACCAUGGUGAGUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1113	G018058 guide RNA targeting CIITA gene	mC*mA*mU*CUACCAUGGUGAGUGCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1114	G018059 guide RNA targeting CIITA gene	mA*mC*mC*AUGGUGAGUGCGGGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1115	G018060 guide RNA targeting CIITA gene	mG*mC*mC*AGGCCCCGCACUCACCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1116	G018061 guide RNA targeting CIITA gene	mC*mC*mA*CUCUCCACCCCCAAUGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1117	G018062 guide RNA targeting CIITA gene	mC*mU*mC*CACCCCCAAUGUAGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1118	G018063 guide RNA targeting CIITA gene	mC*mA*mC*CUCACCUACAUUGGGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1119	G018064 guide RNA targeting CIITA gene	mG*mG*mG*CACCUCACCUACAUUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1120	G018065 guide RNA targeting CIITA gene	mU*mG*mG*GGCACCUCACCUACAUUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

1121	G018066 guide RNA targeting CIITA gene	mC*mU*mG*GGGCACCUCACCUACAUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1122	G018067 guide RNA targeting CIITA gene	mA*mC*mC*UCCCGAGCAAACAUGACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1123	G018068 guide RNA targeting CIITA gene	mC*mC*mG*AGCAAACAUGACAGGUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1124	G018069 guide RNA targeting CIITA gene	mG*mU*mA*CCUCUCACAGGCCCUAAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1125	G018070 guide RNA targeting CIITA gene	mA*mG*mU*ACCUCUCACAGGCCCUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1126	G018071 guide RNA targeting CIITA gene	mG*mA*mC*GUCUUGUGCUCUGGAGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1127	G018072 guide RNA targeting CIITA gene	mA*mA*mC*AAGCUUCCAAAAUGGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1128	G018073 guide RNA targeting CIITA gene	mG*mA*mG*AUCCCGCAUCACUCACCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1129	G018074 guide RNA targeting CIITA gene	mG*mC*mC*CCUGGCCUUUGCAGAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1130	G018079 guide RNA targeting CIITA gene	mU*mC*mA*GCUGUGUCACCCGUUUCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1131	G018080 guide RNA targeting CIITA gene	mG*mC*mU*GUGUCACCCGUUUCAGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1132	G018083 guide RNA targeting CIITA gene	mC*mA*mC*CCGUUUCAGGUGGGGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1133	G018086 guide RNA targeting CIITA gene	mG*mU*mC*UGAGGCCCUCCCUCCACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

1134	G018087 guide RNA targeting CIITA gene	mU*mC*mU*GAGGCCCUCCCUCCACAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1135	G018088 guide RNA targeting CIITA gene	mC*mA*mA*GGCAGCCCUGUGGAGGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1136	G018089 guide RNA targeting CIITA gene	mG*mC*mU*CAAGGCAGCCCUGUGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1137	G018090 guide RNA targeting CIITA gene	mC*mG*mC*UCAAGGCAGCCCUGUGGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1138	G018092 guide RNA targeting CIITA gene	mU*mA*mA*CAUUGCCUGUUCUCUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1139	G018093 guide RNA targeting CIITA gene	mC*mU*mU*CUCGUCCUGGAGAGAACGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1140	G018094 guide RNA targeting CIITA gene	mU*mU*mC*CGAGGAACUUCUCGUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1141	G018095 guide RNA targeting CIITA gene	mA*mC*mC*CUUGCUCUUUGCCUCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1142	G018096 guide RNA targeting CIITA gene	mU*mU*mG*CUCUUUGCCUCCUAGGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1143	G018097 guide RNA targeting CIITA gene	mU*mG*mC*UCUUUGCCUCCUAGGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1144	G018098 guide RNA targeting CIITA gene	mC*mC*mU*GAGACAGGGCCCAGCCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1145	G018099 guide RNA targeting CIITA gene	mA*mU*mC*UGAUUCCACCUGCAGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1146	G018101 guide RNA targeting CIITA gene	mA*mC*mU*CAGCGCAUCCAGGCUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

1147	G018102 guide RNA targeting CIITA gene	mA*mA*mA*CCCUCAAGUGAGUGAGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1148	G018103 guide RNA targeting CIITA gene	mA*mA*mC*CCUCAAGUGAGUGAGCUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1149	G018104 guide RNA targeting CIITA gene	mG*mG*mC*CCAGCUCACUCACUUGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1150	G018105 guide RNA targeting CIITA gene	mA*mG*mG*CCCAGCUCACUCACUUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1151	G018106 guide RNA targeting CIITA gene	mC*mA*mG*ACUGCGGGGACACAGUGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1152	G018107 guide RNA targeting CIITA gene	mC*mU*mG*CAUCCCUGCUCAGGCUAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1153	G018108 guide RNA targeting CIITA gene	mC*mC*mU*GCUCAGGCUAAGGUGAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1154	G018109 guide RNA targeting CIITA gene	mC*mU*mG*CUCAGGCUAAGGUGAGUGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1155	G018110 guide RNA targeting CIITA gene	mC*mA*mG*CACCUGACCGGUAUCCGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1156	G018111 guide RNA targeting CIITA gene	mG*mU*mA*CAAGCUGUCGGAAACAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1157	G018112 guide RNA targeting CIITA gene	mG*mG*mA*GACGCUGGCGUAAGUCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1158	G018113 guide RNA targeting CIITA gene	mG*mG*mC*GUAAGUCCAGGCAACCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1159	G018114 guide RNA targeting CIITA gene	mC*mU*mC*CACCCACCAGGGUUGCCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

1160	G018115 guide RNA targeting CIITA gene	mU*mG*mA*GUCCCAUCCCCCCUUGCGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1161	G018116 guide RNA targeting CIITA gene	mC*mC*mA*CAUCCUGCAAGGGGGGAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1162	G018119 guide RNA targeting CIITA gene	mG*mG*mG*CGUCCACAUCCUGCAAGGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU
1163	G018121 guide RNA targeting CIITA gene	mG*mU*mG*GGCGUCCACAUCCUGCAGUUUUAGAmGmCmUmAmGmAmAmAmUmAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCAmAmCm UmUmGmAmAmAmAmAmGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU*mU*mU*mU

(In each of the sequences in the Table above or described herein, a modified sequence can be unmodified or modified in an alternative way)

EXAMPLES

[00489] The following examples are provided to illustrate certain disclosed embodiments and are not to be construed as limiting the scope of this disclosure in any way.

[00490] As used in the Examples below, the term "editor" refers to an agent comprising a polypeptide that is capable of deaminating a base within a DNA molecule and it is a base editor. The editor may be capable of deaminating a cytidine (C) in DNA. The editor may include an RNA-guided nickase (e.g., Cas9 nickase) fused to a cytidine deaminase (e.g., an APOBEC3A deaminase (A3A)) by an optional linker. In some cases, the editor includes a UGI. In some embodiments, the editor lacks a UGI.

[00491] An exemplary editor used in the Examples below is BC22n (SEQ ID NO:3) which consists of a *H. sapiens* APOBEC3A fused to *S. pyogenes*-D10A SpyCas9 nickase by an XTEN linker, and mRNA encoding BC22n. An mRNA encoding BC22n (SEQ ID NO: 1) is also used.

Example 1. General Methods

1.1. Preparation of lipid nanoparticles

[00492] In general, the lipid components were dissolved in 100% ethanol at various molar ratios. The RNA cargos (*e.g.*, Cas9 mRNA and sgRNA) were dissolved in 25 mM citrate buffer, 100 mM NaCl, pH 5.0, resulting in a concentration of RNA cargo of approximately 0.45 mg/mL.

[00493] Unless otherwise specified, the lipid nucleic acid assemblies contained ionizable Lipid A ((9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-((((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-((((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl (9Z,12Z)-octadeca-9,12-dienoate), cholesterol, DSPC, and PEG2k-DMG in a 50:38:9:3 molar ratio, respectively. The lipid nucleic acid assemblies were formulated with a lipid amine to RNA phosphate (N:P) molar ratio of about 6, and a ratio of gRNA to mRNA of 1:1 by weight, unless otherwise specified.

[00494] LNPs were prepared using a cross-flow technique utilizing impinging jet mixing of the lipid in ethanol with two volumes of RNA solutions and one volume of water. The lipids in ethanol were mixed through a mixing cross with the two volumes of RNA solution. A fourth stream of water was mixed with the outlet stream of the cross through an inline tee (*See* WO2016010840 Fig. 2.). The LNPs were held for 1 hour at room

temperature, and further diluted with water (approximately 1:1 v/v). LNPs were concentrated using tangential flow filtration on a flat sheet cartridge (Sartorius, 100kD MWCO) and buffer exchanged using PD-10 desalting columns (GE) into 50 mM Tris, 45 mM NaCl, 5% (w/v) sucrose, pH 7.5 (TSS). Alternatively, the LNP's were optionally concentrated using 100 kDa Amicon spin filter and buffer exchanged using PD-10 desalting columns (GE) into TSS. The resulting mixture was then filtered using a 0.2 μm sterile filter. The final LNP was stored at 4°C or -80°C until further use.

1.2. In vitro transcription ("IVT") of mRNA

[00495] Capped and polyadenylated mRNA containing N1-methyl pseudo-U was generated by in vitro transcription using a linearized plasmid DNA template and T7 RNA polymerase. Plasmid DNA containing a T7 promoter, a sequence for transcription, and a polyadenylation region was linearized by incubating at 37°C for 2 hours with XbaI with the following conditions: 200 ng/μL plasmid, 2 U/μL XbaI (NEB), and 1x reaction buffer. The XbaI was inactivated by heating the reaction at 65°C for 20 min. The linearized plasmid was purified from enzyme and buffer salts. The IVT reaction to generate modified mRNA was performed by incubating at 37°C for 1.5-4 hours in the following conditions: 50 ng/μL linearized plasmid; 2-5 mM each of GTP, ATP, CTP, and N1-methyl pseudo-UTP (Trilink); 10-25 mM ARCA (Trilink); 5 U/μL T7 RNA polymerase (NEB); 1 U/μL Murine RNase inhibitor (NEB); 0.004 U/μL Inorganic E. coli pyrophosphatase (NEB); and 1x reaction buffer. TURBO DNase (ThermoFisher) was added to a final concentration of 0.01 U/μL, and the reaction was incubated for an additional 30 minutes to remove the DNA template. The mRNA was purified using a MegaClear Transcription Clean-up kit (ThermoFisher) or a RNeasy Maxi kit (Qiagen) per the manufacturers' protocols. Alternatively, the mRNA was purified through a precipitation protocol, which in some cases was followed by HPLC-based purification. Briefly, after the DNase digestion, mRNA is purified using LiCl precipitation, ammonium acetate precipitation and sodium acetate precipitation. For HPLC purified mRNA, after the LiCl precipitation and reconstitution, the mRNA was purified by RP-IP HPLC (see, *e.g.*, Kariko, et al. Nucleic Acids Research, 2011, Vol. 39, No. 21 e142). The fractions chosen for pooling were combined and desalted by sodium acetate/ethanol precipitation as described above. In a further alternative method, mRNA was purified with a LiCl precipitation method followed by further purification by tangential flow filtration. RNA concentrations were determined by measuring the light absorbance at 260 nm (Nanodrop), and transcripts were analyzed by capillary electrophoresis by Bioanlayzer (Agilent).

[00496] Nme2Cas9 mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID NOs: 360 (see sequences in **Table 5C**). *Streptococcus pyogenes* ("Spy") Cas9 mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID NOs: 8, 11, or 23 (see sequences in **Table 5C**). BC22n mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID NOs: 2 or 5. BC22 mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID NO: 20. BC22 with 2x UGI mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID No: 29. UGI mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID NOs: 26 or 35. BE3 mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID NO: 14 or 17. BE4Max mRNA was generated from plasmid DNA encoding an open reading frame according to SEQ ID NO: 32. When the sequences cited in this paragraph are referred to below with respect to RNAs, it is understood that Ts should be replaced with Us (which were N1-methyl pseudouridines as described above). Messenger RNAs used in the Examples include a 5' cap and a 3' polyadenylation region, *e.g.*, up to 100 nts.

1.3. Next-generation sequencing ("NGS") and analysis for on-target editing efficiency

[00497] Genomic DNA was extracted using QuickExtract™ DNA Extraction Solution (Lucigen, Cat. QE09050) according to the manufacturer's protocol. To quantitatively determine the efficiency of editing at the target location in the genome, deep sequencing was utilized to identify the presence of insertions and deletions introduced by gene editing. PCR primers were designed around the target site within the gene of interest (e.g., TRAC) and the genomic area of interest was amplified. Primer sequence design was done as is standard in the field.

[00498] Additional PCR was performed according to the manufacturer's protocols (Illumina) to add chemistry for sequencing. The amplicons were sequenced on an Illumina MiSeq instrument. The reads were aligned to the human reference genome (e.g., hg38) after eliminating those having low quality scores. Reads that overlapped the target region of interest were re-aligned to the local genome sequence to improve the alignment. Then the number of wild type reads versus the number of reads which contain C-to-T mutations, C-to-A/G mutations or indels was calculated. Insertions and deletions were scored in a 20 bp region centered on the predicted Cas9 cleavage site. Indel percentage is defined as the total number of sequencing reads with one or more base inserted or deleted within the 20

bp scoring region divided by the total number of sequencing reads, including wild type. C-to-T mutations or C-to-A/G mutations were scored in a 40 bp region including 10 bp upstream and 10 bp downstream of the 20 bp sgRNA target sequence. The C-to-T editing percentage is defined as the total number of sequencing reads with either one or more C-to-T mutations within the 40 bp region divided by the total number of sequencing reads, including wild type. The percentage of C-to-A/G mutations are calculated similarly.

Example 1A - C-to-T conversion

[00499] APOBEC3A deaminase-Cas9D10A editor was evaluated for efficiency of C-to-T conversion activity and the span of the C-to-T conversion window. C-to-T conversion activity and window results were compared against construct BC27 encoding BE3 (Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016;533(7603):420-424.) Constructs were each evaluated in triplicate against 5 sgRNA in a single experiment.

[00500] Plasmid A, using a pUC19 backbone (GenBank® Accession Number U47119), expresses an S. pyogenes single-guide RNA (sgRNA) from a U6 promoter. Plasmid B, called pCI, expresses base editor constructs from a CMV promoter which are comprised of the candidate deaminase fused to S. pyogenes-D10A-Cas9 by an XTEN linker which is subsequently fused to one copy of UGI and one copy of SV40 NLS. U-2OS cells growing in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) in 96-well plates were transfected using Mirus *Trans*IT-X2® with 100 ng each of plasmid A and plasmid B. Cells were washed and resuspended in fresh media 24 hours after initial transfection. After an additional 48 hours, the media was removed and the cells were lysed with QuickExtract™ DNA Extraction Solution (Lucigen, Cat. QE09050).

[00501] Construct BC22 encoding the *H. sapiens APOBEC3A* deaminase (Uniprot ID P31941) in the fusion had significantly greater C-to-T conversion activity than BC27 encoding the rat APOBEC1 deaminase (Uniprot ID P38483) with more than one guide (sg000296: P value 0.0303; sg001373, P value 0.0263). The average activity of BC22 was comparable to that of BC27 across all guides (Table 6, Figs. 1A-1E).

[00502] Across all 5 sgRNAs tested, BC22 was significantly more likely to convert all target cytosines to thymidines than BC27 (Table 7, Figs. 2A-2E, P value <0.005). A target cytosine is any cytosine within positions 1-10 of the protospacer target sequence (underlined) or in first position 5' of the target sequence The guide sequence for sg000296 is CCUUCCGAAAGAGGCCCCCC (SEQ ID NO: 156), and the locus has 4 target cytosines.

The target guide sequence for sg001373 is UCCCUGGCUGAGGAUCCCCA (SEQ ID NO: 157), and the locus has 5 target cytosines, including the first position 5' of the target sequence. The target guide sequence for sg001400 is ACUCACGAUGAAAUCCUGGA (SEQ ID NO: 158), and the locus has 4 target cytosines including the first position 5' of the target sequence. The target guide sequence for sg003018 is GAGCCCCCCACUGUGGUGAC (SEQ ID NO: 160), and the locus has 6 target cytosines. The guide sequence for sg005883 is CCCCCCGCCGUGUUUGUGGG (SEQ ID NO: 159), and the locus has 8 target cytosines.

[00503] BC22 converted a significantly larger proportion and positional range of target cytosine than BC27 (Figs. 3A-3E). converted a significantly larger proportion and positional range of target cytosines than BC27 (Figs. 3A-3E). This wider C-to-T conversion window held true even for guides that BC22 had less total C-to-T conversion activity on such as sg001400, sg003018 (Table 7, Figure Figs. 3C-3E).

[00504] Table 6. Percent of total reads containing at least 1 cytosine to thymidine conversion. (n=3)

	sg000296		**sg0001373**		**sg001400**		**sg003018**		**sg005883**	
Construct	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
1XNLS Cas9	0.3	0.2	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.1
3XNLS Cas9	0.1	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0
BC22	16.0	0.9	20.3	2.6	24.2	1.1	12.6	1.7	46.0	7.7
BC27	14.2	0.3	13.0	2.6	30.6	1.4	19.6	1.9	40.8	5.2
GFP	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.1

[00505] Table 7. Percent of C-to-T edited reads in which all target cytosines have been converted to thymidines (sg000296, sg001373, sg001400, sg003018). For sg005883, percent of edited reads in which at least 6 of 8 target cytosines have been converted to thymidines. (n/d = not done)

	sg000296		**sg001373**		**sg001400**		**sg003018**		**sg005883**	
Construct	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
1XNLS Cas9	0	0	n/d	n/d	0	0	0	0	9.4	9.5
3XNLS Cas9	0	0	0	0	0	0	0	0	7.0	3.9
BC22	45.1	2.1	34.6	5.5	11.1	0.9	80.2	1.4	61.8	10.1
BC27	0.6	0	3.3	0.3	0.1	0.2	43.2	6.7	15.2	4.8
GFP	0	0	n/d	n/d	0	0	0	0	0	0

Example 2 - Effects of base excision repair genes on editing

[00506] The effects of additional UGI gene expression *in trans* on C-to-T conversion activity relative to unwanted base excision repair activity were investigated in various cell lines.

[00507] U-2OS and HuH-7 cells grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FBS in 96-well plates were co-transfected using Mirus *Trans*IT-X2® with 100 ng of BC22 or BC27, 100 ng of another CMV-driven overexpression plasmid (pcDNA3.1) or a control plasmid (pMAX) and 6.25 pmol of single-guide G000297. The ORFs tested in tandem with BC22 and BC27 were green fluorescent protein (pMAX-GFP, negative control), Uracil DNA glycosylase (pCDNA3.1-UNG), Single-Strand-Selective Monofunctional Uracil-DNA Glycosylase 1 (pcDNA3.1-SMUG1) and Uracil Glycosylase Inhibitor (pCDNA3.1-UGI). UNG and SMUG1 are base excision repair proteins that remove uracil from DNA. UGI is the small protein that binds and inhibits UDG. It has been reported that when UGI is fused to the constructs it increases the rate of C-to-T mutations relative to the other outcomes of C-to-A/G mutations and indels (Liu et al, Nature 2016). To determine whether the addition of UNG transcript knockdown further enriched for C-to-T-only edits, an additional transfection condition included either BC22 or BC27, pcDNA3.1-UGI, G000297 and a pool of siRNA targeting UNG (Dharmacon, # M-011795-00). Three days after transfection, media was removed and the cells were lysed with QuickExtract™ DNA Extraction Solution (Lucigen, Cat. QE09050).

[00508] In HuH-7 cells, overexpressing base excision repair proteins UNG or SMUG1 with either BC22 or BC27 led to a 1.3-1.7-fold significant increase in the proportion of C-to-A/G mutations and indels relative to C-to-T only mutations when compared to overexpression of the negative control GFP (Table 11, Figure 5B). Conversely, blocking UNG with UGI overexpression led to a 4.5-10.8-fold decrease in reads containing C-to-A/G mutations and indels with a significant 1.56-2.2-fold increase in C-to-T-only mutations (Table 11, Figure 5B). Additional UGI overexpression led to, on average, 94% of total edits being C-to-T-only edits whether using BC22 or BC27 (Figure 4B). Similar trends were seen with U-2OS (Table 8, Figure Figures 4A, 5A) with the one outstanding difference that SMUG1 and UNG overexpression did not lead to significant changes in relative proportions of editing outcomes. In both cell lines the addition of siRNAs targeting UNG gene to the UGI overexpression plasmid did not increase the proportion of C-to-T-only edits over UGI overexpression alone (Table 8). Finally, in both cell lines BC22 total C-to-T conversion

activity was higher than BC27 (Figures 4A, 4B), suggesting BC22 has higher innate C-to-T conversion activity than BC27.

[00509] Table 8 – Editing as a % of total reads in U-2OS cells (n=3)

U-2OS		C-to-T %		C-to-A/G %		Indel %	
Construct	**Overexpression/ siRNA**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
BC27	GFP+	1.57	0.99	0.86	0.14	1.18	0.16
	SMUG1+	2.13	0.45	2.4	0.39	2.59	0.09
	UDG+	0.3	0.26	0.36	0.22	1.05	0.04
	UGI+	4.77	0.75	0.29	0.08	0.8	0.34
	siUDG, UGI+	4.9	0.75	0.22	0.05	0.76	0.03
BC22	GFP+	10.87	1.58	1.78	0.61	1.64	0.8
	SMUG1+	6.17	0.4	6.21	1.42	4.85	0.9
	UDG+	1.57	0.06	0.4	0.13	1.21	0.32
	UGI+	12.7	1.93	0.24	0.13	0.86	0.05
	siUDG, UGI+	16.97	0.93	0.36	0.25	0.9	0.04

[00510] Table 9. Fold change relative to GFP in % of edited reads in U-2OS cells

Construct	Overexpression/ siRNA	C-to-T	P value	C-to-A/G or Indel	P value
BC27	GFP+	not applicable			
	SMUG1+	-1.37	0.1953	1.20	0.1953
	UDG+	-1.75	0.0864	1.31	0.0864
	UGI+	1.92	0.0091	-3.09	0.0091
	siUDG, UGI+	1.97	0.0042	-3.48	0.0042
BC22	GFP+	not applicable			
	SMUG1+	-2.08	0.0004	2.67	0.0004
	UDG+	-1.52	0.0031	2.10	0.0031
	UGI+	1.21	0.0048	-2.97	0.0048
	siUDG, UGI+	1.22	0.0041	-3.46	0.0041

[00511] Table 10. Percent editing under different DNA repair pathway conditions in HuH-7 cells

HuH-7		C-to-T %		C-to-A/G %		Indel %	
Construct	**Overexpression/ siRNA**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
BC27	**GFP+**	**18.70**	**0.70**	**13.24**	**0.95**	**10.55**	**1.11**

HuH-7		C-to-T %		C-to-A/G %		Indel %	
Construct	Overexpression/ siRNA	Mean	SD	Mean	SD	Mean	SD
	SMUG1+	9.57	1.91	14.78	1.64	13.23	1.24
	UDG+	4.57	0.46	7.45	0.23	7.74	0.49
	UGI+	42.40	2.10	0.82	0.34	1.88	0.47
	siUDG, UGI+	41.37	2.73	0.94	0.37	1.32	0.30
BC22	GFP+	41.77	2.30	15.02	0.73	15.00	1.06
	SMUG1+	19.20	1.61	21.89	2.17	25.98	1.86
	UDG+	12.97	1.89	8.27	0.82	12.11	1.14
	UGI+	64.53	1.32	2.50	0.30	4.05	1.45
	siUDG, UGI+	65.33	2.28	2.05	1.45	4.24	1.53

[00512] **Table 11. Fold change relative to GFP in % of edited reads in Huh-7 cells**

Construct	Overexpression/ siRNA	C-to-T	P value	C-to-A/G or indel	P value
BC27	GFP+	not applicable			
	SMUG1+	-1.73	0.0014	1.33	0.001
	UDG+	-1.86	<0.0001	1.37	<0.0001
	UGI+	2.13	<0.0001	-9.30	<0.0001
	siUDG, UGI+	2.15	<0.0001	-10.81	<0.0001
BC22	GFP+	not applicable			
	SMUG1+	-2.03	<0.0001	1.71	<0.0001
	UDG+	-1.50	0.0001	1.46	0.0001
	UGI+	1.56	<0.0001	-4.53	<0.0001
	siUDG, UGI+	1.57	<0.0001	-4.76	<0.0001

Example 3 - UGI mRNA titration in T cells

[00513] The mRNA encodes a fusion protein, BC22n (SEQ ID NO: 3), which is, from N- terminus to C-terminus, a *H. sapiens* APOBEC3A, an XTEN linker, a D10ACas9 nickase, a linker, and an SV40 NLS**.** Notably the BC22n polypeptide lacks a UGI. T cells were edited with BC22n using and a variable amount of UGI mRNA to determine the impact

of trans UGI levels on editing profile. This experiment was performed using 2 different UGI mRNAs each encoding the same protein using a different open reading frame (SEQ ID Nos: 26 and 35).

3.1 T cell preparation

[00514] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. No. 130-070-525) on the LOVO device. T cells were isolated via positive selection using CD4 and CD8 magnetic beads (Miltenyi Biotec Cat. No. 130-030-401/130-030-801) using the CliniMACS® Plus and CliniMACS® LS disposable kit. T cells were aliquoted into vials and cryopreserved in a 1:1 formulation of Cryostor® CS10 (StemCell Technologies Cat. No. 07930) and Plasmalyte A (Baxter Cat. No. 2B2522X) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell basal media composed of X-VIVO 15™ serum-free hematopoietic cell medium (Lonza Bioscience) containing 5% (v/v) of fetal bovine serum, 50 μM of 2-Mercaptoethanol, 10 mM of N-Acetyl-L-(+)-cysteine, 10 U/mL of Penicillin-Streptomycin, in addition to 1X cytokines (200 U/mL of recombinant human interleukin-2, 5 μg/mL of recombinant human interleukin-7 and 5 μg/mL of recombinant human interleukin-15). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec). Cells were expanded in T cell basal media for 72 hours prior to electroporation.

3.2 Electroporation of T cells using mRNA and sgRNA

[00515] A solution containing mRNAs encoding BC22n (SEQ ID NO: 2) and one species of UGI (SEQ ID NO: 26 or 35) was prepared in sterile water. 50 μM TRAC targeting sgRNA G016017 (SEQ ID NO: 184) were removed from their storage plates and denatured for 2 minutes at 95°C before cooling on ice. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of BC22n mRNA, 20 pmols of sgRNA and UGI mRNA ranging from 0.02 ng to 26 ng (dilution factor of 12.24), in a final volume of 20 μL of P3 electroporation buffer. This mRNA + sgRNA + T cell mix was transferred in triplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 μL of T cell basal media without cytokines for 10 minutes before being transferred to a new flat-bottom 96-well plate containing an additional 100 μL of T cell basal media with 2X concentration of cytokines.

Electroporated T cells were subsequently cultured for 3 additional days and were collected for NGS sequencing as described in Example 1.

[00516] The constant dose of 200 ng of BC22n mRNA drove high total editing (>92%) across all conditions (Tables 12 and 13, sum of indel, C-to-A/G and C-to-T). The proportion of edits that were C-to-T-only increased in a dose-responsive manner to UGI mRNA concentration in the electroporation. At the highest dose of either UGI mRNAs SEQ ID NOs: 25 or 34, the percentage of sequencing reads containing indels and C-to-A/G mutations dropped to 7.4±0.4, 7.4±0.6 and 2.9±0.2, 3.6±0.3, respectively (Tables 12 and 13).

[00517] **Table 12- Editing as percent of total reads with UGI mRNA (SEQ ID NO: 26)**

UGI mRNA (ng)	% Editing					
	Indel		C-to-A/G		C-to-T	
	Mean	SD	Mean	SD	Mean	SD
Untreated	0.2	0.1	2.1	0.0	0.2	0.0
0.000	51.2	0.8	27.2	0.4	18.3	0.3
0.025	50.0	1.1	28.2	0.6	18.6	0.6
0.051	48.2	2.0	28.4	0.3	19.0	1.0
0.102	49.8	0.5	28.1	0.3	18.7	0.8
0.203	49.5	1.1	27.9	0.6	18.7	0.3
0.406	47.7	2.0	27.5	0.5	19.4	0.7
0.813	47.8	0.6	28.0	0.5	19.5	0.9
1.625	47.9	0.1	26.6	0.5	21.5	0.3
3.250	41.9	0.6	24.7	0.8	28.9	0.7
6.500	30.0	0.5	21.1	0.6	44.1	1.0
13.000	18.7	2.6	15.5	0.6	58.3	1.0
26.000	7.4	0.4	7.4	0.6	80.1	2.1

[00518] **Table 13 - Editing as percent of total reads with UGI mRNA (SEQ ID NO: 35)**

UGI mRNA (ng)	% Editing					
	Indel		C-to-A/G		C-to-T	
	Mean	SD	Mean	SD	Mean	SD
Untreated	0.2	0.1	2.1	0.0	0.2	0.0
0.000	51.2	0.8	27.2	0.4	18.3	0.3
0.025	48.2	1.7	27.0	0.7	19.0	0.5
0.051	48.7	1.3	27.3	0.7	19.4	0.4
0.102	47.5	2.3	27.1	0.2	18.5	0.5

0.203	49.7	1.3	27.4	0.9	18.7	0.7
0.406	49.4	0.7	28.0	0.6	19.4	0.4
0.813	47.1	2.3	26.6	2.2	19.6	0.5
1.625	43.8	2.5	27.3	2.0	25.2	1.0
3.250	33.1	3.0	22.8	0.8	38.4	0.7
6.500	24.0	0.4	18.2	0.3	53.9	0.8
13.000	11.2	1.9	10.5	0.6	73.1	1.6
26.000	2.9	0.2	3.6	0.3	89.1	1.6

Example 4: *In vivo* editing with UGI in cis

[00519] LNPs formulated with editor mRNA and sgRNA G000282 at a 1:1 RNA weight ratio were tested for editing efficacy and editing outcomes *in vivo.* Experimental groups included TSS buffer only; mRNA encoding cleavase-spCas9 (SEQ ID NO: 8), mRNA encoding BC22 which includes human APOBEC3A fused to D10A SpyCas9 and one copy of UGI (SEQ ID NO: 20); and mRNAs encoding BE3 which includes rat APOBEC1 fused to D10A SpyCas9 and one copy of UGI (SEQ ID NOs: 14 and 17).

[00520] CD-1 female mice ranging from 6-10 weeks of age (n=5/group) were used in this study. LNPs were administered intravenously via tail vein injection at a dose of 1 mg/kg of total RNA to body weight. The animals were periodically observed for adverse effects for at least 24 hours post dose. Six days after treatment, animals were euthanized by cardiac puncture under isoflurane anesthesia; blood and liver tissue were collected for downstream analysis. Blood was collected into serum separator tubes or into tubes containing buffered sodium citrate for plasma. Liver punches weighting between 5 and 15 mg were collected for isolation of genomic DNA and total RNA.

[00521] Table 14 and Figure 7A describe the editing results in mouse liver. The construct containing human APOBEC3A (BC22) showed total editing levels comparable to those of the Cas9 cleavase construct, while the rat APOBEC1 (BE3) constructs showed less than half that total editing activity. The construct containing human APOBEC3A also achieved more than twice the absolute percentage of C-to-T base conversion compared to the rat APOBEC1 constructs (BE3).

[00522] Table 14. TTR editing in liver tissue (Indel = insertions or deletions; SD = standard deviation)

	% Editing					
	C-to- T		C-to-A/G		Indel	
Editor	Mean	SD	Mean	SD	Mean	SD
TSS	0.15	0.04	1.95	0.26	0.31	0.11
Cas9 (SEQ ID NO: 8)	0.05	0.02	0.72	0.23	69.59	1.28
BE3 (SEQ ID NO: 14)	16.53	7.49	8.69	3.01	4.31	1.79
BE3 (SEQ ID NO: 17)	10.23	2.72	7.41	1.90	4.71	1.92
BC22 (SEQ ID NO: 20)	45.75	1.17	4.98	0.49	22.0	1.82

Example 5 – In vivo editing with UGI in trans

[00523] *In vivo* editing profiles of deaminase containing constructs were compared to Cas9 when UGI was delivered *in trans* (as a separate mRNA). The constructs used encoded a fusion protein including D10A SpyCas9 with a deaminase. Further, gene expression differences between these editing conditions were analyzed through transcriptomic analysis.

5.1 *In vivo* editing as assayed by NGS

[00524] Forty-five commercially available CD-1 female mice ranging from 6-10 weeks of age (n=5 per group) were used in this study. Animals were weighed pre-dose for dosing calculations. Each RNA species was formulated separately in an LNP. Formulations containing editor mRNA, UGI mRNA and sgRNA were mixed in a w/w ratio of RNA cargos of 6:3:2 (editor mRNA:sgRNA:UGI mRNA). The formulation mixture for Group 3 contained only editor mRNA and sgRNA and these were mixed in a w/w ratio of 2:1 (editor mRNA:sgRNA). Apart from the negative control group, which was dosed with TSS buffer only, all groups received sgRNA G000282. Formulations were administered intravenously via tail vein injection according to the doses listed in Table 15. Animals were periodically observed for adverse effects for at least 24 hours post-dose. Six days after treatment, animals

were euthanized by cardiac puncture under isoflurane anesthesia; liver tissue were collected for downstream analysis. Liver punches weighing between 5 and 15 mg were collected for isolation of genomic DNA and total RNA. Genomic DNA samples were analyzed with NGS sequencing as described in Example 1.

[00525] Table 15 – Animal groups and the respective LNPs.

Sample	LNP cargos	Editor mRNA + sgRNA dose (mg/kg)
TSS	None	0
Cas9, UGI	Cas9 (SEQ ID NO: 11)	0.3 mg/kg
	UGI (SEQ ID NO: 26)	
	G0000282	
BC22n	BC22n (SEQ ID NO: 2)	0.3 mg/kg
	G0000282	
BC22n, UGI	BC22n (SEQ ID NO: 2)	0.3 mg/kg Or 0.1 mg/kg
	UGI (SEQ ID NO: 26)	
	G0000282	

[00526] Editing data are shown in Table 16 and Figure 8. Treatment with Cas9 mRNA led to 58.2% editing of the *Ttr* gene. Animals treated with BC22n mRNA in the absence of UGI mRNA *in trans* displayed 29.26% C-to-T conversions and 28% indel, while those treated with BC22n and UGI mRNA showed a higher C-to-T editing purity (56.67% C-to-T conversions and only 5.6% indel).

[00527] Lowering the lipid dose from 0.3 to 0.1 mg/kg led to lower editing levels (31.39% C-to-T conversions and 3.67% indel).

[00528] Table 16 – Ttr editing levels in mice treated with different LNP combinations (SD = standard deviation).

		mRNA(s)	**TSS**	**Cas9, UGI**	**BC22n**	**BC22n, UGI**	**BC22n, UGI**
		Dose (editor +sgRNA)	**0**	**0.3 mg/kg**	**0.3 mg/kg**	**0.3 mg/kg**	**0.1 mg/kg**
% *Ttr* Editing	**C-to-T**	Average	0.2%	0.1%	29.3%	56.7%	31.4%
		SD	0.0%	0.1%	5.7%	2.9%	4.9%
	C-to-A/G	Average	0.5%	0.3%	8.1%	1.8%	1.7%
		SD	0.1%	0.1%	1.3%	0.2%	0.1%
	Indel	Average	0.2%	58.2%	28.0%	5.7%	3.7%
		SD	0.0%	4.0%	5.0%	0.5%	0.1%
		SD	14.1%	6.7%	9.6%	3.5%	9.0%

5.2 Whole transcriptome sequencing

[00529] Liver punches were mixed with 800 µL of TRIzol reagent (Thermo Fisher Scientific, Cat No. 15596026) in Lysing Matrix D tubes (MPBio, Cat. No. 116913100) which contain ceramic beads. Tissue was homogenized in a bead beater for 45 seconds and transferred to ice. Lysing Matrix D tubes were spun down at maximum speed for 5 min at 4C and ~600 µL of TRIzol (without tissue debris) was mixed with an equal volume of absolute ethanol. The mixture was loaded in the Directzol RNA miniprep column (Zymo Research, Cat No. R2051) and RNA was extracted following the manufacturer's protocol. Purified RNA samples were quantified in a NanoDrop™ 8000 spectrophotometer (Thermo Fisher Scientific) and diluted to 41.67 ng/µL using nuclease-free water. From each experimental group, 2 samples were randomly chosen for further transcriptomic analysis. 500 ng (12 µL) of purified total RNA were depleted of ribosomal RNA (rRNA) components using the NEBNext® rRNA Depletion Kit (New England Biolabs, Cat. No. E6350L) according to the manufacturer's instructions. rRNA-depleted samples were converted into double-stranded DNA libraries using NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England Biolabs, Cat No. E7765S) following the manufacturer's protocol. Amplified libraries were quantified in a Qubit 4 fluorometer and the average fragment size of each library was obtained by capillary electrophoresis. Libraries were pooled at an equimolar concentration of 4 nM and pair-end sequenced using a high-output 300-cycle kit (Illumina, Cat No. 20024908) in a NextSeq550 sequencing platform (Illumina).

[00530] Data processing for differential gene expression analysis

[00531] Sequencing reads in FASTQ format were generated and demultiplexed using the bcl2fastq program (Illumina, v2.20). Reads were assigned to a sample if the Hamming distance (Hamming, R.W. Bell Syst. Tech. J. 29, 147–160) between each index read and the sample indexes was less than or equal to one. The sequencing quality was examined with FastQC program (v0.11.9) (Andrews S. Babraham Inst.). Ribosomal RNA reads were identified by aligning all reads to human rRNA sequences (GenBank U13369.1) with Bowtie2 (v2.3.5.1) (Langmead, B. and Salzberg, S.L. Nat. Methods 9, 357–359). Transcriptome quantification was performed using Salmon (v0.14.1) (Patro R., et al. Nat. Methods 14, 417-419) with non-ribosomal RNA reads. Differential gene expression analysis was carried out using DESeq2 (v1.26.0) (Love, M.I., et al. Genome Biol. 15, 550) on the outputs of Salmon. Genes or transcripts with Benjamini-Hochberg adjusted p-values less than 0.05 were determined to be differentially expressed.

[00532] RNA-Seq analysis revealed that treatment with BC22n mRNA and UGI mRNA *in trans* led to only 53 differentially expressed genes, compared to 223 events in those animals treated with BC22n alone (i.e. no UGI mRNA) and 127 events in animals treated with both Cas9 and UGI mRNA *in trans* (Figs. 9A-9C).

Example 6 – Editing in T Cells with UGI in trans

6.1 Editing in T cells

[00533] T cells were edited at the CIITA locus with UGI in trans and either BC22n or Cas9 to assess the impact on editing type on MHC II antigens.

[00534] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. No. 130-070-525) on the LOVO device. T cells were isolated via positive selection using CD4 and CD8 magnetic beads (Miltenyi Biotec Cat. No. 130-030-401/130-030-801) using the CliniMACS® Plus and CliniMACS® LS disposable kit. T cells were aliquoted into vials and cryopreserved in a 1:1 formulation of Cryostor® CS10 (StemCell Technologies Cat. No. 07930) and Plasmalyte A (Baxter Cat. No. 2B2522X) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell basal media composed of X-VIVO 15™ serum-free hematopoietic cell medium (Lonza Bioscience) containing 5% (v/v) of fetal bovine serum, 50 μM of 2-Mercaptoethanol, 10 mM of N-Acetyl-L-(+)-cysteine, 10 U/mL of Penicillin-Streptomycin, in addition to 1X cytokines (200 U/mL of recombinant human interleukin-2, 5 μg/mL of recombinant human interleukin-7 and 5 μg/mL of recombinant human interleukin-15). T-cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec). Cells were expanded in T cell basal media containing TransAct™ for 72 hours prior to electroporation.

Electroporation of T cells

[00535] A solution containing mRNAs encoding Cas9 (SEQ ID NO. 11), BC22n (SEQ ID NO: 2) or UGI (SEQ ID NO: 26) was prepared in sterile water. 50 μM CIITA sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C before cooling on ice. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of editor mRNA, 200 ng of UGI mRNA and 20 pmols of sgRNA as described in Table 17 in a

final volume of 20 uL of P3 electroporation buffer. This mix was transferred in triplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 µL of T cell basal media without cytokines for 10 minutes before being transferred to a new flat-bottom 96-well plate containing an additional 100 µL of T cell basal media supplemented with 2X cytokines. The resulting plate was incubated at 37°C for 4 days. After 96 hours, T cells were diluted 1:3 into fresh T cell basal media with 1X cytokines. Electroporated T cells were subsequently cultured for 3 additional days and were collected for flow cytometry analysis, NGS sequencing, and transcriptomics as described in Example 1.

Flow cytometry and NGS sequencing

[00536] On day 7 post-editing, T cells were phenotyped by flow cytometry to determine MHC class II protein expression. Briefly, T cells were incubated in a cocktail of antibodies targeting HLA-DR, DQ, DP-PE (BioLegend® Cat. No. 361704) and Isotype Control-PE (BioLegend® Cat. No. 400234). Cells were subsequently washed, processed on a Cytoflex flow cytometer (Beckman Coulter) and analyzed using the FlowJo software package. T cells were gated based on size, shape, viability, and MHC II expression. DNA samples were subjected to PCR and subsequent NGS analysis, as described in Example 1. Table 17 and Figure 10 show CIITA gene editing. For both Cas9 and BC22n conditions, total editing went to near completion, above 95%. Table 17 and Figure 11 show MHC class II protein expression following electroporation with UGI mRNA combined with Cas9 or BC22n mRNA. For G018117, editing with BC22n lead to 80.50% MHC II negative cells, while editing with Cas9 lead to 51.63% MHC II antigen negative cells.

[00537] Table 17– CIITA editing in T cells presented at a percentage of total NGS reads; Flow cytometry assessment of the percentage of cells lacking surface markers HLA DQ/DP/DR

Editor	Guide	% C-to-T		% C-to-A/G		% Indel		% MHC II neg		
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	N
Cas9	G018117	0.0	0.0	0.1	0.0	98.0	0.2	70.3	1.5	3
	G018118	0.0	0.0	0.1	0.1	97.5	0.5	73.7	0.5	3
	G018120	0.1	0.1	0.2	0.0	92.3	0.3	83.1	1.1	3
	G018076	0.0	0.0	0.1	0.0	98.3	0.1	62.8	1.1	3
	G018100	0.0	0.0	0.1	0.0	98.9	0.2	84.1	1.3	3
	G018091	0.0	0.0	0.0	0.0	99.4	0.2	96.2	0.4	3
	No guide	not reported						10.5	0.0	1
BC22n	G018117	95.7	0.2	2.3	0.1	1.1	0.1	99.2	0.1	3

	G018118	95.8	0.3	2.3	0.1	1.2	0.3	99.0	0.1	3
	G018120	95.7	0.5	2.6	0.4	0.7	0.1	99.0	0.1	3
	G018076	96.5	0.4	1.2	0.2	1.1	0.3	98.5	0.2	3
	G018100	90.7	0.3	2.8	0.3	5.1	0.7	95.7	0.6	3
	G018091	95.4	0.6	2.8	0.1	1.1	0.3	98.7	0.2	3
	No guide	not reported						10.8	0.0	1

6.2 – Gene expression analysis in T cells

Whole transcriptome sequencing

[00538] On day 7 post-editing, T cells treated with G018117 and G18078 were harvested and preserved at -80C for future processing. Total RNA was extracted from samples in TRIzol™ reagent using the Direct-zol RNA microprep kit (Zymo Research, Cat No. R2062) following the manufacturer's protocol. Purified RNA samples were quantified in a NanoDrop™ 8000 spectrophotometer (Thermo Fisher Scientific) and diluted to 41.67 ng/μL using nuclease-free water. From each experimental triplicate shown in Figs. 13A-15B, samples per group were randomly chosen for transcriptomic analysis. 500 ng (12 μL) of purified total RNA were depleted of ribosomal RNA (rRNA) components using the NEBNext® rRNA Depletion Kit (New England Biolabs, Cat. No. E6350L) according to the manufacturer's instructions. rRNA-depleted samples were converted into double-stranded DNA libraries using NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England Biolabs, Cat No. E7765S) following the manufacturer's protocol. Amplified libraries were quantified in a Qubit 4 fluorometer and the average fragment size of each library was obtained by capillary electrophoresis. Libraries were pooled at an equimolar concentration of 4 nM and pair-end sequenced using a high-output 300-cycle kit (Illumina, Cat No. 20024908) in a NextSeq550 sequencing platform (Illumina).

Data processing for differential gene expression analysis

[00539] Sequencing reads in FASTQ format were generated and demultiplexed using the bcl2fastq program (Illumina, v2.20). Reads were assigned to a sample if the Hamming distance (Hamming, R.W. Bell Syst. Tech. J. 29, 147–160) between each index read and the sample indexes was less than or equal to one. The sequencing quality was examined with FastQC program (v0.11.9) (Andrews S. Babraham Inst.). Ribosomal RNA reads were identified by aligning all reads to human rRNA sequences (GenBank U13369.1) with Bowtie2 (v2.3.5.1) (Langmead, B. and Salzberg, S.L. Nat. Methods 9, 357–359). Transcriptome quantification was performed using Salmon (v0.14.1) (Patro R., et al. Nat.

Methods 14, 417-419) with non-ribosomal RNA reads. Differential gene expression analysis was carried out using DESeq2 (v1.26.0) (Love, M.I., et al. Genome Biol. 15, 550) on the outputs of Salmon. Genes or transcripts with Benjamini-Hochberg adjusted p-values less than 0.05 were determined to be differentially expressed. Lists of differentially expressed genes were analyzed in terms of gene ontology using Metascape (Zhou, Y., et al. Nat. Comm. 10, 1523). Protein-protein interactions were determined using the BioGrid, InWeb_IM and OmniPath8 databases (Li, T., et al. Nat. Methods 14, 61-64; Stark, C., et al. Nucleic Acids Res. 34, 535-539; Türei, D., et al. Nat. Methods 13, 966-967). Densely connected networks were identified using the molecular complex detection (MCODE) algorithm (Bader, G.D., et al. BMC Bioinformatics 4, 1-27) and the three best-scoring terms by p-value were retained as the functional description of the corresponding network components.

[00540] Compared to samples treated with Cas9 mRNA, T cells electroporated with BC22n mRNA displayed a significantly stronger downregulation of MHC class II genes and the HLA-associated CD74 gene (Table 18 and 19). Minimal effects on class I MHC genes were observed (Table 20 and Table 21). In terms of transcriptome-wide differential gene expression events, treatment with BC22n mRNA led to fewer differentially expressed genes (p. adjusted < 0.05) when compared to Cas9 mRNA. In T cells electroporated with sgRNA G018076, a total of 553 and 65 differential gene expression events were observed for Cas9 and BC22n mRNA treatments, respectively (Figs. 12A-12B). A similar trend was observed in T cells electroporated with sgRNA G018117, which displayed 303 and 30 differential gene expression events when treated with Cas9 and BC22n mRNA, respectively (Figs. 13A-13B). Fewer protein-protein interaction networks were identified among the list of differentially expressed genes in T cells treated with BC22n mRNA when compared to those treated with Cas9 mRNA (Figs. 14A-14B and 15A-15B).

[00541] **Table 18 – Differential gene expression of class II HLA genes in T cells** (ns = not significant, * = p.adj.<0.05, ** = p.adj.<0.01, *** = p.adj.<0.001). For transcript quantification data, refer to Table 19.

	G018076 sgRNA				G018117 sgRNA			
	Cas9 vs control		BC22n vs control		Cas9 vs control		BC22n vs control	
Gene	Fold change	p. adj.	Fold change	p. adj.	Fold change	p. adj.	Fold change	p. adj.
CD74	0.446	***	0.116	***	0.360	***	0.100	***
HLA-DMA	0.457	***	0.150	***	0.356	***	0.146	***
HLA-DMB	0.363	***	0.113	***	0.267	***	0.091	***
HLA-DOA	0.450	***	0.299	***	0.420	***	0.280	***
HLA-DPA1	0.474	***	0.181	***	0.385	***	0.170	***

	G018076 sgRNA				G018117 sgRNA			
	Cas9 vs control		BC22n vs control		Cas9 vs control		BC22n vs control	
Gene	Fold change	p. adj.	Fold change	p. adj.	Fold change	p. adj.	Fold change	p. adj.
HLA-DPB1	0.381	***	0.087	***	0.300	***	0.073	***
HLA-DQA1	0.316	***	0.017	***	0.214	***	0.009	***
HLA-DQA2	0.215	***	0.007	***	0.221	***	0.011	***
HLA-DQB1	0.383	***	0.069	***	0.290	***	0.069	***
HLA-DQB1-AS1	0.323	***	0.102	***	0.311	***	0.097	***
HLA-DRA	0.288	***	0.004	***	0.205	***	0.002	***
HLA-DRB1	0.287	***	0.029	***	0.207	***	0.027	***
HLA-DRB3	0.268	***	0.012	***	0.219	***	0.004	***
HLA-DRB4	0.282	***	0.024	***	0.224	***	0.026	***

[00542] **Table 19 – Transcript quantification of the expression of class II HLA genes in T cells.** Each square contains the average number of transcripts from a given gene per one million of mRNA molecules. For statistical significance, please refer to Table 19.

Guide	none	G0188117		G018076	
mRNAs	UGI	Cas9 UGI	BC22n UGI	Cas9 UGI	BC22n UGI
CD74	385	154	38	174	45
HLA-DMA	27	11	4	12	4
HLA-DMB	11	3	1	4	1
HLA-DOA	3	2	1	2	1
HLA-DPA1	70	30	12	34	13
HLA-DPB1	16	5	1	6	1
HLA-DQA1	12	3	0	4	0
HLA-DQA2	3	1	0	1	0
HLA-DQB1	60	19	4	23	4
HLA-DQB1-AS1	9	3	1	3	1
HLA-DRA	120	27	0	35	1
HLA-DRB1	117	27	3	34	3
HLA-DRB3	17	4	0	5	0
HLA-DRB4	13	3	0	4	0

[00543] **Table 20** - Differential gene expression of class I HLA genes in T cells harvested 7 days post-treatment with different mRNA combinations and CIITA-targeting sgRNAs (ns = not significant, * = p.adj.<0.05, ** = p.adj.<0.01, *** = p.adj.<0.001). For transcript quantification data, refer to Table 21.

	G018076 sgRNA		G018117 sgRNA	
	Cas9 vs control	BC22n vs control	Cas9 vs control	BC22n vs control

Gene	Fold change	p. adj.	Fold change	p. adj.	Fold change	p. adj.	Fold change	p. adj.
HLA-A	0.995	ns	0.910	*	0.969	ns	0.926	ns
HLA-B	1.001	ns	0.881	**	1.043	ns	0.913	ns
HLA-C	1.013	ns	0.917	ns	0.995	ns	0.922	ns
HLA-E	0.925	ns	0.919	ns	0.870	***	0.949	ns
HLA-F	0.897	ns	0.812	**	0.930	ns	0.891	ns

[00544] **Table 21** – Transcript quantification of the expression of class I HLA genes in T cells. Each square contains the average number of transcripts from a given gene per one million of mRNA molecules. For statistical significance, please refer to Table 20.

Guide	**none**	**G0188117**		**G018076**	
mRNAs	**UGI**	**Cas9 UGI**	**BC22n UGI**	**Cas9 UGI**	**BC22n UGI**
HLA-A	880	947	797	887	814
HLA-B	457	528	407	463	409
HLA-C	479	528	430	490	445
HLA-E	166	161	154	156	155
HLA-F	68	70	59	61	56

Example 7 – Editing assessment at APOBEC3A Hotspots

[00545] T cell preparation

[00546] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. No. 130-070-525) on the LOVO device. T cells were isolated via positive selection using CD4 and CD8 magnetic beads (Miltenyi Biotec Cat. No. 130-030-401/130-030-801) using the CliniMACS® Plus and CliniMACS® LS disposable kit. T cells were aliquoted into vials and cryopreserved in a 1:1 formulation of Cryostor® CS10 (StemCell Technologies Cat. No. 07930) and Plasmalyte A (Baxter Cat. No. 2B2522X) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell basal media composed of X-VIVO 15™ serum-free hematopoietic cell medium (Lonza Bioscience) containing 5% (v/v) of fetal bovine serum, 50 μM of 2-Mercaptoethanol, 10 mM of N-Acetyl-L-(+)-cysteine, 10 U/mL of Penicillin-Streptomycin, in addition to 1X cytokines (200 U/mL of recombinant human interleukin-2, 5 μg/mL of recombinant human interleukin-7 and 5 μg/mL of recombinant human interleukin-15). T-cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec). Cells were expanded in T cell basal media containing TransAct™ for 72 hours prior to electroporation.

[00547] mRNA and sgRNA electroporation of T cells

[00548] A solution containing mRNA encoding Cas9 (SEQ ID NO: 11), BC22n (SEQ ID NO: 2) or UGI (SEQ ID NO: 26) was prepared in sterile water. A 50 μM TRAC targeting sgRNA (G016017) was removed from its storage plate and denatured for 2 minutes at 95°C before cooling on ice. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of each mRNA and 20 pmols of sgRNA in a final volume of 20 μL of P3 electroporation buffer. The T cell mix was transferred in triplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 μL of T cell basal media without cytokines for 10 minutes before being transferred to a new flat-bottom 96-well plate containing an additional 100 μL of T cell basal media supplemented with 2X cytokines. The resulting plate was incubated at 37°C for 4 days. After 96 hours, T cells were diluted 1:3 into fresh T cell basal media with 1X cytokines. Electroporated T cells were subsequently cultured for 3 additional days and were collected for NGS sequencing.

[00549] NGS sequencing

[00550] On day 7 post-editing, DNA samples were subjected to PCR and subsequent NGS analysis, as described in Example 1. This study characterized the on-target TRAC locus in addition to 10 genomic loci previously described as mutational hotspots in tumor samples positive for APOBEC enzymes (Buisson et al., 2019). The chromosomal location of these sites is listed on Table 22.

[00551] Tables 23, 24, and 25 show C-to-T, C-to-A/G and indel editing levels in the on-target locus and predicted APOBEC hotspot sites for all sample groups. A graphical representation of these results in shown in Figs. 16A-16C. Compared to control samples treated with Cas9 mRNA, T cells electroporated with BC22n mRNA, UGI mRNA or BC22n mRNA in addition to both UGI mRNA and the sgRNA G016017 did not display any significant changes in the editing profile of 10 genomic loci previously reported as mutational hotspots in tumor samples positive for APOBEC enzymes (Buisson R., et al. (2019). Science 364, 1-8.,). High levels of deamination at the on-target TRAC locus were observed in samples treated with BC22n mRNA, UGI mRNA and sgRNA G016017, but absent in those treated with BC22n mRNA alone, Cas9 mRNA alone or UGI mRNA alone.

[00552] Table 22 – List of mutational hotspots evaluated in this study, their Entrez gene IDs and genomic locations.

Gene symbol	Entrez gene ID	Chromosome	Genomic location	Strand
FAM83G	644815	17	18,907,093	-
PCNT	5116	21	47,783,764	-
CUEDC2	79004	10	104,184,490	-
PANX2	56666	22	50,615,649	-
GSE1	23199	16	85,691,140	-
ZNF672	79894	1	249,142,293	+
NLRP1	22861	17	5,461,801	+
MLLT4/AFDN	4301	6	168,352,586	+
OR5A1	219982	11	59,211,279	+
KIAA1522	57648	1	33,237,497	+

[00553] Table 23 – Levels of C-to-T base conversion in the on-target TRAC locus in addition to 10 previously described APOBEC hotspots in the human genome. SD = standard deviation; ND = not determined.

	C-to-T Editing							
	Cas9 mRNA		BC22n mRNA		UGI mRNA		BC22n mRNA + UGI mRNA + G016017	
Gene	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TRAC (on-target)	0.14%	0.02%	0.13%	0.01%	0.19%	0.02%	96.26%	0.25%
FAM83G	0.23%	0.05%	0.22%	0.03%	0.22%	0.04%	0.28%	0.05%
PCNT	2.02%	ND	1.97%	0.09%	1.92%	0.04%	1.99%	0.00%
CUEDC2	0.32%	0.12%	0.36%	0.04%	0.27%	0.01%	0.28%	0.00%
PANX2	0.16%	0.01%	0.19%	0.00%	0.21%	ND	0.26%	0.01%
GSE1	0.20%	0.03%	0.11%	ND	0.15%	ND	0.12%	0.02%
ZNF672	0.15%	0.02%	0.19%	0.03%	0.20%	0.01%	0.18%	0.02%
NLRP1	0.14%	0.00%	0.15%	0.01%	0.15%	0.00%	0.16%	0.02%
MLLT4	0.12%	0.01%	0.13%	0.03%	0.13%	0.01%	0.14%	0.00%
OR5A1	0.12%	0.03%	0.13%	0.01%	0.13%	0.02%	0.14%	0.04%
KIAA1522	0.24%	0.03%	0.29%	0.08%	0.25%	0.02%	0.24%	0.01%

[00554] Table 24– Levels of C-to-A/G transversions in the on-target TRAC locus in addition to 10 previously described APOBEC hotspots in the human genome. SD = standard deviation; ND = not determined.

	C-to-A/G Editing							
	Cas9 mRNA		BC22n mRNA		UGI mRNA		BC22n mRNA + UGI mRNA + G016017	
Gene	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TRAC (on-target)	0.29%	0.03%	0.30%	0.03%	2.73%	0.69%	2.38%	0.20%

FAM83G	4.22%	0.11%	4.11%	0.22%	4.66%	0.03%	4.13%	0.06%
PCNT	0.45%	ND	0.53%	0.01%	0.52%	0.13%	0.53%	0.04%
CUEDC2	1.28%	0.01%	1.34%	0.11%	1.27%	0.20%	1.36%	0.01%
PANX2	2.18%	0.15%	2.32%	0.24%	2.48%	ND	2.37%	0.27%
GSE1	0.72%	0.06%	0.58%	ND	0.70%	ND	0.66%	0.03%
ZNF672	0.39%	0.03%	0.42%	0.02%	0.36%	0.04%	0.39%	0.02%
NLRP1	0.62%	0.06%	0.62%	0.01%	0.60%	0.02%	0.57%	0.08%
MLLT4	1.27%	0.09%	1.21%	0.01%	1.26%	0.03%	1.32%	0.03%
OR5A1	0.72%	0.02%	0.75%	0.06%	0.70%	0.04%	0.73%	0.03%
KIAA1522	0.72%	0.08%	0.70%	0.07%	0.76%	0.04%	0.74%	0.04%

[00555] Table 25 – Levels of insertions and deletions (indel) in the on-target TRAC locus in addition to 10 previously described APOBEC hotspots in the human genome. SD = standard deviation; ND = not determined.

	Indel							
	Cas9 mRNA		**BC22n mRNA**		**UGI mRNA**		**BC22n mRNA + UGI mRNA + G016017**	
Gene	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TRAC (on-target)	0.08%	0.00%	0.11%	0.01%	0.11%	0.02%	0.95%	0.26%
FAM83G	3.78%	0.10%	3.63%	0.11%	3.90%	0.13%	3.67%	0.50%
PCNT	0.11%	ND	0.23%	0.00%	0.18%	0.02%	0.21%	0.02%
CUEDC2	0.09%	0.05%	0.14%	0.03%	0.13%	0.03%	0.14%	0.02%
PANX2	0.47%	0.01%	0.54%	0.10%	0.26%	ND	0.52%	0.02%
GSE1	0.23%	0.06%	0.13%	ND	0.15%	ND	0.36%	0.05%
ZNF672	0.71%	0.06%	0.80%	0.08%	0.72%	0.01%	0.78%	0.15%
NLRP1	0.07%	0.03%	0.07%	0.03%	0.08%	0.01%	0.07%	0.01%
MLLT4	0.26%	0.00%	0.25%	0.01%	0.25%	0.02%	0.27%	0.01%
OR5A1	0.07%	0.02%	0.08%	0.01%	0.08%	0.01%	0.07%	0.01%
KIAA1522	0.38%	0.01%	0.46%	0.02%	0.42%	0.02%	0.41%	0.00%

Example 8 - LNP titration in T cells with fixed ratio of BC22n:UGI

[00556] Using LNP delivery to activated human T cells, the potency of single-target and multi-target editing was assessed with either Cas9 or BC22n.

8.1 T cell preparation

[00557] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. No. 130-070-525) on the LOVO device. T cells were isolated via positive selection using CD4 and CD8 magnetic beads (Miltenyi Biotec Cat. No. 130-030-

401/130-030-801) using the CliniMACS® Plus and CliniMACS® LS disposable kit. T cells were aliquoted into vials and cryopreserved in a 1:1 formulation of Cryostor® CS10 (StemCell Technologies Cat. No. 07930) and Plasmalyte A (Baxter Cat. No. 2B2522X) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell basal media composed of X-VIVO 15™ serum-free hematopoietic cell medium (Lonza Bioscience) containing 5% (v/v) of fetal bovine serum, 50 µM of 2-Mercaptoethanol, 10 mM of N-Acetyl-L-(+)-cysteine, 10 U/mL of Penicillin-Streptomycin, in addition to 1X cytokines (200 U/mL of recombinant human interleukin-2, 5 µg/mL of recombinant human interleukin-7 and 5 µg/mL of recombinant human interleukin-15). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec). Cells were expanded in T cell basal media for 72 hours prior to LNP transfection.

8.2 T cell editing

[00558] Each RNA species, i.e. UGI mRNA, sgRNA or editor mRNA, was formulated separately in an LNP as described in Example 1. Editor mRNAs encoded either BC22n (SEQ ID NO: 5) or Cas9 (SEQ ID NO: 23). Guides targeting B2M (G015995), TRAC (G016017), TRBC1/2 (G016206) and CIITA (G018117 and G016086) were used either singly or in combination. Messenger RNA encoding UGI (SEQ ID NO: 26) is delivered in both Cas9 and BC22n arms of the experiment to normalize lipid amounts. Previous experiments have established UGI mRNA does not impact total editing or editing profile when used with Cas9 mRNA. LNPs were mixed to fixed total RNA weight ratios of 6:3:2 for editor mRNA, guide RNA, and UGI mRNA respectively as described in Table 26. In the 4-guide experiment described in Table 27, doses of LNPs with individual guides are decreased 4-fold to maintain the overall 6:3 editor mRNA: guide weight ratio and to allow comparison to individual guide potency based on total lipid delivery. LNP mixtures were incubated for 5 minutes at 37°C in T cell basal media substituting 6% cynomolgus monkey serum (Bioreclamation IVT, Cat. CYN220760) for fetal bovine serum.

[00559] Seventy-two hours post activation, T cells were washed and suspended in basal T cell media. Pre-incubated LNP mix was added to the each well with 1x10e5 T cells/well. T cells were incubated at 37°C with 5% C02 for the duration of the experiment. T cell media was changed 6 days and 8 days after activation and on tenth day post activation, cells were harvested for analysis by NGS and flow cytometry. NGS analysis was performed as described in Example 1.

[00560] Table 26 and Figs. 17A-E describe the editing profile of T cells when an individual guide was used for editing. Total editing and C to T editing showed direct, dose responsive relationships to increasing amounts of BC22n mRNA, UGI mRNA and guide across all guides tested. Indel and C conversions to A or G are in an inverse relationship with dose where lower doses resulted in a higher percentage of these mutations. In samples edited with Cas9, total editing and indel activity increase with the total RNA dose.

[00561] Table 26 Editing as a percent of total reads – single guide delivery

Guide	Editor	Total RNA (ng)	% C-to-T		% C-to-A/G		% Indel		N
			mean	SD	mean	SD	mean	SD	
G015995 B2M	BC22n	0.0	0.3	0.0	1.5	0.1	0.2	0.0	2
		8.6	49.5	3.5	7.7	0.6	6.0	0.4	2
		17.2	68.5	1.7	6.7	1.3	4.3	0.1	2
		34.4	79.0	0.9	5.7	0.3	3.8	0.0	2
		68.8	88.2	0.8	4.6	0.0	2.5	0.2	2
		137.5	90.6	1.8	4.1	0.4	2.2	0.5	2
		275.0	92.6	0.8	3.7	0.3	2.2	0.3	2
		550.0	95.2	0.4	2.8	0.0	1.6	0.2	2
	Cas9	0.0	0.3	0.0	1.5	0.2	0.2	0.0	2
		8.6	0.3	0.0	1.2	0.1	23.7	2.1	2
		17.2	0.3	0.0	0.9	0.1	41.1	0.2	2
		34.4	0.3	0.0	0.6	0.0	59.4	0.6	2
		68.8	0.2	0.1	0.4	0.0	76.8	1.2	2
		137.5	0.1	0.1	0.2	0.0	88.2	2.0	2
		275.0	0.1	0.0	0.1	0.1	95.1	0.5	2
		550.0	0.1	0.0	0.1	0.0	97.5	0.3	2
G016017 TRAC	BC22n	0.0	0.2	0.0	2.2	0.1	0.2	0.1	2
		8.6	34.6	1.1	5.6	0.8	6.6	0.2	2
		17.2	51.3	0.8	5.7	0.1	6.7	1.0	2
		34.4	66.9	2.6	5.4	0.2	4.7	0.4	2
		68.8	79.0	0.6	4.4	0.7	4.5	0.9	2
		137.5	89.2	0.4	3.6	0.9	2.5	0.2	2
		275.0	92.8	0.9	2.9	0.0	2.3	0.0	2
		550.0	94.5	1.3	3.4	1.0	1.6	0.2	2
	Cas9	0.0	0.2	0.0	2.3	0.1	0.1	0.0	2
		8.6	0.2	0.0	2.1	0.2	20.7	0.5	2
		17.2	0.1	0.0	1.4	0.0	34.6	0.7	2
		34.4	0.1	0.0	1.5	0.4	49.8	0.4	2
		68.8	0.1	0.0	1.0	0.0	62.3	0.1	2
		137.5	0.1	0.0	0.6	0.1	77.0	0.1	2
		275.0	0.0	0.0	0.3	0.0	87.8	0.2	2
		550.0	0.0	0.0	0.2	0.0	93.8	0.6	2

Guide	Editor	Total RNA (ng)	% C-to-T		% C-to-A/G		% Indel		N
			mean	SD	mean	SD	mean	SD	
G016206 TRBC1/2	BC22n	0.0	0.4	0.1	0.6	0.1	0.1	0.1	2
		8.6	23.7	1.3	6.1	0.0	6.1	0.8	2
		17.2	42.4	2.2	6.8	0.1	6.8	0.3	2
		34.4	60.1	2.2	5.7	0.3	5.9	0.7	2
		68.8	73.2	4.2	4.3	0.1	4.7	1.1	2
		137.5	81.7	0.8	3.6	0.2	3.7	0.4	2
		275.0	91.0	1.7	2.3	0.1	2.8	0.8	2
		550.0	93.6	1.9	2.0	0.2	1.7	0.6	2
	Cas9	0.0	0.3	0.0	0.5	0.0	0.1	0.0	1
		8.6	0.3	0.2	0.5	0.1	8.1	0.2	2
		17.2	0.3	0.1	0.7	0.1	14.9	0.6	2
		34.4	0.2	0.0	0.8	0.0	24.1	0.0	1
		68.8	0.2	0.0	0.4	0.0	35.9	0.0	1
		137.5	0.2	0.0	0.5	0.0	48.6	2.1	2
		275.0	0.1	0.0	0.4	0.0	63.8	0.0	1
		550.0	Not assayed						
G018117 CIITA	BC22n	0.0	0.3	0.0	2.7	0.1	0.3	0.0	2
		8.6	14.5	1.5	3.8	0.3	3.5	0.3	2
		17.2	28.1	0.6	3.5	0.3	3.9	1.0	2
		34.4	45.9	0.4	3.3	0.4	3.6	0.0	2
		68.8	62.8	5.3	3.6	0.1	3.7	1.2	2
		137.5	78.9	1.3	2.7	0.1	2.7	0.7	2
		275.0	86.3	1.8	2.6	0.1	2.0	0.1	2
		550.0	92.3	1.2	2.6	0.2	1.1	0.2	2
	Cas9	0.0	0.2	0.0	2.8	0.1	0.3	0.0	2
		8.6	0.3	0.0	2.5	0.0	6.0	0.2	2
		17.2	0.2	0.0	2.4	0.1	11.2	1.6	2
		34.4	0.2	0.0	2.1	0.0	20.8	0.3	2
		68.8	0.2	0.0	1.9	0.1	33.2	0.4	2
		137.5	0.1	0.0	1.3	0.1	51.2	0.0	2
		275.0	0.1	0.0	0.9	0.2	64.5	0.9	2
		550.0	0.1	0.0	0.6	0.0	78.4	1.1	2
G016086 CIITA	BC22n	0.0	0.2	0.0	1.0	0.1	0.1	0.0	2
		8.6	23.5	1.8	3.2	0.1	3.7	0.1	2
		17.2	40.9	1.1	4.4	0.7	4.6	1.0	2
		34.4	58.0	0.5	4.6	0.3	3.8	0.6	2
		68.8	73.5	0.7	3.7	0.0	2.8	0.5	2
		137.5	83.8	1.1	3.7	0.5	2.0	0.7	2
		275.0	90.1	2.4	3.1	0.1	1.9	0.8	2
		550.0	93.4	0.9	3.0	0.2	1.2	0.3	2
	Cas9	0.0	0.2	0.0	1.0	0.1	0.1	0.0	2
		8.6	0.2	0.0	1.1	0.2	7.4	0.7	2

Guide	Editor	Total RNA (ng)	% C-to-T		% C-to-A/G		% Indel		N
			mean	SD	mean	SD	mean	SD	
		17.2	0.2	0.0	1.1	0.3	17.7	1.0	2
		34.4	0.2	0.0	0.8	0.1	32.1	0.1	2
		68.8	0.2	0.0	0.7	0.2	51.5	0.8	2
		137.5	0.2	0.0	0.4	0.0	69.3	0.1	2
		275.0	0.3	0.1	0.3	0.1	84.2	0.1	2
		550.0	0.3	0.0	0.1	0.1	90.0	0.7	2

[00562] Table 27 and Figs. 18A-18D describe the editing profile for T cells in percent of total reads when four guides were used simultaneously for editing. In this arm of the experiment, each guide is used at 25% the concentration compared to the single guide editing experiment. In total, T cells were exposed to 6 different LNPs simultaneously (editor mRNA, UGI mRNA, 4 guides). Editing with BC22n and trans UGI lead to higher percentages of maximum total editing for each locus compared to editing with Cas9.

[00563] Table 27– Editing as a percentage of total reads – multiple guide delivery

Locus Assayed	Editor	Total RNA (ng)	% C-to-T		% C-to-A/G		% Indel		N
			mean	SD	mean	SD	Mean	SD	
G015995 B2M	BC22n	0.0	0.3	0.0	1.5	0.2	0.2	0.0	2
		8.6	27.3	0.2	3.8	0.1	2.6	0.1	2
		17.2	47.2	2.2	4.1	0.4	3.0	0.1	2
		34.4	61.2	3.0	3.9	0.1	2.6	0.3	2
		68.8	81.4	0.1	2.9	0.1	1.4	0.1	2
		137.5	90.0	1.1	2.6	0.3	1.3	0.5	2
		275.0	94.7	0.1	2.2	0.1	0.8	0.0	2
		550.0	95.9	0.9	2.9	1.0	0.4	0.3	2
	Cas9	0.0	0.3	0.0	1.4	0.1	0.2	0.0	2
		8.6	0.3	0.0	1.4	0.0	5.0	0.1	2
		17.2	0.3	0.0	1.3	0.0	10.5	0.4	2
		34.4	0.3	0.0	1.1	0.0	19.3	0.6	2
		68.8	0.3	0.0	0.9	0.0	34.4	0.1	2
		137.5	0.2	0.0	0.7	0.0	51.1	1.3	2
		275.0	0.2	0.1	0.5	0.0	68.0	0.1	2
		550.0	0.3	0.1	0.4	0.1	76.7	2.0	2
G016017 TRAC	BC22n	0.0	0.1	0.1	1.9	0.6	0.2	0.0	2
		8.6	12.1	1.3	4.3	0.2	2.4	0.2	2
		17.2	25.7	2.2	4.2	0.5	3.8	0.7	2
		34.4	44.7	1.4	4.7	1.0	3.0	0.3	2
		68.8	64.2	1.9	4.4	0.6	2.5	0.1	2
		137.5	79.3	1.1	3.6	0.4	2.1	0.1	2

Locus Assayed	Editor	Total RNA (ng)	% C-to-T		% C-to-A/G		% Indel		N
			mean	SD	mean	SD	Mean	SD	
		275.0	90.7	0.0	3.0	0.1	1.5	0.0	2
		550.0	93.3	0.6	2.4	0.1	0.9	0.4	2
	Cas9	0.0	0.1	0.1	2.1	0.2	0.1	0.0	2
		8.6	0.2	0.1	2.3	0.2	6.1	0.2	2
		17.2	0.1	0.0	1.8	0.2	11.5	0.5	2
		34.4	0.1	0.0	2.0	0.4	21.0	0.4	2
		68.8	0.1	0.0	1.4	0.0	33.5	0.1	2
		137.5	0.1	0.0	1.2	0.1	47.5	0.5	2
		275.0	0.1	0.0	0.9	0.1	64.8	0.2	2
		550.0	0.1	0.0	0.6	0.1	76.1	1.3	2
G016206 TRBC1/2	BC22n	0.0	No data						
		8.6	11.6	0.3	2.6	0.2	2.8	0.3	2
		17.2	23.4	0.4	3.6	0.3	2.6	0.5	2
		34.4	38.5	1.4	3.7	0.2	2.9	0.7	2
		68.8	55.6	1.7	2.3	0.4	2.4	0.0	2
		137.5	72.4	1.2	1.8	0.5	1.7	0.5	2
		275.0	85.1	1.0	1.9	0.5	1.7	0.6	2
		550.0	89.8	2.8	2.2	0.1	0.9	0.3	2
	Cas9	0.0	0.2	0.0	0.6	0.0	0.1	0.0	1
		8.6	0.2	0.1	0.7	0.1	2.3	0.3	2
		17.2	0.3	0.0	0.7	0.3	4.2	0.4	2
		34.4	0.1	0.0	0.5	0.1	6.6	0.5	2
		68.8	0.4	0.0	0.5	0.0	12.3	0.0	1
		137.5	0.2	0.0	0.5	0.0	17.8	0.0	1
		275.0	0.1	0.0	0.5	0.0	33.0	0.0	1
		550.0	0.3	0.2	0.3	0.0	43.3	1.7	2
G018117 CIITA	BC22n	0.0	0.2	0.0	2.6	0.1	0.3	0.0	2
		8.6	4.6	0.9	3.1	0.2	0.8	0.2	2
		17.2	10.5	0.2	2.9	0.1	1.1	0.2	2
		34.4	18.8	0.3	2.9	0.2	1.6	0.2	2
		68.8	35.1	0.6	2.7	0.2	1.6	0.7	2
		137.5	52.9	0.2	2.9	0.3	1.5	0.0	2
		275.0	71.9	2.4	2.5	0.3	1.3	0.1	2
		550.0	81.1	1.9	2.6	0.1	1.1	0.6	2
	Cas9	0.0	0.3	0.0	2.7	0.1	0.3	0.0	2
		8.6	0.2	0.0	2.6	0.2	1.4	0.0	2
		17.2	0.2	0.0	2.5	0.0	2.1	0.3	2
		34.4	0.3	0.0	2.5	0.0	3.9	0.1	2
		68.8	0.2	0.0	2.5	0.2	7.7	0.6	2
		137.5	0.2	0.0	2.2	0.1	13.3	0.2	2
		275.0	0.1	0.0	1.9	0.0	26.7	1.3	2
		550.0	0.1	0.0	1.7	0.1	42.3	0.3	2

[00564] On day 10 post-activation, T cells were phenotyped by flow cytometry to determine if editing resulted in loss of cell surface proteins. Briefly, T cells were incubated in a mix of the following antibodies: B2M-FITC (BioLegend, Cat. 316304), CD3-AF700 (BioLegend, Cat. 317322), HLA DR DQ DP-PE (BioLegend, Cat 361704) and DAPI (BioLegend, Cat 422801). A subset of unedited cells was incubated with Isotype Control-PE (BioLegend® Cat. No. 400234). Cells were subsequently washed, processed on a Cytoflex instrument (Beckman Coulter), and analyzed using the FlowJo software package. T cells were gated based on size, shape, viability, and antigen expression.

[00565] Table 28 and Figs. 19A-19I report phenotyping results as percent of cells negative for antibody binding. The percentage of antigen negative cells increased in a dose responsive manner with increasing total RNA for both BC22n and Cas9 samples. Cells edited with BC22n showed comparable or higher protein knockout compared to cells edited with Cas9 for all guides tested. In multi-edited cells, BC22n with trans UGI showed substantially higher percentages of antigen negative cells than Cas9 with trans UGI. For example, BC22n edited samples at the highest total RNA dose of 550 ng showed 84.2% of cells lacking all three antigens, while Cas9 editing led to only 46.8% such triple knockout cells. For samples treated with one guide only, the correlation between DNA editing and antigen reduction was robust. BC22n had an R square measurement of 0.93 when comparing C to T conversions to antigen knockout. Cas9 had an R square measurement of 0.95 when comparing indels to antigen knockout.

[00566] Table 28 – Flow cytometry data – percent cells negative for antigen (n=2)

Guide(s)	Phenotype	Total RNA (ng)	BC22n		Cas9	
			Mean %	SD	Mean %	SD
G015995 B2M	B2M neg	550.0	95.7	0.1	91.3	0.6
		275.0	94.4	0.4	89.3	0.1
		137.5	91.2	0.1	82.1	3.3
		68.8	83.9	0.4	68.7	3.3
		34.4	75.7	1.4	53.4	0.2
		17.2	60.8	2.0	30.7	1.3
		8.6	44.0	2.3	13.9	2.0
		0.0	14.1	4.1	9.9	1.9
G015995 G016017 G016206 G018117	B2M neg	550.0	94.4	0.1	74.2	0.4
		275.0	91.3	0.1	65.2	0.1
		137.5	84.3	0.2	45.4	1.9
		68.8	72.7	0.4	24.5	0.8

		34.4	56.2	1.2	14.1	2.3
		17.2	38.5	0.2	9.9	0.8
		8.6	20.6	0.7	7.6	2.4
		0.0	14.1	4.1	9.9	1.9
G016017 TRAC	CD3 neg	550.0	97.3	0.3	94.8	0.4
		275.0	96.0	0.2	87.0	4.9
		137.5	91.9	0.2	72.7	0.9
		68.8	85.7	0.5	65.6	0.1
		34.4	76.6	0.8	51.7	3.0
		17.2	61.8	1.7	35.7	1.1
		8.6	42.1	0.7	20.1	1.5
		0.0	1.0	0.1	0.9	0.1
G016206 TRBC1/2	CD3 neg	550.0	97.9	0.1	86.6	0.3
		275.0	96.0	0.1	77.3	0.1
		137.5	90.4	0.8	59.4	0.4
		68.8	82.9	0.1	40.6	1.2
		34.4	71.9	1.5	27.0	1.6
		17.2	53.4	0.3	16.1	0.1
		8.6	32.6	0.6	7.9	0.4
		0.0	0.8	0.0	0.9	0.4
G015995 G016017 G016206 G018117	CD3 neg	550.0	98.3	0.2	84.2	0.1
		275.0	96.3	0.1	74.6	0.5
		137.5	90.4	0.3	57.4	1.0
		68.8	81.3	0.3	39.4	0.1
		34.4	66.3	1.6	25.6	0.8
		17.2	48.2	1.0	15.3	0.5
		8.6	27.3	0.7	8.6	0.5
		0.0	0.9	0.1	0.9	0.2
G018117 CIITA	MHC II neg	550.0	95.7	0.4	72.0	0.1
		275.0	92.5	1.1	65.6	0.4
		137.5	85.2	0.6	55.5	0.6
		68.8	74.5	1.1	48.9	0.0
		34.4	65.8	3.7	40.7	0.6
		17.2	49.9	0.1	36.2	0.6
		8.6	41.6	0.8	34.2	1.3
		0.0	30.1	1.6	35.2	0.4
G015995 G016017 G016206 G018117	MHC II neg	550.0	88.0	0.2	52.8	1.1
		275.0	81.2	0.2	46.4	0.4
		137.5	70.4	1.3	39.9	1.8
		68.8	60.0	0.4	39.1	3.3
		34.4	48.8	0.6	37.7	2.9
		17.2	43.0	4.2	37.5	0.6
		8.6	37.8	2.1	35.0	0.0
		0.0	33.0	1.9	37.3	2.1
G015995	B2M neg	550.0	84.2	0.0	46.8	1.1

G016017 G016206 G018117	CD3 neg MHC II neg	275.0	76.2	0.0	37.8	0.2
		137.5	63.0	1.3	23.4	2.4
		68.8	48.2	0.2	10.8	0.9
		34.4	31.5	1.1	3.6	0.9
		17.2	17.8	1.7	1.1	0.2
		8.6	6.4	0.0	0.4	0.1
		0.0	0.1	0.0	0.1	0.0
G016086 CIITA	HLA DR DP DQ neg	550.0	96.0	0.1	90.9	0.7
		275.0	93.7	0.1	87.4	0.3
		137.5	88.4	0.5	76.3	0.6
		68.8	80.0	0.7	66.1	1.8
		34.4	69.2	1.5	53.4	1.1
		17.2	56.4	0.4	41.9	0.8
		8.6	45.2	2.9	37.3	0.1
		0.0	30.1	0.9	36.8	0.4

Example 9 – Editing in T cells with trans UGI titration

[00567] Editing profiles were assessed to determine the concentration of UGI mRNA necessary to achieve highly pure C-to-T editing in T cells with different editing constructs. C-to-T editing purity (% of edited reads containing only C-to-T conversions) was measured using a saturating dose of editor mRNA and sgRNA and varying amounts of UGI mRNA. Editor mRNAs included an mRNA encoding BC22n (SEQ ID NO: 5), an mRNA encoding BC22 with a total of 2 fused UGI moieties (SEQ ID NO: 29), and an mRNA encoding BE4Max which includes 2 fused UGI moieties (SEQ ID NO: 32) (Koblan LW, Doman JL, Wilson C, et al. *Nat Biotechnol*. 2018;36(9):843-846.).

[00568] Solutions containing editor mRNA as listed in Table 29 or UGI mRNA (SEQ ID NO: 26) were prepared in sterile water. 50 μM B2M targeting sgRNA G015995 were removed from their storage plates and denatured for 2 minutes at 95°C before cooling on ice. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of editor mRNAs, 20 pmols of sgRNA and different concentrations of UGI mRNA ranging from 0.8 ng to 600.0 ng (0.8 nM to 597.0 nM), in a final volume of 20 uL of P3 electroporation buffer. This mRNA + sgRNA + T cell mix was transferred in triplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in CTS™ OpTmizer™ T Cell Expansion serum-free media (SFM) (ThermoFisher Cat. A1048501) without cytokines for 10 minutes before being

transferred to a new flat-bottom 96-well plate containing an additional 100 uL of the same media with 2X concentration of cytokines. The resulting plate was incubated at 37°C for 9 days, during which time the cells were split and media refreshed two times. T cells were collected and processed for NGS sequencing as described in Example 1.

[00569] Table 29 and Figs. 20A-20C show the percentage of reads with each edit type. Increasing levels of UGI mRNA in trans decreased indels for all three editor constructs. With 199 nM UGI mRNA or higher, indels were decreased to less than 2% of total reads for each editor construct tested. At low concentrations of UGI mRNA in trans, editor constructs that encoded UGI showed more C to T conversions than the BC22n construct that lacked an encoded UGI. Table 30 and Fig. 21 represents the data captured in Table 29 and Figs. 20A-20C as percentage of total edits that were C to T conversions, also known as the C to T purity. Increasing UGI mRNA in trans increased C to T total editing and increased C to T purity for all constructs. All three constructs showed over 95% C to T editing purity with 199 nM UGI mRNA, which corresponds to a molar ratio of UGI mRNA to editor mRNA of about 7:1.

[00570] Table 29 – Editing type as a percentage of total reads with increasing UGI mRNA

		% C-to-T		% C-to-A/G		% Indel	
Editor	**nM UGI mRNA**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**
BC22n	No UGI	19.9	1.1	31.8	1.4	47.4	0.4
	0.82	23.2	0.6	29.7	1.6	45.9	2.8
	2.46	32.0	0.2	24.5	2.3	42.6	3.3
	7.37	45.0	1.1	23.6	1.6	30.4	2.1
	22.11	59.8	1.0	18.1	0.1	20.6	0.4
	66.33	82.7	2.6	9.3	1.7	7.2	1.2
	199	93.8	1.4	3.2	0.4	1.5	0.3
	597	95.7	0.1	1.9	0.2	1.3	0.4
BC22-2xUGI	No UGI	54.2	0.4	19.2	2.3	21.0	2.1
	0.82	58.6	0.1	18.1	1.9	21.0	2.3
	2.46	61.7	2.0	18.0	2.9	17.5	1.8
	7.37	68.5	0.1	14.5	1.1	15.1	0.8
	22.11	75.7	1.1	12.1	2.3	10.1	2.1
	66.33	88.4	1.2	5.4	0.6	4.7	1.1
	199	94.6	0.1	2.1	0.1	1.5	0.1
	597	96.9	0.6	1.4	0.1	0.6	0.1
BE4Max	No UGI	54.7	1.3	17.2	0.3	9.1	0.1
	0.82	60.3	0.5	15.8	0.5	7.3	0.7

		% C-to-T		% C-to-A/G		% Indel	
Editor	**nM UGI mRNA**	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**
	2.46	62.4	0.1	16.2	0.2	6.7	0.6
	7.37	65.2	0.6	15.1	2.0	5.8	0.4
	22.11	71.8	0.4	10.0	0.2	4.2	0.6
	66.33	78.6	0.6	4.9	1.1	2.1	0.4
	199	83.8	1.0	2.3	0.8	0.5	0.1
	597	85.9	1.6	1.2	0.1	0.4	0.1

[00571] Table 30 – C-to-T purity: Percent of edited sequencing reads containing only C-to-T conversions, no indels, no C to A/G conversions (n=2)

nM UGI mRNA	BC22n		BC22-2XUGI		BE4Max	
	mean	SD	mean	SD	mean	SD
597	96.8	0.7	98.0	0.1	98.2	0.1
199	95.3	0.7	96.4	0.1	96.8	1.1
66.33	83.4	2.9	89.8	1.6	91.9	0.9
22.11	60.8	0.6	77.4	0.4	83.6	0.8
7.37	45.4	0.8	69.9	0.2	75.8	1.6
2.46	32.3	0.5	63.5	1.4	73.2	0.7
0.82	23.5	0.8	60.0	0.1	72.4	0.1
0	19.1	0.4	57.4	0.1	67.5	0.8

Example 10. Screening of CIITA Guide RNAs

[00572] CIITA gRNAs were screened for efficacy in T cells by assessing knockdown of MHC class II cell surface expression using both Cas9 and BC22. The percentage of T cells negative for MHC class II protein was assayed following CIITA editing by electroporation with RNP.

10.1. RNP electroporation of T cells

[00573] Cas9 editing activity was assessed using electroporation of Cas9 ribonucleoprotein (RNP). Upon thaw, Pan CD3+ T cells (StemCell, HLA-A*02.01/ A*03.01) were plated at a density of 0.5 x 10^6 cells/mL in T cell RPMI media composed of RPMI 1640 (Invitrogen, Cat. 22400-089) containing 5% (v/v) of fetal bovine serum, 1x

Gluatmax (Gibco, Cat. 35050-061), 50 μM of 2-Mercaptoethanol, 100 uM non-essential amino acids (Invitrogen, Cat. 11140-050), 1 mM sodium pyruvate, 10 mM HEPES buffer, 1% of Penicillin-Streptomycin, and 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec). Cells were expanded in T cell RPMI media for 72 hours prior to RNP transfection.

[00574] CIITA targeting sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C before cooling at room temperature for 10 minutes. RNP mixture of 20 uM sgRNA and 10 uM recombinant Cas9-NLS protein (SEQ ID No. 36) was prepared and incubated at 25°C for 10 minutes. Five μL of RNP mixture was combined with 100,000 cells in 20 μL P3 electroporation Buffer (Lonza). 22 μL of RNP/cell mix was transferred to the corresponding wells of a Lonza shuttle 96-well electroporation plate. Cells were electroporated in duplicate with the manufacturer's pulse code. T cell RPMI media was added to the cells immediately post electroporation. Electroporated T cells were subsequently cultured and collected for NGS sequencing as described in Example 1 at 2 days post edit.

10.2 Flow cytometry

[00575] On day 10 post-edit, T cells were phenotyped by flow cytometry to determine HLA class II protein expression. Briefly, T cells were incubated in cocktails of antibodies targeting HLA-DR, DQ, DP-PE (BioLegend® Cat. No. 361704) and Isotype Control-AF647 (BioLegend® Cat. No. 400234). Cells were subsequently washed, processed on a Cytoflex flow cytometer (Beckman Coulter) and analyzed using the FlowJo software package. T cells were gated based on size, shape, viability, and MHC II expression. DNA samples were subjected to PCR and subsequent NGS analysis, as described in Example 1.

10.3. mRNA electroporation of T cells

[00576] BC22 editing activity was assayed following CIITA editing by electroporation with mRNA and guide. Upon thaw, Pan CD3+ T cells isolated from a commercially obtained leukopak (StemCell) were plated at a density of 0.5 x 10^6 cells/mL in T cell R10 media composed of RPMI 1640 (Invitrogen, Cat. 22400-089) containing 10% (v/v) of fetal bovine serum, 2 mM Glutamax (Gibco, Cat. 35050-061), 22 μM of 2-Mercaptoethanol, 100 uM non-essential amino acids (Invitrogen, Cat. 11140-050), 1 mM sodium pyruvate, 10mM HEPES buffer, 1% of Penicillin-Streptomycin, plus100 U/mL of

recombinant human interleukin-2 (Peprotech, Cat. 200-02). T cells were activated with Dynabeads® Human T-Activator CD3/CD28 (ThermoFisher). Cells were expanded in T cell for 72 hours prior to mRNA transfection.

[00577] CIITA sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C before cooling at room temperature for 10 minutes. Fifty microliter electroporation mix was prepared with 100,00 T cells in P3 buffer (Lonza) and 10 ng/uL mRNA encoding UGI (SEQ ID NO: 26), 10 ng/uL mRNA encoding BC22 (SEQ ID NO: 20) and 2 uM sgRNA. This mix was transferred to the corresponding wells of a Lonza shuttle 96-well electroporation plate. Cells were electroporated in duplicate wells using Lonza shuttle 96w program with the manufacturer's pulse code. R10 media with IL-2 was added to the cells immediately post electroporation. Electroporated T cells were subsequently cultured and collected for NGS sequencing and flow cytometry 10 days post edit. Flow cytometry was performed as described for Cas9 RNP treated cells above. DNA samples were subjected to PCR and subsequent NGS analysis, as described in Example 1.

[00578] Table 31 and Figure 22 show mean percentage of T cells negative for cell surface expression of MHC Class II proteins HLA-DR, DQ, DP using Cas9 and BC22 in relation to the distance from the cut site to the splice site boundary nucleotide. For each guide, the genomic coordinate of the cut site with SpCas9 is shown, as well as the distance (# of nucleotides) between the acceptor splice site boundary nucleotide or the donor splice site boundary nucleotide and the cut site. Positive numerical values show the number of nucleotides in the 5'directed between a splice site boundary nucleotide and cut site, whereas the negative numerical values show the number of nucleotides in the 3' direction between a splice site boundary nucleotide and cut site.

[00579] **Table 31: Guide position vs. protein knockdown efficiency**

Guide	Distance from Cut Site (bp)	BC22		Cas9	
		Mean % MHC II negative	SD	Mean % MHC II negative	SD
G018075	11	88.3	1.56	54.7	3.54
G018076	10	88.45	0.64	51.45	6.29
G018077	9	69.15	2.48	55.9	0.14
G018078	4	71.35	4.17	83.65	3.04

G018081	0	37.9	1.27	91.1	0.42
G018082	-1	36.75	1.91	93.6	2.26
G018084	-7	45.25	3.75	44	3.39
G018085	-8	45.75	2.19	32.95	0.92

Example 11 – Screening of HLA-A Guides with BC22n and Cas9

[00580] HLA-A guide RNAs were screened for efficacy in T cells by assessing loss of HLA-A cell surface expression. The percentage of T cells negative for HLA-A protein in an HLA-A2 background ("% HLA-A2-") was assayed by flow cytometry following HLA-A editing by mRNA delivery.

11.1. mRNA electroporation of T cells

[00581] Cas9 and BC22n editing activity was assessed using electroporation of mRNA encoding Cas9 (SEQ ID NO: 11), mRNA encoding BC22n (SEQ ID NO: 2), or mRNA encoding UGI (SEQ ID NO: 26), as provided below. Upon thaw, Pan CD3+ T cells (StemCell, HLA-A*02.01/ A*02.01) were plated at a density of 1x 10^6 cells/mL in TCGM composed of CTS OpTmizer T Cell Expansion SFM (Thermofisher, Cat. A3705001) supplemented with 5% human AB serum (Gemini, Cat. 100-512), 1X GlutaMAX (Thermofisher, Cat.35050061), 10 mM HEPES (Thermofisher, Cat. 15630080), 1x of Penicillin-Streptomycin, further supplemented with 200 U/mL IL-2 (Peprotech, Cat. 200-02), 10 ng/ml IL-7 (Peprotech, Cat. 200-07), 10 ng/ml IL-15 (Peprotech, Cat. 200-15). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec). Cells were expanded in T cell RPMI media for 72 hours at 37°C prior to mRNA electroporation.

[00582] HLA-A sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C before incubating at room temperature for 5 minutes. BC22n electroporation mix was prepared with 100,000 T cells in P3 buffer (Lonza), 200 ng of mRNA encoding UGI, 200 ng of mRNA encoding BC22n and 20 pmoles of sgRNA. Cas9 electroporation mix was prepared with 100,000 T cells in P3 buffer (Lonza), 200 ng of mRNA encoding UGI, 200 ng ofmRNA encoding Cas9 and 20 pmoles of sgRNA. Each mix was transferred to the corresponding wells of a Lonza shuttle 96-well electroporation plate. Cells were electroporated in duplicate using Lonza shuttle 96w using manufacturer's pulse code. Immediately post electroporation, cells were recovered in pre-warmed TCGM without

cytokines and incubated at 37°C for 15 minutes. Electroporated T cells were subsequently cultured in TCGM with further supplemented with 200 U/mL IL-2 (Peprotech, Cat. 200-02), 10 ng/ml IL-7 (Peprotech, Cat. 200-07), 10 ng/ml IL-15 (Peprotech, Cat. 200-15) and collected for flow cytometry 8 days post edit.

11.2. Flow cytometry

[00583] On day 8 post-edit, T cells were phenotyped by flow cytometry to determine HLA-A protein expression. Briefly, T cells were incubated with antibodies targeting HLA-A2, (eBioscience Cat. No. 17-9876-42). Cells were subsequently washed, processed on a Cytoflex flow cytometer (Beckman Coulter) and analyzed using the FlowJo software package. T cells were gated based on size, shape, viability, and HLA-A2 expression. **Table 32** shows the percentage of cells negative for HLA-A surface proteins following genomic editing of *HLA-A* with BC22n or Cas9.

[00584] **Table 32 – Percentage of cells negative for HLA-A surface protein following genomic editing of *HLA-A* with BC22n or Cas9.**

	BC22n		Cas9	
Guide ID	**Mean %A2-**	**SD % A2-**	**Mean %A2-**	**SD % A2-**
G018932	20.15	2.76	43.30	1.70
G018933	10.35	1.20	74.00	0.57
G018934	0.50	0.14	15.30	1.56
G018935	0.00	0.00	69.30	0.28
G018936	0.10	0.00	29.65	2.62
G018937	0.15	0.07	50.50	0.71
G018938	0.00	0.00	0.00	0.00
G018939	0.00	0.00	44.90	1.27
G018940	0.00	0.00	12.00	0.42
G018941	0.00	0.00	2.65	0.35
G018942	0.10	0.00	2.15	0.07
G018943	0.00	0.00	16.20	0.42
G018944	0.00	0.00	3.00	0.28
G018945	0.05	0.07	3.20	0.42

	BC22n		Cas9	
Guide ID	**Mean %A2-**	**SD % A2-**	**Mean %A2-**	**SD % A2-**
G018946	0.00	0.00	2.30	0.14
G018947	0.00	0.00	1.55	0.49
G018949	0.00	0.00	47.10	0.57
G018950	0.00	0.00	0.30	0.00
G018951	0.00	0.00	13.30	0.28
G018952	0.00	0.00	0.50	0.00
G018953	0.00	0.00	3.65	0.64
G018955	0.20	0.14	5.20	0.28
G018958	0.00	0.00	1.30	0.28
G018959	0.00	0.00	3.70	0.14
G018960	0.00	0.00	0.35	0.07
G018961	0.00	0.00	0.40	0.00
G018962	0.00	0.00	2.90	0.42
G018963	0.00	0.00	12.50	0.14
G018964	0.00	0.00	6.45	0.64
G018965	0.00	0.00	0.90	0.00
G018966	0.00	0.00	1.30	0.14
G018968	0.10	0.00	0.10	0.00
G018969	0.00	0.00	0.80	0.14
G018970	0.00	0.00	0.95	0.07
G018971	0.00	0.00	0.10	0.00
G018972	0.05	0.07	3.40	0.28
G018973	0.00	0.00	1.35	0.07
G018974	0.00	0.00	0.45	0.07
G018976	0.05	0.07	2.45	0.07
G018977	0.00	0.00	12.45	1.06
G018978	0.00	0.00	1.75	0.07
G018979	0.05	0.07	37.40	0.71
G018980	0.05	0.07	32.40	2.40

	BC22n		Cas9	
Guide ID	**Mean %A2-**	**SD % A2-**	**Mean %A2-**	**SD % A2-**
G018981	0.00	0.00	17.45	0.35
G018982	0.00	0.00	26.35	0.92
G018983	0.00	0.00	0.25	0.07
G018984	0.00	0.00	0.65	0.07
G018986	0.00	0.00	1.85	0.21
G018987	0.00	0.00	2.25	0.07
G018988	0.00	0.00	0.15	0.07
G018989	0.00	0.00	1.85	0.07
G018990	0.25	0.07	17.45	1.06
G018991	0.20	0.00	23.15	0.92
G018992	0.20	0.14	38.15	0.07
G018993	0.15	0.07	12.15	1.34
G018994	4.35	0.35	23.75	0.49
G018995	0.55	0.07	94.27	0.30
G018996	0.85	0.07	92.39	0.83
G018997	97.80	0.08	95.03	1.87
G018998	74.75	7.71	93.33	0.18
G018999	98.26	0.30	96.05	2.27
G019000	9.05	0.35	94.67	0.74
G019001	0.05	0.07	4.05	0.64
G019002	0.00	0.00	0.05	0.07
G019003	0.00	0.00	11.10	0.00
G019004	0.00	0.00	30.70	0.00
G019005	0.00	0.00	1.65	0.35
G019006	0.00	0.00	4.75	0.49
G019007	0.00	0.00	5.35	0.78
G019008	0.00	0.00	55.20	3.54
G019009	0.00	0.00	19.55	2.19
G019010	0.05	0.07	5.40	0.14

	BC22n		Cas9	
Guide ID	**Mean %A2-**	**SD % A2-**	**Mean %A2-**	**SD % A2-**
G019011	0.00	0.00	4.40	0.85
G019012	0.05	0.07	22.90	2.55
G019013	0.00	0.00	30.60	2.40
G019014	0.05	0.07	14.65	0.49
G019015	0.00	0.00	44.70	1.70
G019016	0.00	0.00	13.95	0.35
G019017	0.00	0.00	2.35	0.35
G019018	0.00	0.00	19.90	0.00
G019019	0.00	0.00	3.20	0.14
G021205	0.00	0.00	0.00	0.00
G021206	0.00	0.00	4.10	0.28
G021207	0.00	0.00	2.80	0.28
G021208	84.75	2.05	58.50	0.28
G021209	97.96	0.16	83.35	1.77
G021210	71.45	2.90	75.20	1.70
G021211	0.10	0.00	67.80	1.70

Example 12 – T cell editing, CIITA guide RNAs with Cas9 and BC22

12.1 T cell Preparation

[00585] T cells were edited at the CIITA locus with UGI in trans and either BC22 or Cas9 to assess the impact on editing type on MHC class II antigens.

[00586] T cells were prepared from a leukopak using the EasySep Human T cell Isolation Kit (Stem Cell Technology, Cat. 17951) following the manufacturers protocol. T cells were cryopreserved in Cryostor CS10 freezing media (Cat. 07930) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell R10 media composed of RPMI 1640 (Corning, Cat. 10-040-CV) containing 10% (v/v) of fetal bovine serum, 2 mM Glutamax (Gibco, Cat. 35050-061), 22 μM of 2-Mercaptoethanol, 100 uM non-essential amino acids (Corning, Cat. 25-025-Cl), 1 mM sodium pyruvate, 10 mM HEPES buffer, 1% of Penicillin-Streptomycin, plus 100 U/mL of recombinant human interleukin-2

(Peprotech, Cat. 200-02). T cells were activated with Dynabeads® Human T-Activator CD3/CD28 (Gibco, Cat. 11141D). Cells were expanded in T cell media for 72 hours prior to mRNA transfection.

12.2 T cell editing with RNA electroporation

[00587] Solutions containing mRNA encoding Cas9 protein (SEQ ID NO: 8), BC22 (SEQ ID NO: 20) or UGI (SEQ ID NO 26:) were prepared in sterile water. 50 μM CIITA targeting sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C before cooling on ice. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of editor mRNA, 200 ng of UGI mRNA and 20 pmols of sgRNA as described in **Table 33** in a final volume of 20 uL of P3 electroporation buffer. This mix was transferred in duplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were rested in 180 ul of R10 media plus 100 U/mL of recombinant human interleukin-2 before being transferred to a new flat-bottom 96-well plate. The resulting plate was incubated at 37°C for 4 days. On day 10 post-editing cells were collected for flow cytometry analysis and NGS sequencing.

12.3 Flow cytometry and NGS sequencing

[00588] On day 10 post-editing, T cells were phenotyped by flow cytometry to determine MHC class II protein expression as described in **Example 6** using antibodies targeting HLA-DR, DQ, DP-PE (BioLegend® Cat. No. 361704) and Isotype Control-PE (BioLegend® Cat. No. 400234). DNA samples were subjected to PCR and subsequent NGS analysis, as described in Example 1. **Table 33** shows CIITA gene editing and MHC class II negative results for cells edited with BC22. **Table 34** shows CIITA gene editing and MHC class II negative results for cells edited with Cas9.

[00589] Table 33 – Percent editing and percent of MHC-II negative cells following *CIITA* editing with BC22

	% C to T			% A to G			% Indel			% MHC Class II negative		
Guide	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**
G016030	40.2	14.2	2	3.5	1.1	2	2.4	1.1	2	39.7	5.6	2
G016031	58.8	0.0	1	3.4	3.2	2	0.6	0.8	2	41.5	3.0	2
G016032	1.8	0.6	2	18.2	1.3	2	75.2	0.6	2	45.9	4.2	2

Guide	% C to T			% A to G			% Indel			% MHC Class II negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
G016033	1.6	0.1	2	30.7	2.2	2	18.0	5.2	2	38.5	1.5	2
G016034	52.8	14.8	2	1.5	0.5	2	2.0	0.3	2	41.8	2.2	2
G016035	50.3	14.5	2	1.6	0.6	2	1.4	0.4	2	40.8	2.5	2
G016036	14.2	4.9	2	2.2	0.3	2	2.4	0.3	2	40.1	2.4	2
G016037	10.1	4.5	2	1.3	0.4	2	0.3	0.1	2	38.2	0.8	2
G016038	71.3	6.5	2	3.2	0.1	2	3.0	0.6	2	45.6	2.6	2
G016039	66.0	5.9	2	5.0	0.1	2	10.5	2.9	2	38.6	0.5	2
G016040	No data			0.0	0.0	1	0.0	0.0	1	38.4	0.7	2
G016041	No data			3.1	4.3	2	3.3	1.2	2	40.4	2.1	2
G016042	21.5	7.9	2	3.2	0.1	2	1.6	0.9	2	44.4	3.7	2
G016043	44.7	11.5	2	2.3	0.1	2	3.2	0.5	2	40.5	1.6	2
G016044	93.4	2.3	2	1.8	0.4	2	4.9	0.7	2	39.9	0.9	2
G016045	7.1	3.0	2	1.2	0.1	2	2.2	0.2	2	39.8	0.7	2
G016046	63.7	11.5	2	2.9	0.1	2	3.3	0.1	2	46.4	1.4	2
G016047	72.7	3.0	2	2.9	0.2	2	4.6	0.9	2	45.4	2.0	2
G016048	6.2	2.4	2	2.4	0.1	2	0.2	0.1	2	38.4	0.5	2
G016049	66.8	9.0	2	5.4	0.1	2	2.8	0.1	2	42.4	2.5	2
G016050	45.4	9.0	2	2.3	0.3	2	3.0	0.6	2	39.1	2.8	2
G016051	86.2	5.9	2	3.7	0.1	2	1.3	0.3	2	39.5	0.6	2
G016052	53.7	13.7	2	4.2	0.9	2	7.4	0.4	2	42.1	1.8	2
G016053	38.6	18.7	2	1.3	0.3	2	3.1	0.8	2	40.4	1.7	2
G016054	42.0	10.7	2	1.9	0.1	2	2.1	0.4	2	42.1	4.2	2
G016055	36.6	13.1	2	3.5	0.9	2	8.2	2.3	2	41.3	1.5	2
G016056	78.0	9.1	2	3.4	0.1	2	2.0	0.1	2	39.8	3.4	2
G016057	73.3	9.1	2	3.4	0.6	2	5.3	0.0	2	39.7	1.3	2
G016058	75.0	9.1	2	1.7	0.1	2	4.2	0.0	2	46.0	2.3	2
G016059	66.5	12.0	2	4.3	0.2	2	4.7	0.7	2	41.0	0.8	2
G016060	55.6	5.2	2	3.5	0.4	2	10.7	1.0	2	44.2	1.6	2
G016061	65.5	9.3	2	2.5	0.4	2	1.2	0.0	2	38.6	2.0	2
G016062	65.8	9.1	2	3.0	0.2	2	4.3	0.3	2	39.8	0.1	2
G016063	10.2	4.2	2	0.9	0.2	2	0.3	0.1	2	39.9	2.5	2
G016064	66.8	12.2	2	3.5	0.0	2	4.4	1.1	2	59.2	2.2	2
G016065	13.5	5.5	2	0.6	0.2	2	1.1	0.4	2	40.4	2.5	2
G016066	0.1	0.0	2	0.7	0.1	2	0.3	0.0	2	35.4	1.5	2
G016067	82.7	6.4	2	2.6	0.6	2	4.6	0.2	2	59.8	1.2	2
G016068	60.1	8.9	2	2.5	0.8	2	1.1	0.1	2	54.0	0.7	2
G016069	55.8	11.6	2	2.6	0.6	2	2.6	1.1	2	34.3	4.0	2
G016070	76.5	9.6	2	3.5	1.0	2	4.4	0.3	2	53.3	1.7	2
G016071	54.3	6.8	2	4.1	0.5	2	1.4	0.6	2	45.8	0.1	2
G016072	82.0	0.0	1	4.1	0.0	1	5.1	0.0	1	25.1	1.1	2

Guide	% C to T			% A to G			% Indel			% MHC Class II negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
G016073	63.4	11.1	2	3.5	0.5	2	0.9	0.0	2	38.7	1.3	2
G016074	62.7	13.0	2	3.9	0.2	2	4.2	0.7	2	53.3	2.0	2
G016075	41.2	17.2	2	1.0	0.4	2	8.0	1.7	2	48.1	0.7	2
G016076	42.8	14.3	2	0.9	0.1	2	10.2	2.8	2	55.6	2.9	2
G016077	65.1	10.6	2	5.4	0.2	2	2.4	0.1	2	46.0	0.2	2
G016078	44.1	15.1	2	2.0	0.6	2	6.6	2.0	2	41.6	0.1	2
G016079	79.8	4.9	2	5.6	0.5	2	4.6	0.9	2	53.4	2.8	2
G016080	39.0	12.2	2	4.3	0.6	2	13.1	3.1	2	49.5	3.1	2
G016081	9.6	4.9	2	0.5	0.3	2	2.5	0.8	2	39.0	0.3	2
G016082	20.3	8.6	2	2.0	0.3	2	7.0	3.3	2	40.1	4.1	2
G016083	74.5	9.8	2	3.6	0.1	2	8.0	0.6	2	48.6	2.3	2
G016084	46.0	8.5	2	3.9	0.0	2	23.5	3.2	2	54.5	0.7	2
G016085	35.6	6.7	2	1.4	0.1	2	1.6	0.1	2	41.9	3.0	2
G016086	75.0	10.3	2	3.9	0.4	2	3.5	0.2	2	63.1	1.3	2
G016087	45.3	9.9	2	1.2	0.1	2	3.4	0.5	2	40.3	3.7	2
G016088	67.8	10.5	2	5.6	1.4	2	6.9	1.4	2	44.1	3.3	2
G016089	64.4	10.5	2	4.5	0.1	2	1.5	0.3	2	38.3	1.8	2
G016090	67.1	7.1	2	1.7	0.1	2	16.9	2.2	2	61.3	2.3	2
G016091	47.4	12.0	2	2.1	0.9	2	2.9	2.2	2	54.0	3.9	2
G016092	71.4	11.5	2	3.3	0.1	2	6.5	0.5	2	54.9	0.9	2
G016093	76.6	8.3	2	3.3	0.2	2	3.7	0.1	2	52.2	2.5	2
G016094	75.5	7.5	2	3.6	0.4	2	1.7	0.2	2	44.0	2.5	2
G016095	76.1	7.2	2	7.6	0.4	2	2.8	0.4	2	44.1	0.4	2
G016096	77.6	8.5	2	2.1	0.2	2	4.6	0.4	2	41.2	3.0	2
G016097	44.7	15.0	2	2.2	0.1	2	0.7	0.1	2	38.6	3.2	2
G016098	28.9	8.8	2	1.6	0.4	2	1.6	0.2	2	40.1	4.7	2
G016099	68.8	11.9	2	2.2	0.1	2	3.9	0.7	2	44.8	1.0	2
G016100	85.4	6.9	2	2.4	0.6	2	3.3	0.4	2	43.5	1.9	2
G016101	4.8	1.1	2	0.8	0.1	2	0.2	0.0	2	38.0	3.3	2
G016102	57.5	14.4	2	1.9	0.1	2	2.3	0.4	2	42.6	3.3	2
G016103	69.4	12.8	2	2.4	0.0	2	5.6	0.5	2	39.1	2.7	2
G016104	66.5	12.2	2	1.6	0.7	2	11.1	0.4	2	49.0	3.3	2
G016105	58.4	14.3	2	4.4	0.6	2	8.3	0.7	2	38.4	3.0	2
G016106	74.8	5.7	2	3.3	0.3	2	7.3	0.3	2	51.8	0.1	2
G016107	45.2	12.2	2	5.8	1.1	2	5.4	0.8	2	39.6	1.0	2
G016108	15.3	3.5	2	1.2	0.1	2	1.2	0.4	2	41.5	0.1	2
G016109	77.5	3.7	2	4.5	0.4	2	3.4	0.6	2	48.3	2.5	2
G016110	43.1	15.3	2	1.7	0.2	2	6.9	1.6	2	45.4	0.6	2
G016111	89.2	1.3	2	4.2	0.1	2	5.6	0.8	2	40.6	4.4	2
G016112	68.8	13.2	2	3.8	0.4	2	1.4	0.1	2	43.4	2.7	2

Guide	% C to T			% A to G			% Indel			% MHC Class II negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
G016113	65.2	9.8	2	3.6	0.2	2	6.9	1.1	2	58.3	2.5	2
G016114	75.5	11.5	2	2.7	0.1	2	5.1	0.6	2	38.9	4.0	2
G016115	72.1	8.6	2	1.2	0.1	2	7.9	0.3	2	60.2	0.9	2
G016116	36.3	7.5	2	2.3	0.8	2	3.0	0.4	2	41.1	0.4	2
G016117	75.4	7.7	2	3.5	0.4	2	5.9	0.4	2	37.1	2.8	2

[00590] Table 34 – Percent editing and percent of MHC-II negative cells following CIITA editing with Cas9

Guide	% C to T			% A to G			% Indel			% MHC Class II negative		
	Mean	SD	n	Mean	SD	N	Mean	SD	n	Mean	SD	n
G016030	0.0	0.0	2	0.4	0.0	2	30.1	8.3	2	42.7	0.4	2
G016031	0.0	0.0	1	0.1	0.0	1	79.5	0.0	1	45.9	1.0	2
G016032	0.1	0.1	2	14.8	2.6	2	77.7	4.8	2	43.6	0.1	2
G016033	0.1	0.1	2	16.1	2.4	2	61.5	5.3	2	44.1	1.8	2
G016034	0.0	0.0	2	0.1	0.0	2	49.7	6.2	2	46.1	2.7	2
G016035	0.0	0.0	2	0.1	0.0	2	44.7	6.4	2	46.2	0.6	2
G016036	0.0	0.0	2	1.8	0.3	2	5.6	0.1	2	38.3	1.8	2
G016037	0.1	0.1	2	1.2	0.1	2	3.4	0.4	2	35.9	1.1	2
G016038	0.0	0.0	2	0.6	0.1	2	88.3	3.5	2	65.7	2.9	2
G016039	0.0	0.0	2	0.1	0.0	2	91.9	2.6	2	62.9	2.5	2
G016040	No data									63.2	0.6	2
G016041	No data									62.3	0.6	2
G016042	0.0	0.0	2	1.5	0.3	2	40.6	10.8	2	43.9	0.6	2
G016043	0.0	0.0	2	1.3	0.1	2	26.7	4.9	2	42.9	0.4	2
G016044	No data			0.2	0.0	1	74.4	0.0	1	54.9	0.8	2
G016045	0.1	0.1	2	0.8	0.0	2	13.9	4.3	2	40.6	0.8	2
G016046	0.0	0.0	2	0.2	0.1	2	92.8	2.4	2	62.5	0.6	2
G016047	0.1	0.1	2	0.2	0.0	2	80.0	2.0	2	59.8	0.9	2
G016048	0.0	0.0	2	2.1	0.0	2	9.3	2.5	2	41.0	0.8	2
G016049	0.0	0.0	2	0.1	0.0	2	85.8	3.0	2	56.9	0.3	2
G016050	0.1	0.1	2	0.6	0.1	2	38.4	1.6	2	45.0	0.4	2
G016051	0.0	0.0	2	0.4	0.0	2	82.6	3.6	2	58.8	1.4	2
G016052	0.1	0.1	2	0.6	0.1	2	69.8	7.0	2	53.8	0.4	2
G016053	No data									43.4	1.2	2
G016054	0.0	0.0	2	1.1	0.1	2	28.1	7.0	2	44.9	3.2	2
G016055	0.0	0.0	2	0.6	0.1	2	37.1	7.7	2	47.5	0.4	2
G016056	0.0	0.0	2	0.1	0.0	2	89.5	5.3	2	63.8	0.6	2
G016057	0.0	0.0	2	0.1	0.0	2	84.7	4.0	2	61.6	3.1	2

	% C to T			% A to G			% Indel			% MHC Class II negative		
Guide	**Mean**	**SD**	**n**	**Mean**	**SD**	**N**	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**
G016058	0.0	0.0	2	0.2	0.1	2	82.3	5.9	2	60.4	2.2	2
G016059	0.0	0.0	2	0.1	0.0	2	75.1	3.7	2	56.6	1.6	2
G016060	0.0	0.0	2	0.2	0.0	2	84.3	3.8	2	61.5	1.8	2
G016061	0.0	0.0	2	0.1	0.0	2	55.2	2.9	2	47.4	0.6	2
G016062	0.0	0.0	2	0.5	0.1	2	71.1	5.9	2	46.4	0.7	2
G016063	0.0	0.0	2	0.6	0.0	2	5.1	1.4	2	36.1	1.4	2
G016064	0.0	0.0	2	1.3	0.1	2	42.7	5.7	2	49.1	0.7	2
G016065	0.0	0.0	2	0.1	0.0	2	11.0	2.0	2	40.0	0.9	2
G016066	0.0	0.0	2	0.6	0.0	2	0.4	0.1	2	36.8	0.6	2
G016067	0.1	0.0	2	0.1	0.0	2	85.4	3.3	2	59.7	3.3	2
G016068	0.0	0.0	2	0.1	0.0	2	59.2	6.9	2	54.0	0.2	2
G016069	0.0	0.0	2	0.4	0.0	2	39.5	3.7	2	40.7	0.0	2
G016070	0.1	0.1	2	0.1	0.1	2	92.3	2.3	2	66.4	3.1	2
G016071	0.0	0.0	2	0.2	0.0	2	73.1	2.2	2	55.6	3.0	2
G016072	0.0	0.0	2	0.1	0.1	2	91.1	1.4	2	51.6	1.1	2
G016073	0.0	0.0	2	1.5	0.1	2	38.3	4.2	2	43.5	0.8	2
G016074	0.2	0.1	2	0.6	0.0	2	72.4	7.5	2	56.1	1.4	2
G016075	0.1	0.1	2	0.3	0.1	2	72.9	10.3	2	57.8	4.4	2
G016076	0.0	0.0	2	0.3	0.1	2	66.7	13.4	2	58.4	4.6	2
G016077	0.0	0.0	2	0.5	0.1	2	80.0	5.2	2	62.5	1.5	2
G016078	0.1	0.1	2	0.3	0.2	2	59.3	10.5	2	51.6	4.5	2
G016079	0.0	0.0	2	0.7	0.1	2	81.9	4.0	2	58.4	1.0	2
G016080	0.0	0.0	2	0.5	0.1	2	71.8	6.2	2	44.9	1.4	2
G016081	0.0	0.0	2	0.4	0.0	2	8.1	1.3	2	39.1	0.7	2
G016082	0.1	0.1	2	2.1	0.0	2	10.0	2.5	2	39.0	0.8	2
G016083	0.1	0.1	2	0.2	0.1	2	92.2	1.5	2	63.6	0.7	2
G016084	0.0	0.0	2	0.4	0.0	2	70.7	6.4	2	56.1	2.6	2
G016085	0.0	0.0	2	0.3	0.1	2	17.5	0.7	2	42.3	0.4	2
G016086	0.1	0.1	2	0.2	0.1	2	85.8	6.1	2	62.2	3.0	2
G016087	0.0	0.0	2	0.2	0.0	2	89.5	2.1	2	56.1	0.1	2
G016088	0.8	0.0	2	0.3	0.1	2	76.8	4.6	2	58.1	0.5	2
G016089	0.1	0.1	2	0.3	0.1	2	73.3	6.4	2	54.2	0.0	2
G016090	0.2	0.0	2	0.3	0.0	2	88.3	5.1	2	61.2	2.3	2
G016091	0.0	0.0	1	0.7	0.0	1	42.0	0.0	1	49.5	3.8	2
G016092	0.1	0.1	2	0.5	0.1	2	60.9	10.0	2	52.0	2.9	2
G016093	0.0	0.0	2	0.5	0.1	2	68.8	8.1	2	50.6	2.5	2
G016094	0.1	0.1	2	0.1	0.0	2	71.3	6.5	2	50.5	3.3	2
G016095	0.0	0.0	2	0.6	0.1	2	70.5	5.4	2	51.6	5.0	2
G016096	0.2	0.1	2	0.1	0.0	2	94.9	2.0	2	51.2	0.1	2
G016097	0.1	0.1	2	0.3	0.0	2	39.7	12.4	2	50.9	4.5	2

	% C to T			% A to G			% Indel			% MHC Class II negative		
Guide	**Mean**	**SD**	**n**	**Mean**	**SD**	**N**	**Mean**	**SD**	**n**	**Mean**	**SD**	**n**
G016098	0.1	0.1	2	0.2	0.0	2	23.4	7.5	2	47.2	0.5	2
G016099	0.1	0.0	2	0.2	0.0	2	84.7	5.8	2	63.2	2.5	2
G016100	0.0	0.0	2	0.3	0.1	2	79.8	7.1	2	60.3	0.6	2
G016101	0.1	0.1	2	0.6	0.1	2	2.3	0.8	2	38.8	1.5	2
G016102	0.0	0.0	2	0.4	0.1	2	75.7	8.9	2	59.9	7.1	2
G016103	0.2	0.1	2	0.6	0.1	2	76.8	4.7	2	46.9	3.4	2
G016104	1.4	0.0	1	1.1	0.0	1	66.8	0.0	1	56.1	3.1	2
G016105	0.1	0.1	2	0.7	0.3	2	90.7	5.1	2	58.0	3.0	2
G016106	0.0	0.0	2	0.2	0.0	2	95.1	2.1	2	62.2	3.0	2
G016107	0.1	0.1	2	0.2	0.0	2	84.9	2.6	2	59.5	1.1	2
G016108	0.0	0.0	2	0.6	0.1	2	19.1	4.8	2	43.3	0.8	2
G016109	0.0	0.0	2	0.1	0.0	2	86.5	3.3	2	62.9	3.5	2
G016110	0.0	0.0	2	0.6	0.1	2	34.9	10.0	2	48.0	4.2	2
G016111	65.6*	3.7	2	1.4	0.1	2	32.9	3.7	2	44.5	1.3	2
G016112	0.0	0.0	2	0.8	0.1	2	60.7	6.7	2	54.1	5.3	2
G016113	1.1	0.4	2	0.2	0.0	2	84.2	6.7	2	60.5	5.2	2
G016114	0.1	0.1	2	0.4	0.0	2	87.6	3.5	2	50.2	2.5	2
G016115	0.0	0.0	2	0.3	0.1	2	69.3	7.4	2	52.8	4.8	2
G016116	0.0	0.0	2	0.4	0.0	2	16.6	5.0	2	38.7	0.4	2
G016117	0.0	0.0	2	0.4	0.1	2	85.6	4.9	2	49.0	3.0	2

*There is a naturally occurring C/T single nucleotide polymorphism for G016111 target sequence.

Example 13 – Dose Response and Multiplexed Editing

[00591] Three guides from **Table 33**, G016086, G016092, and G016067, were further characterized for editing efficacy with increasing amounts of guide and in combination with guides targeting TRAC (G013009, G016016, or G016017) and B2M (G015991, G015995, or G015996).

[00592] Cell preparation, activation, and electroporation were performed as described in **Example 6** with the following deviations. Editing was performed using two mRNA species encoding BC22 (SEQ ID NO: 20) and UGI (SEQ ID NO:26) respectively. Editing was assessed at multiple concentrations of sgRNA, as indicated in **Table 35** and **Table 36**. When multiple guides were used in a single reaction, each guide represented one quarter of the total guide concentration.

[00593] On day 10 post-editing, T cells were phenotyped by flow cytometry to determine MHC class II protein expression as described in **Example 6**. In addition, B2M

detection was performed with B2M-FITC antibody (BioLegend, Cat. 316304) and CD3 expression was assayed using CD3-BV605 antibody (BioLegend, Cat. 317322).

DNA samples were subjected to PCR and subsequent NGS analysis, as described in **Example 1**. **Table 35** provides MHC Class II negative flow cytometry results and NGS editing for cells edited with BC22 and individual guides targeting CIITA, with **Fig. 24A** graphing the percent C-to-T conversion and **Fig. 24B** graphing the percent MHC class II negative. **Table 36** shows MHC Class II negative results for cells edited simultaneously with CIITA, B2M, and TRAC guides.

[00594] Table 35 – Percent MHC-II negative cells and NGS outcomes following *CIITA* editing (n=2)

Guide Concentration		**2 uM**		**1 uM**		**0.5 uM**		**0.25 uM**	
Assay	**Guide**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
MHC-II neg	G016086	80.2	12.0	72.2	19.7	60.5	14.3	49.7	16.1
	G016092	64.5	11.4	59.3	11.8	49.0	5.7	42.6	15.1
	G016067	77.3	4.4	76.8	2.1	64.5	6.3		
C-to-T	G016086	75.4	12.7	66.1	25.0	53.9	20.1	38.3	28.3
	G016092	71.5	18.5	62.5	25.5	45.2	13.0	36.0	28.1
	G016067	83.1	6.8	82.9	5.7	66.9	8.8	50.1	28.1
C-to-A/G	G016086	1.8	0.1	1.6	0.1	1.7	0.1	1.1	0.6
	G016092	1.6	0.1	1.3	0.2	1.2	0.0	1.2	0.5
	G016067	1.0	0.3	1.2	0.1	1.5	0.5	0.9	0.5
Indel	G016086	2.5	1.4	1.4	0.5	1.5	0.3	1.2	0.5
	G016092	2.9	0.3	3.0	0.6	3.1	0.5	2.5	1.7
	G016067	3.3	0.1	2.4	0.0	3.4	1.4	2.1	1.2

[00595] Table 36 – Percent antigen negative cells following CIITA, TRAC, and B2M editing

	Concentration per guide:	**0.5 uM**		**0.25 uM**		**0.125 uM**	
Assay	**Guide**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Triple Neg	G015995 G016086 G016017	62.1	9.3	51.9	9.4	26.1	13.6
	G015991 G016092 G016016	43.8	15.6	20.2	7.4	8.2	7.6
	G015996 G016067 G013009	35.3	17.7	15.4	12.3	6.8	7.3
MHC Class II Neg	G015995 G016086 G016017	67.0	8.0	57.6	7.8	38.6	5.4
	G015991 G016092 G016016	57.2	8.2	45.2	3.2	36.8	5.3
	G015996 G016067 G013009	53.1	7.9	69.1	43.1	41.3	7.1
CD3 Neg	G015995 G016086 G016017	92.5	2.3	90.5	3.2	77.3	15.1
	G015991 G016092 G016016	88.1	6.2	87.6	3.2	74.2	14.0
	G015996 G016067 G013009	92.8	2.0	89.5	4.8	79.3	14.4
B2M	G015995 G016086 G016017	94.9	2.5	90.0	6.4	63.2	28.0

	Concentration per guide:	0.5 uM		0.25 uM		0.125 uM	
Assay	**Guide**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Neg	G015991 G016092 G016016	73.4	19.2	29.1	13.1	14.3	15.2
	G015996 G016067 G013009	60.8	25.7	29.1	21.4	14.3	15.0

Example 14. Screening of TRAC Guide RNAs

[00596] TRAC gRNAs were screened for efficacy in T cells by assessing knockdown of CD3 surface expression using both Cas9 with UGI in trans and BC22 with UGI in trans. The percentage of T cells negative for CD3 protein and the percentage of editing at the TRAC locus was assayed following TRAC editing by electroporation with mRNA and gRNA.

Example 14.1. T cell Preparation

[00597] T cells were prepared from a leukopak using the EasySep Human T cell Isolation Kit (Stem Cell Technology, Cat. 17951) following the manufacturers protocol. T cells were cryopreserved in Cryostor CS10 freezing media (Cat. 07930) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell R10 media composed of RPMI 1640 (Corning, Cat. 10-040-CV) containing 10% (v/v) of fetal bovine serum, 2 mM Glutamax (Gibco, Cat. 35050-061), 22 μM of 2-Mercaptoethanol, 100 uM non-essential amino acids (Corning, Cat. 25-025-Cl), 1 mM sodium pyruvate, 10 mM HEPES buffer, 1% of Penicillin-Streptomycin, plus 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02). T cells were activated with Dynabeads® Human T-Activator CD3/CD28 (Gibco, Cat. 11141D). Cells were expanded in T cell media for 72 hours prior to mRNA transfection.

Example 14.2. T cell editing with RNA electroporation

[00598] Solutions containing mRNA encoding Cas9 protein (SEQ ID NO: 8), BC22 (SEQ ID NO: 20) or UGI (SEQ ID NO 26:) were prepared in sterile water. 50 μM TRAC targeting sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C before cooling on ice. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of editor mRNA, 200 ng of UGI mRNA and 20 pmols of sgRNA as described in **Table 37** in a final volume of 20 uL of P3 electroporation buffer. This mix was

transferred in triplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were rested in 180 ul of R10 media plus 100 U/mL of recombinant human interleukin-2 before being transferred to a new flat-bottom 96-well plate. The resulting plate was incubated at 37°C for 6 days. On day 9 post-editing cells were collected for flow cytometry analysis and NGS sequencing.

Example 14.3. Flow cytometry and NGS sequencing

[00599] On day 9 post-editing, T cells were phenotyped by flow cytometry to determine CD3 protein expression as described in **Example 6** using antibodies targeting CD3 (BioLegend® Cat. No. 317322) and Isotype Control-PE (BioLegend® Cat. No. 400269). DNA samples were subjected to PCR and subsequent NGS analysis, as described in **Example 1**. The mean percentage of each editing type at the TRAC locus and the mean number of CD3 negative cells after editing with BC22 and UGI are showing in **Table 37**; results after editing with Cas9 and UGI are shown in **Table 38.** C-to-T editing purity is calculated as the percentage of edited reads containing only C-to-T conversions.

Table 37 – Percent editing and percent of CD3 negative cells following *TRAC* editing with BC22 and UGI

	% C to T		% A to G		% Indel		% C to T purity		n	% CD3 negative		
Guide	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**		**Mean**	**SD**	**n**
G012086	95.21	0.42	2.46	0.16	1.24	0.04	96.26	0.12	3	1.22	0.55	3
G013004	74.7	14.43	1.6	0.32	3.25	1.05	93.97	0.52	3	17.17	7.69	3
G013005	89.47	0.6	3.24	0.11	4.04	0.49	92.48	0.53	3	8.52	1.59	3
G013006	90.56	3.87	2.82	0.23	1.74	0.15	95.21	0.16	3	2.77	0.34	3
G013007	90.96	3.68	3.05	0.37	1.51	0.35	95.21	0.37	3	3.07	1.2	3
G013008	67.48	24.46	2.15	0.67	1.07	0.51	95.45	0.64	3	1.2	0.38	3
G013009	90.05	4.17	2.29	0.27	2.09	0.17	95.36	0.36	3	87.87	5.02	3
G014831	92.48	1.55	2.26	0.22	2.93	0.25	94.68	0.54	3	50.5	3.25	3
G015997	86.72	9.45	1.34	0.09	1.18	0.21	97.14	0.50	3	2.45	0.75	3
G015998	79.5	15.18	2.12	0.63	0.98	0.12	96.25	0.13	3	2.8	1.35	3
G015999	87.16	1.81	3.57	0.03	5.8	0.77	90.30	0.57	3	14.16	6.97	3
G016000	74.27	15.17	2.79	0.16	1.6	0.16	94.24	1.47	3	6.64	0.97	3
G016001	96.49	1.69	0.88	0.14	0.42	0.06	98.67	0.22	3	15.93	3.4	3
G016002	96.11	0.65	1.18	0.1	0.87	0.07	97.90	0.09	3	4.18	2.8	3
G016003	84.79	6.22	1.04	0.06	7.1	1.92	91.24	1.99	3	2.17	1.04	3
G016004	91.78	3.75	2.16	0.18	0.95	0.23	96.72	0.36	3	1.53	0.28	3
G016005	91.02	3.02	2.16	0.08	3.06	0.13	94.58	0.11	3	15.37	2.72	3

Guide	% C to T		% A to G		% Indel		% C to T purity		n	% CD3 negative		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		Mean	SD	n
G016006	92.18	1.4	1.91	0.09	2.97	0.05	94.98	0.09	3	22.03	0.9	3
G016007	96.71	0.32	1.14	0.12	1.05	0.2	97.78	0.22	3	5.78	0.2	3
G016008	95.81	0.94	1.29	0.07	1.33	0.2	97.34	0.16	3	36.67	0.45	3
G016009	38.69	13.44	2.08	0.22	0.4	0.17	93.73	1.11	3	0.79	0.2	3
G016010	91.39	0.55	2.22	0.06	5.02	0.43	92.67	0.47	3	9.21	0.44	3
G016011	94.67	0.78	2.26	0.11	1.04	0.1	96.63	0.16	3	0.32	0.06	3
G016012	87.92	1.23	2	0.17	7.99	1.51	89.80	1.36	3	10.11	0.54	3
G016013	91.68	3.86	1.77	0.06	2.45	0.28	95.59	0.43	3	59.4	4.51	3
G016014	90.7	6.53	1.6	0.02	1.44	0.26	96.75	0.30	3	6.21	2.09	3
G016015	93.07	2.04	2.84	0.16	1.2	0.23	95.83	0.41	3	0.88	0.67	3
G016016	94.2	0.99	2.11	0.11	0.96	0.18	96.84	0.28	3	90.8	1.59	3
G016017	95.51	0.63	1.86	0.1	1.04	0.17	97.05	0.25	3	91.1	1.35	3
G016018	94.5	1.03	1.48	0.16	1.61	0.24	96.83	0.34	3	6.02	1.95	3
G016019	71.54	5.93	4.69	0.23	3.03	0.53	90.27	0.14	2	78.13	9.8	3
G016020	29.14	9.23	1.64	0.35	0.44	0.18	93.23	0.87	3	1.56	0.8	3
G016021	32.46	10.32	2.07	0.54	1.42	0.21	90.09	0.99	3	0.65	0.13	3
G016022	89.76	4.52	1.98	0.03	2.74	0.16	95.00	0.11	3	15.97	6.07	3
G016023	47.93	14.11	2.17	0.64	1.72	0.39	92.48	0.21	2	5.12	2.45	3
G016024	91.23	3.82	1.77	0.27	2.48	0.28	95.55	0.37	3	3.56	2.12	3
G016025	59.93	18.15	1.18	0.3	0.71	0.06	96.86	0.40	3	1.86	0.37	3
G016026	93.52	3.27	1.45	0.29	0.8	0.09	97.65	0.29	3	82.27	3.13	3
G016027	92.7	1.71	1.67	0.04	3.09	0.19	95.12	0.16	3	5.05	4.02	3
G016028	84.88	9.36	1.29	0.11	0.88	0.05	97.49	0.29	3	2.01	1.07	3
G016029	51.08	14.8	1.57	0.37	0.8	0.08	95.48	0.50	3	1.29	1.14	3

Table 38 – Percent editing and percent of CD3 negative cells following *TRAC* editing with Cas9

Guide	% C to T			% A to G			% Indel			% CD3 negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
G012086	0.01	0.01	3	0.10	0.02	3	98.18	0.28	3	94.87	0.25	3
G013004	0.05	0.01	3	0.17	0.08	3	64.90	20.08	3	61.17	19.91	3
G013005	0.03	0.00	3	0.05	0.01	3	94.14	1.85	3	83.43	2.05	3
G013006	0.04	0.03	3	0.11	0.07	3	96.29	2.22	3	92.90	1.32	3
G013007	0.05	0.01	3	0.11	0.02	3	96.13	1.88	3	92.93	1.69	3
G013008	0.04	0.01	3	0.23	0.11	3	83.46	9.24	3	79.53	7.89	3
G013009	0.19	0.05	3	0.20	0.02	3	94.74	2.23	3	80.87	1.65	3
G014831	0.02	0.02	3	0.07	0.02	3	99.13	0.19	3	94.87	0.21	3
G015997	0.19	0.07	3	0.16	0.01	3	90.71	4.41	3	88.77	3.65	3
G015998	0.22	0.06	3	0.15	0.04	3	93.52	4.46	3	80.93	1.07	3

Guide	% C to T			% A to G			% Indel			% CD3 negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
G015999	0.02	0.02	3	0.01	0.01	3	98.15	0.89	3	87.00	0.78	3
G016000	0.05	0.03	3	0.08	0.01	3	97.63	0.90	3	92.87	0.93	3
G016001	0.01	0.01	3	0.01	0.01	3	95.81	1.95	3	83.63	1.05	3
G016002	0.02	0.02	3	0.08	0.07	3	97.47	0.51	3	32.97	3.19	3
G016003	0.17	0.08	3	0.08	0.03	3	95.15	3.25	3	70.80	2.09	3
G016004	0.06	0.06	3	0.13	0.07	3	97.56	0.99	3	93.83	0.81	3
G016005	0.22	0.07	3	0.21	0.07	3	98.75	0.01	3	87.40	1.47	3
G016006	0.03	0.03	3	0.14	0.06	3	97.76	0.66	3	94.93	0.47	3
G016007	0.09	0.04	3	0.06	0.03	3	98.40	0.46	3	94.73	0.40	3
G016008	0.00	0.01	3	0.01	0.01	3	98.92	0.35	3	94.67	0.55	3
G016009	0.19	0.05	3	0.77	0.11	3	13.55	3.43	3	68.13	3.92	3
G016010	0.02	0.00	3	0.19	0.04	3	98.20	0.40	3	93.63	0.21	3
G016011	0.06	0.02	3	0.18	0.03	3	97.57	0.73	3	93.03	63.50	3
G016012	0.04	0.02	3	0.12	0.05	3	97.37	0.15	3	93.87	0.76	3
G016013	0.03	0.04	3	0.05	0.00	3	97.49	1.15	3	82.40	2.11	3
G016014	0.00	0.00	3	0.04	0.06	3	98.15	1.00	3	92.43	0.90	3
G016015	0.07	0.05	3	0.19	0.03	3	98.09	0.20	3	94.03	0.29	3
G016016	0.01	0.01	3	0.05	0.01	3	97.03	0.81	3	84.17	0.29	3
G016017	0.03	0.03	3	0.05	0.01	3	97.64	0.43	3	77.80	4.97	3
G016018	0.05	0.03	3	0.03	0.01	3	98.44	0.32	3	61.57	6.82	3
G016019	0.01	0.00	3	0.06	0.02	3	88.15	5.31	3	31.43	5.54	3
G016020	0.03	0.02	3	0.10	0.03	3	79.18	7.16	3	93.30	1.28	3
G016021	0.03	0.01	3	0.23	0.04	3	42.54	6.06	3	70.13	8.45	3
G016022	0.00	0.00	3	0.01	0.01	3	98.01	1.02	3	94.23	0.72	3
G016023	0.07	0.03	3	0.54	0.10	3	88.15	7.96	3	51.13	15.47	3
G016024	0.02	0.02	3	0.24	0.13	3	98.29	0.54	3	76.57	2.92	3
G016025	0.15	0.02	3	0.24	0.04	3	58.36	15.14	3	62.27	3.41	3
G016026	0.00	0.01	3	0.03	0.01	3	95.34	1.69	3	65.37	10.10	3
G016027	0.08	0.01	3	0.27	0.14	3	92.27	3.82	3	15.07	5.42	3
G016028	0.07	0.06	3	0.16	0.05	3	73.05	9.53	3	90.10	0.28	2
G016029	0.09	0.01	3	0.51	0.04	3	19.24	5.53	3	95.00	0.28	2

Example 15. Screening of TRBC Guide RNAs

[00600] TRBC gRNAs were screened for efficacy in T cells by assessing knockdown of CD3 surface expression using both Cas9 with UGI in trans and BC22 with UGI in trans. The percentage of T cells negative for CD3 protein and the percentage of editing of each type at the TRBC1 locus was assayed following TRBC editing by electroporation with mRNA and gRNA.

Example 15.1. T cell Preparation

[00601] T cells were prepared from a leukopak using the EasySep Human T cell Isolation Kit (Stem Cell Technology, Cat. 17951) following the manufacturers protocol. T cells were cryopreserved in Cryostor CS10 freezing media (Cat. 07930) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell R10 media composed of RPMI 1640 (Corning, Cat. 10-040-CV) containing 10% (v/v) of fetal bovine serum, 2 mM Glutamax (Gibco, Cat. 35050-061), 22 μM of 2-Mercaptoethanol, 100 uM non-essential amino acids (Corning, Cat. 25-025-Cl), 1 mM sodium pyruvate, 10 mM HEPES buffer, 1% of Penicillin-Streptomycin, plus 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02). T cells were activated with Dynabeads® Human T-Activator CD3/CD28 (Gibco, Cat. 11141D). Cells were expanded in T cell media for 72 hours prior to mRNA transfection.

Example 15.2. T cell editing with RNA electroporation

[00602] Solutions containing mRNA encoding Cas9 protein (SEQ ID NO: 8), BC22 (SEQ ID NO: 20) or UGI (SEQ ID NO 26:) were prepared in sterile water. 50 μM TRBC targeting sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C before cooling on ice. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of editor mRNA, 200 ng of UGI mRNA and 20 pmols of sgRNA as described in **Table 39** in a final volume of 20 uL of P3 electroporation buffer. This mix was transferred in triplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were rested in 180 ul of R10 media plus 100 U/mL of recombinant human interleukin-2 before being transferred to a new flat-bottom 96-well plate. The resulting plate was incubated at 37°C for 6 days. On day 9 post-editing cells were collected for flow cytometry analysis and NGS sequencing.

Example 15.3. Flow cytometry and NGS sequencing

[00603] On day 9 post-editing, T cells were phenotyped by flow cytometry to determine CD3 protein expression as described in **Example 6** using antibodies targeting CD3 (BioLegend® Cat. No. 317322) and Isotype Control-PE (BioLegend® Cat. No. 400269). DNA samples were subjected to PCR and subsequent NGS analysis, as described in **Example 1**. The mean percentage of each editing type at the TRAC locus and the mean number of CD3

negative cells after editing with BC22 and UGI are showing in **Table 39**; results after editing with Cas9 and UGI are shown in **Table 40**. C-to-T editing purity is calculated as the percentage of edited reads containing only C-to-T conversions.

Table 39 – Percent editing and percent of CD3 negative cells following *TRBC* editing with BC22

Guide	% C to T		% A to G		% Indel		% C to T purity		n	% CD3 negative		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		Mean	SD	N
G013015	93.67	0.06	3.16	0.19	2.46	0.19	94.34	0.00	2	32.55	1.77	2
G014832	90.99	0.09	3.08	0.24	3.66	0.49	93.10	0.23	2	13.80	0.85	2
G016167	88.56	0.62	5.06	0.26	5.76	0.47	89.11	0.72	2	30.25	7.28	2
G016168	93.21	0.42	2.12	0.09	4.19	0.09	93.67	0.20	2	35.75	0.07	2
G016171	92.69	0.50	3.59	0.15	3.09	0.33	93.28	0.48	2	57.90	0.28	2
G016172	81.81	1.17	5.54	0.28	10.17	1.74	83.90	1.93	2	50.70	2.83	2
G016174	92.42	1.66	0.99	0.09	4.74	0.19	94.17	0.20	2	92.75	0.78	2
G016177	96.72	0.07	0.91	0.05	1.64	0.18	97.43	0.13	2	21.10	4.67	2
G016179	82.05	1.68	6.67	0.28	9.58	0.83	83.46	1.22	2	40.00	2.12	2
G016180	94.32	0.31	1.46	0.22	1.06	0.23	97.39	0.01	2	91.70	0.71	2
G016181	92.19	0.51	3.77	0.30	2.04	0.14	94.07	0.18	2	19.20	1.27	2
G016183	92.38	0.73	2.99	0.31	3.42	0.43	93.51	0.75	2	24.60	0.85	2
G016187	92.08	2.19	3.14	0.33	1.88	0.47	94.82	0.90	2	13.90	0.85	2
G016188	79.29	2.46	2.78	0.00	3.06	0.15	93.14	0.03	2	90.00	0.57	2
G016189	93.26	0.49	2.79	0.19	2.76	0.45	94.38	0.64	2	92.05	0.64	2
G016190	87.95	0.88	2.56	0.19	3.41	0.14	93.64	0.11	2	5.44	0.35	2
G016191	51.28	4.28	2.37	0.21	2.65	0.05	91.06	0.41	2	3.88	0.59	2
G016192	90.50	1.28	3.62	0.10	5.08	1.13	91.23	1.24	2	14.10	2.26	2
G016193	88.60	0.69	3.61	0.10	4.95	0.05	91.19	0.11	2	8.06	0.46	2
G016194	88.95	0.62	0.67	0.02	4.55	0.47	94.45	0.49	2	19.70	4.81	2
G016195	91.17	0.74	2.57	0.15	3.19	0.23	94.05	0.04	2	13.65	1.91	2
G016196	91.58	0.80	2.16	0.32	5.05	0.11	92.70	0.46	2	18.25	0.78	2
G016197	89.55	0.90	2.72	0.17	6.99	0.68	90.22	0.87	2	16.25	7.71	2
G016199	91.01	1.34	4.05	0.15	3.10	0.53	92.71	0.75	2	93.75	0.21	2
G016200	94.68	0.86	2.09	0.34	2.68	0.60	95.20	0.95	2	93.95	0.07	2
G016201	89.27	0.55	5.04	0.63	4.80	0.14	90.07	0.76	2	19.40	2.97	2
G016203	79.59	1.09	3.78	0.55	9.40	0.25	85.80	0.57	2	12.95	0.49	2
G016204	93.89	0.26	3.36	0.07	2.06	0.08	94.54	0.01	2	42.85	1.06	2
G016205	67.12	3.18	3.84	0.00	5.66	0.09	87.59	0.62	2	41.35	1.48	2
G016206	92.29	0.93	2.63	0.42	1.65	0.12	95.56	0.34	2	93.00	0.28	2
G016207	85.70	2.33	5.16	0.32	8.56	2.71	86.20	2.39	2	44.60	0.71	2
G016211	84.04	2.44	1.20	0.09	2.92	0.02	95.32	0.06	2	67.50	3.96	2

Guide	% C to T		% A to G		% Indel		% C to T purity		n	% CD3 negative		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		Mean	SD	N
G016212	95.39	0.87	0.78	0.11	2.11	0.44	97.05	0.58	2	48.30	0.00	2
G016214	88.50	3.77	2.03	0.05	1.53	0.38	96.12	0.51	2	83.75	3.75	2
G016216	91.37	0.45	3.67	0.40	4.65	0.84	91.65	0.44	2	60.65	0.92	2
G016217	90.89	0.05	3.59	0.01	5.13	0.04	91.25	0.05	2	35.75	3.46	2
G016219	89.15	1.07	3.93	0.16	5.59	0.64	90.35	0.83	2	47.85	2.33	2
G016222	82.02	0.15	4.25	0.09	12.30	0.07	83.21	0.05	2	22.55	0.35	2
G016227	67.64	0.78	4.66	0.28	14.36	0.70	78.05	0.57	2	38.20	1.27	2
G016228	93.12	1.35	4.09	0.32	1.11	0.13	94.71	0.11	2	14.30	0.99	2
G016230	88.34	1.74	3.06	0.11	5.98	1.48	90.71	1.45	2	13.25	2.76	2
G016232	71.60	n/a	6.32	n/a	21.40	n/a	72.09	n/a	1	39.80	0.71	2
G016233	87.08	1.67	2.87	0.18	8.87	2.03	88.13	1.86	2	22.20	8.91	2
G016238	90.49	1.00	2.74	0.20	4.68	0.22	92.41	0.10	2	25.10	6.79	2
G016239	87.37	0.96	5.11	0.06	7.02	1.07	87.81	1.01	2	29.95	1.63	2
G016240	76.58	1.58	3.86	0.23	16.03	0.07	79.39	0.21	2	55.55	3.32	2
G016241	69.44	1.35	3.89	0.04	24.19	1.07	71.21	1.21	2	61.50	1.84	2
G016242	91.60	0.84	2.92	0.06	4.83	0.56	92.20	0.64	2	3.37	0.28	2
G016243	93.32	2.21	2.74	0.02	1.69	0.36	95.46	0.48	2	33.75	5.16	2
G016245	81.87	0.70	6.03	0.31	8.96	0.11	84.52	0.28	2	40.00	2.40	2
G016248	81.23	2.35	1.14	0.04	2.95	0.34	95.21	0.20	2	5.33	0.21	2
G016249	No data									6.87	2.02	2
G016250	No data									12.50	2.12	2
G016251	92.57	0.21	2.66	0.37	4.40	0.21	92.91	0.16	2	31.45	1.48	2
G016252	84.54	0.88	1.23	0.06	1.93	0.03	96.40	0.14	2	2.47	0.17	2
G016253	93.22	0.40	1.19	0.00	2.83	0.38	95.86	0.35	2	28.55	0.21	2
G016254	92.52	0.86	2.96	0.16	3.53	0.57	93.45	0.75	2	2.02	0.44	2
G016255	75.91	0.26	3.64	0.06	18.80	0.10	77.19	0.06	2	21.70	0.42	2
G016256	69.87	5.75	3.53	0.29	1.87	0.48	92.82	0.31	2	6.39	2.44	2
G016257	90.90	0.20	2.36	0.05	5.81	0.08	91.75	0.05	2	5.01	0.46	2
G016258	69.09	1.72	3.05	0.31	24.53	0.99	71.46	1.01	2	26.90	0.71	2
G016259	88.91	1.19	1.99	0.29	5.58	0.45	92.15	0.81	2	8.44	0.88	2
G016260	92.85	1.13	1.82	0.27	2.58	0.11	95.48	0.42	2	4.30	0.07	2
G016261	92.78	1.36	0.94	0.43	4.51	0.44	94.45	0.91	2	16.55	1.34	2
G016262	95.60	0.22	1.36	0.03	2.59	0.37	96.03	0.34	2	4.29	0.81	2
G016263	No data									78.40	0.14	2
G016264	87.24	0.43	1.60	0.02	5.83	0.44	92.15	0.49	2	21.90	0.42	2
G016265	48.92	6.40	2.49	0.13	1.49	0.07	92.41	0.81	2	2.53	0.22	2
G016266	72.19	0.15	6.56	0.20	20.76	0.17	72.55	0.02	2	18.90	3.68	2
G016267	92.15	0.40	4.24	0.38	2.94	0.26	92.78	0.14	2	16.65	8.70	2
G016268	93.60	0.39	3.72	0.25	2.43	0.04	93.83	0.29	2	3.03	0.24	2
G016269	48.11	0.16	4.00	0.55	2.93	0.20	87.42	1.23	2	17.05	1.63	2

Guide	% C to T		% A to G		% Indel		% C to T purity		n	% CD3 negative		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		Mean	SD	N
G016270	87.48	0.28	2.81	0.16	9.04	0.32	88.06	0.46	2	66.90	0.57	2
G016271	91.19	1.41	4.00	0.08	3.94	1.08	91.99	1.19	2	14.50	10.33	2
G016272	93.90	0.09	1.39	0.12	4.07	0.23	94.50	0.11	2	27.20	1.56	2
G016273	87.75	0.20	5.66	0.09	5.36	0.36	88.84	0.26	2	31.50	3.25	2
G016274	83.68	1.26	4.02	0.19	7.12	0.27	88.24	0.59	2	17.50	6.93	2
G016275	93.35	0.24	3.14	0.20	2.52	0.19	94.28	0.00	2	14.65	9.26	2
G016276	94.95	0.27	1.98	0.04	2.65	0.21	95.35	0.26	2	30.35	1.48	2
G016277	90.15	1.29	5.09	0.18	1.82	0.65	92.89	0.54	2	16.13	13.68	2
G016278	83.20	3.62	5.41	0.96	2.92	0.11	90.88	1.20	2	8.88	1.59	2
G016279	90.24	0.72	4.88	0.45	3.23	0.08	91.76	0.40	2	13.70	5.23	2
G016280	91.06	1.61	2.95	0.43	2.76	0.45	94.09	0.95	2	13.81	7.06	2
G016281	94.85	0.41	1.28	0.01	3.31	0.47	95.38	0.46	2	13.85	5.45	2
G016282	93.39	0.27	3.06	0.05	3.13	0.18	93.78	0.24	2	16.95	7.00	2
G016283	77.09	n/a	5.72	n/a	16.13	n/a	77.92	n/a	1	20.40	3.25	2
G016284	87.29	1.73	2.01	1.11	9.08	0.55	89.29	2.50	2	8.38	1.51	2
G016285	60.72	3.51	8.71	0.53	13.95	2.51	72.80	2.87	2	15.70	2.12	2

Table 40 – Percent editing and percent CD3 negative cells following *TRBC* editing with Cas9

Guide	% C to T			% A to G			% Indel			% CD3 negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	N
G013015	0.03	0.04	2	0.09	0.01	2	99.43	0.14	2	93.60	0.28	2
G014832	0.07	0.02	2	0.10	0.00	2	98.24	0.65	2	84.95	0.07	2
G016167	0.01	0.00	2	0.00	0.00	2	99.36	0.26	2	90.10	0.28	2
G016168	0.02	0.01	2	0.18	0.14	2	98.29	0.07	2	95.00	0.28	2
G016171	0.03	0.04	2	0.03	0.02	2	99.11	0.34	2	91.80	0.71	2
G016172	0.01	0.00	2	1.29	0.45	2	96.54	1.15	2	93.60	0.42	2
G016174	0.05	0.05	2	0.14	0.05	2	98.65	0.68	2	94.10	0.42	2
G016177	0.10	0.03	2	0.08	0.03	2	99.02	0.28	2	94.05	0.35	2
G016179	0.09	0.02	2	0.21	0.04	2	97.55	0.18	2	93.20	0.42	2
G016180	0.01	0.00	2	0.19	0.07	2	95.64	0.93	2	91.25	1.20	2
G016181	0.25		1	0.03		1	94.84		1	92.00	0.28	2
G016183	0.04	0.04	2	0.01	0.01	2	98.56	0.53	2	92.90	0.00	2
G016187	0.18	0.00	2	0.05	0.07	2	99.07	0.07	2	93.95	0.64	2
G016188	0.07	0.03	2	0.30	0.07	2	94.24	0.69	2	85.95	0.35	2
G016189	0.12	0.03	2	0.64	0.21	2	97.19	0.73	2	89.75	0.49	2
G016190	6.51	6.05	2	0.15	0.21	2	77.21	4.03	2	70.55	0.35	2
G016191	3.67	0.17	2	0.58	0.13	2	57.71	0.58	2	63.95	0.21	2
G016192	0.11	0.08	2	0.24	0.09	2	98.50	0.20	2	89.50	0.28	2
G016193	0.05	0.02	2	0.15	0.04	2	98.54	0.99	2	92.45	0.35	2

Guide	% C to T			% A to G			% Indel			% CD3 negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	N
G016194	4.24	0.49	2	0.06	0.01	2	85.43	0.76	2	68.00	0.14	2
G016195	0.03	0.04	2	0.00	0.00	2	99.51	0.20	2	93.55	0.35	2
G016196	0.06	0.01	2	0.15	0.05	2	80.55	2.75	2	82.15	1.34	2
G016197	0.00	0.01	2	0.01	0.00	2	99.64	0.02	2	75.20	0.42	2
G016199	0.15	0.01	2	0.22	0.01	2	98.25	0.19	2	91.70	0.57	2
G016200	0.01	0.01	2	0.06	0.02	2	99.45	0.16	2	93.65	0.64	2
G016201	0.05	0.05	2	0.15	0.00	2	93.38	0.20	2	85.10	0.71	2
G016203	0.05	n/a	1	0.09	n/a	1	98.61	n/a	1	92.20	0.28	2
G016204	0.08	0.02	2	0.02	0.01	2	99.31	0.23	2	92.50	0.57	2
G016205	0.05	0.04	2	0.75	0.02	2	77.14	2.06	2	76.20	2.69	2
G016206	0.24	0.09	2	0.24	0.05	2	96.55	0.05	2	91.85	0.07	2
G016207	0.06	0.03	2	0.07	0.01	2	99.25	0.02	2	93.80	0.42	2
G016211	0.04	0.02	2	0.09	0.01	2	84.26	0.46	2	87.75	1.34	2
G016212	0.09	0.00	2	0.10	0.08	2	91.57	1.87	2	90.90	0.42	2
G016214	0.07	0.05	2	0.16	0.02	2	88.68	0.74	2	87.10	0.85	2
G016216	0.08	0.06	2	0.83	0.28	2	98.49	0.46	2	88.50	0.85	2
G016217	0.17	0.06	2	0.08	0.01	2	98.85	0.25	2	93.70	0.85	2
G016219	0.15	0.21	2	0.08	0.04	2	97.51	0.98	2	89.95	0.92	2
G016222	0.03	0.02	2	0.08	0.03	2	98.09	0.23	2	93.35	0.35	2
G016227	0.02	n/a	1	0.12	n/a	1	98.97	n/a	1	84.30	0.71	2
G016228	0.12	0.17	2	0.10	0.08	2	99.05	0.21	2	92.65	0.21	2
G016230	0.10	0.00	2	0.04	0.02	2	98.96	0.30	2	77.90	0.42	2
G016232	0.13	0.04	1	0.02	0.01	1	99.07	0.51	1	94.45	0.21	2
G016233	0.07	0.01	2	0.07	0.04	2	99.00	0.45	2	71.00	1.27	2
G016238	0.09	0.08	2	0.14	0.02	2	93.50	0.81	2	90.90	0.00	2
G016239	0.08	0.08	2	0.09	0.05	2	98.74	0.30	2	85.70	0.71	2
G016240	0.00	n/a	1	0.05	n/a	1	98.91	n/a	1	84.50	0.28	2
G016241	0.00	0.00	2	0.00	0.00	2	99.56	0.05	2	86.05	0.35	2
G016242	0.00	0.00	2	0.03	0.03	2	98.50	0.02	2	87.55	0.64	2
G016243	0.00	0.00	2	0.07	0.01	2	98.44	0.49	2	93.55	0.78	2
G016245	0.45	0.51	2	2.10	0.20	2	95.23	1.04	2	92.70	0.14	2
G016248	2.51	0.44	2	0.09	0.03	2	85.91	2.49	2	78.90	1.13	2
G016249	0.00	0.00	0	0.01	0.00	0	98.87	0.42	0	93.05	0.07	2
G016250	0.00	0.00	0	0.01	0.00	0	99.11	0.30	0	93.35	1.06	2
G016251	0.01	0.02	2	0.04	0.02	2	97.64	1.98	2	93.40	0.42	2
G016252	0.01	0.01	2	0.05	0.01	2	76.81	1.22	2	28.95	3.75	2
G016253	0.11	0.08	2	0.06	0.05	2	95.66	0.25	2	67.70	0.71	2
G016254	0.05	0.02	2	0.07	0.01	2	98.21	0.25	2	25.65	0.92	2
G016255	0.01	0.01	2	1.08	1.33	2	97.81	0.73	2	62.75	0.07	2
G016256	0.06	0.01	2	0.46	0.07	2	80.84	1.58	2	21.05	8.41	2
G016257	0.03	0.02	2	0.01	0.00	2	97.59	0.32	2	85.70	0.00	2
G016258	0.10	0.13	2	0.23	0.04	2	78.74	3.06	2	47.45	1.20	2

Guide	% C to T			% A to G			% Indel			% CD3 negative		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	N
G016259	0.00	0.00	2	0.21	0.04	2	97.82	0.31	2	47.00	2.12	2
G016260	1.16	0.20	2	0.15	0.15	2	94.72	0.27	2	28.30	0.28	2
G016261	0.03	0.04	2	0.03	0.04	2	95.13	0.46	2	90.95	0.07	2
G016262	1.17	0.31	2	0.07	0.05	2	98.39	0.31	2	26.85	0.49	2
G016263	0.05	0.06	0	0.08	0.07	0	97.11	2.82	0	93.50	0.28	2
G016264	0.09	0.01	2	0.14	0.04	2	75.23	0.49	2	48.70	1.70	2
G016265	0.08	0.01	2	0.29	0.01	2	47.59	0.92	2	15.60	1.13	2
G016266	0.02	0.00	2	16.72	0.36	2	82.14	0.85	2	25.25	2.76	2
G016267	0.09	0.01	2	0.15	0.06	2	97.50	0.21	2	34.20	1.56	2
G016268	0.01	0.00	2	0.07	0.00	2	98.95	0.20	2	40.30	1.41	2
G016269	0.14	0.06	2	0.65	0.01	2	38.34	4.64	2	15.35	0.64	2
G016270	0.12	0.07	2	0.32	0.13	2	99.10	0.33	2	93.60	0.85	2
G016271	0.16	0.04	2	0.10	0.01	2	98.21	0.27	2	35.55	5.16	2
G016272	0.09	0.05	2	0.19	0.04	2	94.63	1.28	2	93.45	0.78	2
G016273	0.00	0.00	2	0.01	0.00	2	97.97	0.37	2	92.85	1.20	2
G016274	0.06	0.02	2	0.23	0.12	2	87.53	0.83	2	71.30	1.84	2
G016275	0.24	0.22	2	0.03	0.00	2	98.58	0.13	2	31.50	2.97	2
G016276	0.02	0.01	2	0.34	0.04	2	98.56	0.14	2	24.35	4.03	2
G016277	0.52	0.04	2	0.32	0.04	2	97.86	0.06	2	31.75	4.31	2
G016278	0.30	0.01	2	0.27	0.04	2	98.12	0.07	2	32.10	0.28	2
G016279	0.59	0.15	2	0.14	0.00	2	98.69	0.12	2	30.20	5.37	2
G016280	0.24	0.10	2	0.24	0.01	2	90.41	0.60	2	28.55	3.46	2
G016281	0.65	0.88	2	0.03	0.04	2	98.14	0.28	2	92.90	0.42	2
G016282	0.04	0.02	2	0.01	0.00	2	98.57	0.05	2	93.45	0.21	2
G016283	0.01	n/a	1	0.01	n/a	1	93.16	n/a	1	87.35	2.47	2
G016284	0.02	0.02	2	0.01	0.00	2	97.97	0.35	2	42.05	0.64	2
G016285	0.06	0.08	2	0.01	0.01	2	97.89	0.86	2	89.80	0.14	2

Example 16. Profiling RNA Off-Targets Following B2M Editing

Example 16.1. T cell preparation

[00604] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. No. 130-070-525) on the LOVO device. T cells were isolated via positive selection using CD4 and CD8 magnetic beads (Miltenyi Biotec Cat. No. 130-030-401/130-030-801) using the CliniMACS® Plus and CliniMACS® LS disposable kit. T cells were aliquoted into vials and cryopreserved in Cryostor® CS10 (StemCell Technologies Cat. 07930) for future use. Upon thawing, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth (TCG) media consisting of CTS OpTmizer T cell expansion serum-free

media (Thermofisher, Cat. A3705001) supplemented with 5% human AB serum (Gemini, Cat. 100-512), 1X GlutaMAX (Thermofisher, Cat. 35050061), 10 mM HEPES (Thermofisher, Cat. 15630080) and 1x of Penicillin-Streptomycin, further supplemented with 200 U/mL IL-2 (Peprotech, Cat. 200-02), 10 ng/ml IL-7 (Peprotech, Cat. 200-07), 10 ng/ml IL-15 (Peprotech, Cat. 200-15). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec, Cat. 130-111-160). Cells were expanded for 72 hours at 37°C prior to mRNA electroporation.

Example 16.2. mRNA and sgRNA electroporation of T cells

[00605] A solution containing mRNA encoding Cas9 (SEQ ID NO: 11), BC22n (SEQ ID NO: 2), UGI (SEQ ID NO: 26) or BE4MAX (SEQ ID NO: 32) was prepared in sterile water. A 50 μM B2M targeting sgRNA (G015995) was removed from its storage plate and denatured for 2 minutes at 95°C before cooling at room temperature. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of mRNAs and 20 pmols of sgRNA in a final volume of 20 uL of P3 electroporation buffer. The T cell mix was transferred in 5 replicates to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 μL of TCG media without cytokines for 15 minutes before being transferred to a new flat-bottom 96-well plate containing an additional 100 μL of TCG media supplemented with 2X cytokines cited in Section 16.1.

[00606] To evaluate RNA off-targets when expression level of BC22n peaks, a fraction of edited T cells was collected 24h post-electroporation. This fraction was further divided into 2 plates, one of which was subjected to cell lysis, PCR amplification and NGS analysis, while the other fraction was used for RNA extraction and transcriptome sequencing. On day 3 post-electroporation, the remaining T cells were collected for cell lysis and NGS sequencing, which enabled confirmation of maximum B2M editing in these samples at a timepoint when editing is normally complete.

Example 16.3. NGS sequencing

[00607] At 24 and 72h post-electroporation, T cells were subjected to lysis, PCR amplification of the B2M locus and subsequent NGS analysis, as described in Example

1. **Table 41** and **Fig. 25** show B2M editing levels in T cells collected at both timepoints following B2M editing.

Table 41. Mean editing percentages at the B2M locus following editing

			% C to T		% C toA/G		% Indel	
Time (h)	**mRNA**	**sgRNA**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
24	None	None	0.58	0.06	0.60	0.05	0.22	0.01
	Cas9	G015995	0.48	0.27	0.35	0.05	49.97	9.52
	Cas9, UGI	G015995	0.35	0.09	0.36	0.05	59.44	11.46
	BC22n, UGI	G015995	66.57	5.39	1.27	0.64	1.52	0.61
	BE4MAX, UGI	G015995	36.83	4.89	0.70	0.10	0.38	0.19
72	None	None	0.60	0.09	0.61	0.06	0.23	0.06
	Cas9	G015995	0.33	0.31	0.28	0.16	89.74	1.38
	Cas9, UGI	G015995	0.30	0.24	0.25	0.21	90.02	3.54
	BC22n, UGI	G015995	91.19	1.80	1.84	0.64	1.19	0.37
	BE4MAX, UGI	G015995	67.62	2.38	0.84	0.20	0.49	0.19

Example 16.4. Whole transcriptome sequencing

[00608] At 24h post-electroporation, T cells were centrifuged, and the cell pellet was resuspended in 200 uL of TRIzol™ reagent (Thermo Fisher Scientific, Cat No. 15596026) which was frozen at -80C for future processing. Total RNA was extracted from samples in TRIzol™ reagent using the Direct-zol RNA microprep kit (Zymo Research, Cat No. R2062) following the manufacturer's protocol. Purified RNA samples were quantified in a NanoDrop™ 8000 spectrophotometer (Thermo Fisher Scientific) and diluted to 18.18 ng/uL using nuclease-free water. From each experimental group shown in **Fig. 25**, 3 samples were randomly chosen for transcriptomic analysis. From each sample, 100 ng (5.5 uL) of purified total RNA were depleted of ribosomal RNA (rRNA) components using the NEBNext® rRNA depletion kit (New England Biolabs, Cat. E6350L) according to the manufacturer's instructions. rRNA-depleted samples were converted into double-stranded DNA libraries using NEBNext® Ultra™ II directional RNA library prep kit for Illumina® (New England Biolabs, Cat. E7765S) following the manufacturer's protocol. Amplified libraries were quantified in a Qubit 4 fluorometer (Thermo Fisher Scientific) and the average fragment size of each library was obtained by capillary electrophoresis. Libraries were pooled at an equimolar concentration and sequenced in the Illumina NextSeq550 platform using a high-output 300-cycle kit (Illumina, Cat. 20024908).

Example 16.5. Data processing for single nucleotide variant (SNV) analysis

[00609] Paired-end reads were aligned to human genome GRCh38 with STAR v2.7.1a (Dobin et al., 2013). PCR duplicates were removed with Picard MarkDuplicates v2.19.0 (BroadInstitute, 2019). GATK tools SplitNCigarReads, BaseRecalibrator, ApplyBQSR v4.1.8.1 were deployed consecutively to preprocess alignments. Variations were called with GATK HaplotypeCaller (Auwera et al., 2013; DePristo et al., 2011; McKenna et al., 2010). Variants discovered from replicates of the same sample were merged using bcftools v1.8 (Li, 2011). Sample specific variants were retrieved by excluding variants discovered in controls using vcflib v1.0.0 (Garrison, 2016). Relative C to U frequencies were calculated by dividing the number of C to U variants by the total number of SNVs for each sample. Treatment groups were compared using unpaired t-tests and the statistical significance was determined using the Holm-Sidak method with an alpha of 0.05, without assuming a consistent standard deviation.

[00610] Compared to samples treated with Cas9 mRNA, T cells electroporated with both Cas9 and UGI mRNAs, BC22n and UGI mRNAs or BE4MAX and UGI mRNAs showed no statistically significant ($p < 0.05$) increase in the frequency of C to U transitions, demonstrating the absence of detectable homology-independent cytosine deamination events in the transcriptome of these samples (**Table 42** and **Fig. 26**).

Table 42 – Mean percentage of SNVs in RNA transcripts that are C to U conversions

		% C to U SNVs			
mRNA(s)	**sgRNA**	**Mean**	**SD**	**p value (vs. Cas9)**	**p value (vs. Cas9, UGI)**
Cas9	G015995	10.84	0.25	NA	0.42
Cas9, UGI	G015995	11.07	0.35	0.42	NA
BC22n, UGI	G015995	10.59	0.12	0.19	0.09
BE4MAX, UGI	G015995	10.80	0.29	0.88	0.37

Example 17. Whole Genome Sequencing of Amplified T Cell Genomes Following B2M Editing

Example 17.1. T cell preparation

[00611] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. No. 130-070-525) on the LOVO device. T cells were isolated via positive selection using CD4 and CD8 magnetic beads (Miltenyi Biotec Cat. No. 130-030-

401/130-030-801) using the CliniMACS® Plus and CliniMACS® LS disposable kit. T cells were aliquoted into vials and cryopreserved in Cryostor® CS10 (StemCell Technologies Cat. 07930) for future use. Upon thawing, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth (TCG) media consisting of CTS OpTmizer T cell expansion serum-free media (Thermofisher, Cat. A3705001) supplemented with 5% human AB serum (Gemini, Cat. 100-512), 1X GlutaMAX (Thermofisher, Cat. 35050061), 10 mM HEPES (Thermofisher, Cat. 15630080) and 1x of Penicillin-Streptomycin, further supplemented with 200 U/mL IL-2 (Peprotech, Cat. 200-02), 10 ng/ml IL-7 (Peprotech, Cat. 200-07), 10 ng/ml IL-15 (Peprotech, Cat. 200-15). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec, Cat. 130-111-160). Cells were expanded for 72 hours at 37°C prior to mRNA electroporation.

Example 17.2. mRNA and sgRNA electroporation of T cells

[00612] A solution containing mRNA encoding Cas9 (SEQ ID NO: 11), BC22n (SEQ ID NO: 2) or UGI (SEQ ID NO: 26) was prepared in sterile water. A 50 μM B2M targeting sgRNA (G015995) was removed from its storage plate and denatured for 2 minutes at 95°C before cooling at room temperature. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x $10^{\wedge 6}$ T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of mRNAs and 20 pmols of sgRNA in a final volume of 20 uL of P3 electroporation buffer. The T cell mix was transferred in 8 replicates to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 μL of TCG media without cytokines for 15 minutes before being transferred to a new flat-bottom 96-well plate containing an additional 100 μL of TCG media supplemented with 2X cytokines.

[00613] To promote expansion, T cells were split at the ratios of 1:4 and 1:3 on days 3 and 6 post-electroporation, respectively, using fresh TCG media with 1X cytokines. On day 7 post-electroporation, a fraction of cells was collected for flow cytometry and NGS sequencing, while the remaining cells were frozen for subsequent single-cell whole genome amplification and sequencing.

Example 17.3. Flow cytometry and NGS sequencing

[00614] On day 7 post-electroporation, T cells were phenotyped by flow cytometry to evaluate loss of B2M expression levels following editing with sgRNA

G015995. Briefly, T cells were incubated for 30 min at 4°C with a mixture of antibodies against CD3 (BioLegend® Cat. No. 317340), CD4 (BioLegend® Cat. No. 300537), CD8 (BioLegend® Cat. No. 344706) and B2M (BioLegend® Cat. No. 316314), diluted at 1:200 in cell staining buffer (BioLegend® Cat. No. 420201). Cells were subsequently washed, processed on a Cytoflex flow cytometer (Beckman Coulter) and analyzed using the FlowJo software package. T cells were gated based on size, shape and B2M expression. **Table 43** and **Fig. 27** show the percentage of cells expressing B2M in edited T cells.

Table 43 – Loss of B2M expression (% of cells lacking surface marker) assessed by flow cytometry analysis of T cells harvested 7 days post-treatment with different mRNA combinations and B2M sgRNA G015995 (SD = standard deviation).

			% B2M KO	
Group	**sgRNA**	**mRNA(s)**	**Mean**	**SD**
1	None	None	0.36	0.32
2	G015995	Cas9	95.27	0.59
3	G015995	Cas9, UGI	95.45	0.44
4	G015995	BC22n, UGI	99.37	0.21

[00615] On day 7 post-electroporation, T cells were also subjected to lysis, PCR amplification of the B2M locus and subsequent NGS analysis, as described in Example 1. **Table 44** and **Fig. 28** show B2M editing levels in 32 samples (8 replicates per group) following B2M editing.

Table 44 – Percent B2M editing of T cells harvested 7 days post-treatment with different mRNA combinations and B2M sgRNA G015995 (Indels = insertions or deletions).

Group	Sample	%C to T	%C to A/G	%Indels
1 - Electroporation only	1	0.35	1.47	0.17
	2	0.31	1.35	0.14
	3	0.29	1.49	0.14
	4	0.29	1.48	0.18
	5	0.30	1.52	0.20
	6	0.27	1.51	0.21
	7	0.31	1.55	0.21
	8	0.31	1.57	0.21

Group	Sample	%C to T	%C to A/G	%Indels
2 - Cas9 + G015995	9	0.08	0.10	96.96
	10	0.18	0.07	96.47
	11	0.18	0.08	97.57
	12	0.16	0.04	97.59
	13	0.16	0.06	97.53
	14	0.10	0.05	97.73
	15	0.43	0.11	96.12
	16	0.15	0.09	97.36
3 - Cas9 + UGI + G015995	17	0.09	0.08	97.05
	18	0.42	0.17	96.70
	19	0.23	0.16	96.70
	20	0.28	0.04	97.20
	21	0.20	0.03	97.60
	22	0.06	0.09	97.14
	23	0.32	0.06	97.38
	24	0.16	0.05	97.36
4 - BC22n + UGI + G015995	25	94.29	2.68	0.99
	26	93.74	3.25	1.19
	27	93.61	3.14	0.84
	28	93.94	2.93	1.01
	29	94.47	2.86	1.17
	30	94.55	2.53	1.45
	31	94.63	2.40	0.79
	32	94.14	2.79	0.78

Example 17.4. Single T cell isolation, lysis, whole genome amplification and sequencing

[00616] One sample from each group of 8 replicates was randomly chosen for single cell isolation, whole genome amplification and sequencing. Frozen cells from samples 12, 21 and 30 (**Table 44**) were transferred to a contract research organization (Singulomics Corporation, Inc.) where 10 single T cells were isolated from each sample. These single T cells were lysed, and their genomes were amplified using multiple displacement amplification

(MDA), according to previously published methods (Dong et al., Nat Methods, 2017). Amplified genomes were subjected to PCR amplification of the B2M locus and subsequent NGS analysis, as described in Example 1, aiming to confirm the edited genotype in single T cells. From each group of 10 single cells, 6 DNA samples were converted into whole genome sequencing libraries using the KAPA HyperPlus kit (Roche, Cat. 07962410001), following the manufacturer's protocol. The resulting 18 libraries were sequenced in an Illumina NovaSeq 6000 platform using an S4 reagent kit v1.5 (Illumina, Cat. 20028312).

Example 17.5. Data processing for single nucleotide variant (SNV) analysis

[00617] Paired-end reads were aligned to human genome GRCh38 with BWA-MEM v0.7.17 (Li, 2013). PCR duplicates were removed with Picard MarkDuplicates v2.19.0 (Broad Institute, 2019). Subsequently, base scores were corrected using GATK BaseRecalibrator and ApplyBQSR v4.1.8.1 (Auwera et al., 2013; DePristo et al., 2011; McKenna et al., 2010). Variants were called using DeepVariant v1.0.0 (Poplin et al., 2018). Relative C to T frequencies were calculated by dividing the total number of C to T variants by the total number of SNVs for each sample. Treatment groups were compared using unpaired t-tests and the statistical significance was determined using the Holm-Sidak method with an alpha of 0.05, without assuming a consistent standard deviation.

[00618] Compared to samples treated with Cas9 mRNA, T cells electroporated with both Cas9 and UGI mRNAs or those treated with BC22n and UGI mRNAs, showed no statistically significant ($p < 0.05$) increase in the frequency of C to T transitions in amplified genomic DNA, demonstrating the absence of detectable homology-independent cytosine deamination events in these samples (**Table 45** and **Fig. 29**).

Table 45. Percent C to T transitions of all single nucleotide variants (SNVs) in amplified genomic DNA from single T cells harvested 7 days post-electroporation with different mRNA combinations and B2M targeting sgRNA G015995 (SD = standard deviation; NA = not applicable).

		% C to T SNVs			
sgRNA	**mRNA(s)**	**Mean**	**SD**	**p value (vs. Cas9)**	**p value (vs. Cas9, UGI)**
G015995	**Cas9**	**16.32**	**0.13**	**NA**	**0.21**
G015995	**Cas9, UGI**	**16.46**	**0.21**	**0.21**	**NA**

G015995	BC22n, UGI	16.47	0.15	0.11	0.96

Example 18. Whole Genome Sequencing of Clonally Expanded eHap1 Cells Following B2M Editing

Example 18.1. eHap1 cell culture

[00619] Fully haploid, engineered Hap1 (eHap1) cells were obtained commercially (Horizon Discovery Cat. C669), and cells were cultured in IMDM growth media composed of Iscove's Modified Dulbecco's medium (Thermofisher Cat. 12440053) supplemented with 10% (v/v) of fetal bovine serum (Thermofisher Cat. A3840001) and 1x of Penicillin-Streptomycin (Thermofisher Cat. 15140122). Upon thawing, eHap1 cells were cultured for 48 hours at 37°C and seeded 24 hours prior to LNP treatments at the density of 6 x 10^5 cells/well in 6-well plates, which were incubated at 37°C until treatment.

Example 18.2. eHap1 editing

[00620] The B2M-targeting sgRNA G015991 and mRNAs encoding Cas9 (SEQ ID NO: 11), BC22n (SEQ ID NO: 2), UGI (SEQ ID NO: 26) or BE4MAX (SEQ ID NO: 32) were formulated as individual RNA species in LNPs as described in Example 1. LNPs were applied to T cells in different combinations with at the following concentrations of total RNA: 0.104 µg/mL editor mRNA; 0.55 µg/mL UGI mRNA; 0.4175 µg/mL sgRNA. Different LNP combinations (**Table 46**) were pre-mixed in IMDM growth media supplemented with 10 µg/mL of recombinant human ApoE3 (Peprotech Cat. 350-02) and incubated at 37°C for 15 minutes. The culture media of eHap1 cells was removed and each well received 3 mL of LNP mixture. Untreated controls received 3 mL of IMDM growth media supplemented with 10 µg/mL of ApoE3 without LNPs. Cells were incubated for 24 hours at 37°C and the media was removed and replaced by IMDM growth media.

[00621] Three and five days after treatment, eHap1 cells were detached, re-seeded at a lower density, and returned to the 37°C incubator. At the 5-day timepoint, a fraction of cells was stained with anti-B2M antibodies (BioLegend Cat. No. 316304) to evaluate the loss of B2M expression by flow cytometry. These cells were incubated for 30 min at 4°C with anti-B2M antibodies diluted 1:200 in cell staining buffer (BioLegend® Cat. No. 420201). Cells were subsequently washed, processed on a Cytoflex flow cytometer (Beckman Coulter) and analyzed using the FlowJo software package. eHap1 cells were gated based on size, shape and B2M expression. **Table 46** and **Fig. 30** show B2M protein expression levels in edited eHap1 cells.

Table 46 – Mean percent B2M negative eHap1 cells following editing at B2M with different mRNA combinations.

		% B2M KO	
sgRNA	**mRNA(s)**	**Mean**	**SD**
None	None	0.30	0.00
G015991	Cas9	71.93	1.93
G015991	Cas9, UGI	75.20	0.46
G015991	BC22n	74.87	0.67
G015991	BC22n, UGI	91.63	1.39
G015991	BE4MAX	44.00	2.54
G015991	BE4MAX, UGI	55.70	1.66

[00622] Seven days after treatment, eHap1 cells were detached, the replicates from each treatment group were pooled and the resulting cell suspension was centrifuged. A small fraction of cells from each pool was collected for bulk NGS sequencing as described in Example 1. **Table 47** and **Fig. 31** show B2M editing levels in these samples. The remaining eHap1 cells were mixed with antibodies against B2M (BioLegend Cat. No. 316304), diluted at 1:200 (v/v) in cell staining buffer (BioLegend® Cat. No. 420201) and incubated for 30 min at 4°C. Cells were subsequently washed and processed in the MA900 fluorescence-activated cell sorting (FACS) instrument (Sony Biotechnology). Single cells negative for B2M expression were individually seeded in 96-well plates. For untreated controls, single cells with positive signal for B2M expression were plated instead. Single cells were incubated for 10 days at 37°C to allow the establishment and expansion of single-cell clones.

Table 47 – B2M editing results (NGS sequencing) of eHap1 cells following treatment with different mRNA combinations and B2M sgRNA G015991 (Indels = insertions or deletions).

		NGS Sequencing		
sgRNA	**mRNA(s)**	**%C to T**	**%C to A/G**	**%Indel**
None	None	0.19	0.24	0.27
G015991	Cas9	0.18	0.35	89.90
G015991	Cas9, UGI	0.22	0.23	93.71
G015991	BC22n	12.78	40.92	35.39
G015991	BC22n, UGI	93.08	0.55	1.24
G015991	BE4MAX	25.51	15.65	16.17
G015991	BE4MAX, UGI	63.57	0.28	0.75

Example 18.3. Clonal expansion, extraction of genomic DNA and whole genome sequencing

[00623] After 10 days in culture, clones were visually inspected under an inverted fluorescence microscope. Twelve clones from each treatment group were transferred to 6-well plates to enable further expansion. A small fraction of cells from each clone was collected for NGS sequencing as described in Example 1, while the remaining cells were seeded in 6-plates and cultured at 37°C until confluence was reached. Based on their on-target editing outcomes, 5 clones per group were selected for whole genome sequencing. Cells from all clonal populations, in addition to one non-clonal sample of eHap1 cells, were lysed and their genomic DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen Cat. 69504) following the manufacturer's protocol. DNA samples were converted into whole genome sequencing libraries using the KAPA HyperPlus kit (Roche, Cat. 07962410001), following the manufacturer's instructions. The resulting 36 libraries were sequenced in an Illumina NovaSeq 6000 platform using an S4 reagent kit v1.5 (Illumina, Cat. 20028312).

Example 18.4. Data processing for single nucleotide variant (SNV) analysis

[00624] Reads from each sample were first aligned to the human genome build hg38 with bwa (v0.7.17) (Li, 2013; Li and Durbin, 2010). Alignments were processed with samtools (v1.11) modules fixmate, sort and markdup, consecutively (Kumaran et al., 2019; Li et al., 2009). Variations were called from processed alignments using DeepVariant (v1.0.0) (Poplin et al., 2018). Variants from each sample were then merged using GLnexus (v1.3.1) (Lin et al., 2018; Yun et al., 2021). Variants that occurred in the non-clonal sample of eHap1 cells were excluded from all clonal samples. Variants with read depth below 10 or genotype quality score below 15 were ignored as well. Relative C to T frequencies were calculated by dividing the total number of C to T variants by the total number of SNVs for each sample. Treatment groups were compared using unpaired t-tests and the statistical significance was determined using the Holm-Sidak method with an alpha of 0.05, assuming a consistent standard deviation.

Compared to untreated controls, eHap1 cells treated with Cas9, BC22n or BE4MAX mRNA, in the absence or presence of UGI mRNA, did not show a statistically significant ($p < 0.05$) increase in the frequency of C to T transitions in the genome, demonstrating the absence of detectable homology-independent cytosine deamination events in these samples (**Table 48** and **Fig. 32**).

Table 48 – Percent C to T transitions among all single nucleotide variants (SNVs) in clonally expanded eHap1 cells following editing at B2M with different mRNA combinations (NA = not applicable).

		% C to T SNVs			
sgRNA	**mRNA(s)**	**Mean**	**SD**	**p value (vs. Group 1)**	**Significant? (p<0.05?)**
None	None	12.52	0.43	NA	NA
G015991	Cas9	12.16	0.09	0.105	no
G015991	Cas9, UGI	12.04	0.42	0.112	no
G015991	BC22n	12.44	0.53	0.820	no
G015991	BC22n, UGI	12.08	0.84	0.336	no
G015991	BE4MAX	12.19	0.48	0.284	no
G015991	BE4MAX, UGI	12.59	0.48	0.789	no

Example 19. Simultaneous quadruple edits with BC22n or Cas9 in T cells after delivery via electroporation or LNP

[00625] To assess the amount of structural genomic changes associated with delivery conditions and editing by Cas9 or base editor, T cells treated with electroporation to deliver RNP or lipid nanoparticles (LNP) to deliver four guides and either Cas9 or BC22n were analyzed for cell viability, DNA double-stranded breaks, editing, surface protein expression, and chromosomal structural.

Example 19.1. T cell preparation

[00626] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. No. 130-070-525) on the LOVO device. T cells were isolated via negative selection using EasySep™ Human T Cell Isolation Kit (Stemcell Cat. No. 17951). T cells were aliquoted into vials and cryopreserved in a 1:1 formulation of Cryostor® CS10 (StemCell Technologies Cat. No. 07930) and Plasmalyte A (Baxter Cat. No. 2B2522X) for future use.

[00627] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in OpTmizer-based media containing CTS OpTmizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512) 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml recombinant human

interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/ml recombinant human interleukin 15 (Peprotech, Cat. 200-15). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec) in this media for 72 hours, at which time they were washed and plated in quadruplicate for editing either by electroporation or lipid nanoparticle.

Example 19.2. Single gRNA and 4 gRNA T cell editing with lipid nanoparticles

[00628] LNPs were formulated generally as in Example 1 with a single RNA species cargo. Cargo was selected from an mRNA encoding BC22n, an mRNA encoding Cas9, an mRNA encoding UGI, sgRNA G015995 targeting B2M, sgRNA G016017 targeting TRAC, sgRNA G016200 targeting TRBC or sgRNA G016086 targeting CIITA. Each LNP was incubated in OpTmizer-based media with cytokines as described above supplemented with 20 ug/ml recombinant human ApoE3 (Peprotech, Cat. 350-02) for 15 minutes at 37°C. Seventy-two hours post activation, T cells were washed and suspended in OpTmizer media with cytokines without human serum. For single sgRNA editing conditions, pre-incubated LNP mix was added to the each well of 100,000 cells to yield a final concentration of 2.3 ug/ml editor mRNA (BC22n or Cas9), 1.1 ug/mL UGI and 4.6 ug/uL G016017. For four-plex sgRNA editing LNP mix was added to the each well of 100,000 cells to yield a final concentration of 2.3 ug/ml editor mRNA (BC22n or Cas9), 1.1 ug/mL UGI, 1.15 ug/uL G015995, 1.15 ug/uL G016017, 1.15 ug/uL G016200 and 1.15 ug/uL G016086. A control group including unedited T cells (no LNP) was also included. At 16 hours post-delivery, a subset of cells was used to measure cell viability and another subset of cells was processed for imaging of γH2AX foci. The remaining T cells continued to expand in culture. Media was changed 5 days and 8 days after activation and on the eleventh day post activation, cells were harvested for analysis by NGS, flow cytometry and UnIT. NGS was performed as in Example 1.

Example 19.3. Single gRNA and 4 gRNA T cell editing with mRNA electroporation

[00629] Electroporation was performed 72 hours post activation. sgRNA G015995 (SEQ ID NO: 182) targeting B2M, sgRNA G016017 (SEQ ID NO: 184) targeting TRAC, sgRNA G016200 (SEQ ID NO:801) targeting TRBC and sgRNA G016086 (SEQ ID NO: 586) were denatured for 2 minutes at 95°C before cooling at room temperature for 10 minutes. T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x

10^6 T cells/mL in P3 electroporation buffer (Lonza). For single sgRNA editing conditions, 1 x 10^5 T cells were mixed with 40 ng/uL of editor mRNA (BC22n or Cas9), 10 ng/uL of UGI mRNA and 80 pmols of sgRNA in a final volume of 20 uL of P3 electroporation buffer. For four-plex sgRNA editing conditions, 1 x 10^5 T cells were mixed with 40 ng/uL of editor mRNA (BC22n or Cas9), 10 ng/uL of UGI mRNA and 20 pmols of the four individual sgRNA in a final volume of 20 uL of P3 electroporation buffer. This mix was transferred in quadruplicate to a 96-well Nucleofector™ plate and electroporated using a manufacturer's pulse code. Electroporated T cells were rested in 80 ul of OpTmizer-based media with cytokines before being transferred to a new flat-bottom 96-well plate. A control group including unedited T cells (no EP) was also included. At 16 hours post-delivery, a subset of cells was used to measure cell viability and another subset of cells was processed for imaging of γH2AX foci.

Example 19.4. Relative viability via Cell Titer Glo

[00630] Sixteen hours post electroporation or lipid nanoparticle delivery 20 uL of control or edited cells were removed from original plate and added to a new flat-bottom 96-well plate with black walls (Corning Cat. 3904). CellTiter-Glo® 2.0 (Promega Cat. G9241) was added and samples were processed according to manufacturer's protocol. Relative luminescence units (RLU) were readout by the CLARIstar plus (BMG Labtech) plate reader with gain set at 3600. Relative viability as shown in **Table 49** and **Fig. 33** was calculated by dividing all sample RLU by the average of untreated control RLU. All electroporation conditions had a greater than 5-fold viability drop from untreated control levels whereas LNP treatment, even with 4 guides editing simultaneously, maintained cell viability at close to untreated control samples.

Table 49 – Relative cell viability 16 hours following treatment with various editing and delivery conditions

		EP		LNP	
Editor	**Guide(s)**	**Mean**	**SD**	**Mean**	**SD**
No Editor	No guide	100.00	2.52	100.00	5.90
No Editor	G016017	13.88	1.05	89.48	1.77
Cas9	G016017	13.25	1.18	96.43	9.82
BC22n	G016017	14.78	1.37	91.20	4.67
Cas9	4 guides	13.45	0.65	98.30	4.80
BC22n	4 guides	14.38	1.49	103.40	2.75

Example 19.5. Staining, imaging and quantification of γH2AX foci

[00631] 16 hours post electroporation or lipid nanoparticle delivery T cells were cytospun to a slide using Cytospin 4 (Thermo Fisher). After 5 min pre-extraction in PBS/0.5% Trion X-100 on ice, cells were fixed in 4% paraformaldehyde for 10 min. Then, cells were washed in PBS several times and blocked in PBS/0.1% TX-100/1% BSA for 30 min. Primary antibody (Mouse anti-phospho-Histone H2A.X (Ser139) (Millipore Cat. 05-636) was incubated in the blocking buffer at 4°C overnight. After washed in PBS/0.05% Tween-20 three times, secondary antibody (Goat anti-Mouse IgG Alexa 568 (Thermo Fisher Cat. A31556) was incubated in the blocking buffer at room temperature for 30 min. Cells were washed in PBS/0.05% Tween-20 and nuclei were counter stained with Hoechst 33342. Images were generated by confocal imaging with the Leica SP8. Image analysis was performed via a custom protocol on Thermo Scientific HCS Studio Cell Analysis Software Spot Detector module. **Table 50** and **Fig. 34** show total γH2AX spot intensity per nuclei following treatment with stated editing and delivery conditions. EP Cas9 with 4 guides samples showed a significant increase in gH2AX foci per nuclei over LNP Cas9-4 guide samples.

Table 50 – Mean total γH2AX spot intensity per nuclei following treatment with various editing and delivery conditions

	No editor			**Cas9, 4 guides**			**BC22n, 4 guides**		
Delivery	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**
Untreated	134.33	95.78	5						
EP				21386.62	4336.69	3	Not done		
LNP				2550.88	562.77	3	1770.11	291.97	5

Example 19.6. Flow cytometry and NGS sequencing

[00632] On day 8 post-editing, T cells were phenotyped by flow cytometry to determine B2M, CD3 and HLA II- DR, DP, DQ protein expression as described in **Example 6** using antibodies targeting B2M-APC/Fire™ 750 (BioLegend® Cat. No. 316314), CD3-BV605 (BioLegend® Cat. No. 316314) and HLA II- DR, DP, DQ-PE (BioLegend® Cat. No. 361716). DNA samples were subjected to PCR and subsequent NGS analysis, as described in **Example 1**. **Table 51** and **Fig. 35** show percent editing at loci of interest following

treatment with LNPs. In the condition where 4 guides were delivered by LNP, percent editing is higher at each locus with BC22n than with Cas9. **Table 52** and **Fig. 36** show surface protein expression of interest following LNP treatment. Editing with BC22n resulted in a greater percentage of triple knockout cells than editing with Cas9.

Table 51 – Mean percent editing following treatment with stated editing schemes by LNP delivery.

Editor	Guide(s)	Readout Locus	C-to-T %			C-to-A/G %			Indel %		
			Mean	SD	N	Mean	SD	N	Mean	SD	N
No Editor	No guide	B2M	0.20	0.00	2	0.41	0.01	2	0.13	0.01	2
		TRAC	0.05	0.07	2	0.53	0.01	2	0.09	0.01	2
		TRBC1	0.25	0.07	2	0.86	0.09	2	0.24	0.02	2
		TRBC2	0.30	0.14	2	0.85	0.04	2	0.27	0.01	2
		CIITA	0.15	0.07	2	0.26	0.03	2	0.06	0.00	2
Cas9	G016017	TRAC	0.00	0.00	2	0.05	0.00	2	96.59	0.52	2
	4 guides	B2M	0.30	0.14	2	0.37	0.21	2	83.61	2.72	2
		TRAC	0.00	0.00	2	0.06	0.04	2	89.45	0.25	2
		TRBC1	0.00	0.00	2	0.35	0.01	2	75.98	2.09	2
		TRBC2	0.00	0.00	2	0.38	0.03	2	78.07	1.51	2
		CIITA	0.20	0.14	2	0.38	0.04	2	55.23	0.92	2
BC22n	G016017	TRAC	93.90	0.57	2	2.29	0.29	2	2.34	0.08	2
	4 guides	B2M	95.65	0.07	2	1.42	0.23	2	1.07	0.16	2
		TRAC	95.00	0.57	2	1.65	0.07	2	1.44	0.43	2
		TRBC1	92.40	0.00	2	1.55	0.18	2	1.91	0.10	2
		TRBC2	92.50	0.00	2	1.62	0.02	2	1.87	0.33	2
		CIITA	91.35	0.35	2	0.93	0.04	2	0.64	0.04	2

Table 52 – Mean percentage of cells expressing surface following treatment with stated editing schemes by LNP delivery.

Editor	Guide	Surface protein	Mean	SD	N
No Editor	No guide	CD3-	2.30	2.12	2
		B2M-	4.00	2.26	2
		HLA DP DQ DR-	60.40	1.70	2
Cas9	G016017	CD3-	97.05	0.16	2
	4 guides	B2M-	77.90	4.67	2
		CD3-	94.54	1.34	2
		HLA DP DQ DR-	70.10	4.81	2
		CD3- B2M- HLA DP DQ DR-	61.58	5.42	2
BC22n	G016017	CD3-	96.20	0.13	2
	4 guides	B2M-	97.35	0.37	2

		CD3-	97.34	0.08	2
		HLA DP DQ DR-	93.46	0.89	2
		CD3- B2M- HLA DP DQ DR-	90.63	0.89	2

Example 19.7. Measuring structural variation and translocations by UnIT

[00633] On day 8 post-editing a subset of T cells from the untreated, LNP-Cas9-4 guides and LNP-BC22n-4 guides samples were collected, spun down and resuspended in 100 uL of PBS. gDNA was isolated from the cells using DNeasy Blood & Tissue Kit (Qiagen Cat. 69504). The UnIT structural variant characterization assay was applied to these gDNA samples. High molecular weight genomic DNA is simultaneously fragmented and sequence-tagged ('tagmented') with the Tn5 transposase and an adapter with a partial Illumina P5 sequence and a 12 bp unique molecular identifier (UMI). Two sequential PCRs using a primer to P5 and hemi-nested gene specific primers (GSP) imparting the Illumina the P7 sequence to create two Illumina compatible NGS libraries per sample. Sequencing across both directions of the CRISPR/Cas9 targeted cut site with the two libraries allows the inference and quantification of structural variants in DNA repair outcomes after genome editing. If the two fragments were aligned to different chromosomes, the SV was classified as an "inter-chromosomal translocation." Structural variation results show that interchromosomal translocations are reduced to background levels when multiplex editing is being conducted by BC22n whereas Cas9 multiplex editing leads to significant increases in structural variation, as shown in **Table 53** and **Fig. 37**.

Table 53 – Mean percent interchromosomal translocations among total unique molecule identifiers following treatment with stated editing schemes by LNP delivery.

Edit	**Locus**	**Mean**	**SD**	**N**
Untreated	B2M	0.16	0.09	2
	TRAC	0.21	0.13	2
	CIITA	0.10	0.04	2
Cas9, 4 guides	B2M	2.71	0.52	2
	TRAC	1.55	0.24	2
	CIITA	0.92	0.37	2
BC22n, 4 guides	B2M	0.14	0.06	2
	TRAC	0.21	0.08	2
	CIITA	0.14	0.08	2

Example 20 – CD38 Guide RNA screening in T cells with Cas9 and BC22n

Example 20.1 T cell Preparation

[00634] T cells were edited at the CD38 locus with either Cas9 or with BC22n and UGI mRNAs to assess the editing outcomes and the corresponding loss of CD38 expression.

[00635] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and resuspended in CliniMACS PBS/EDTA buffer (Miltenyi Biotec, Cat. No. 130-070-525) on the LOVO device. T cells were isolated via positive selection using CD4 and CD8 magnetic beads (Miltenyi Biotec, Cat. No. 130- 030-401/130-030-801) using the CliniMACS Plus and CliniMACS LS disposable kit. T cells were aliquoted into vials and cryopreserved in Cryostor CS10 (StemCell Technologies, Cat. No. 07930) for future use. Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell X-VIVO 15 expansion media composed of X-VIVO 15 (Lonza, Cat. No. BE02-06Q) containing 5% (v/v) of fetal bovine serum (ThermoFisher, Cat. No. A3160902), 50 μM (1X) 2-Mercaptothanol (ThermoFisher, Cat. No. 31350010), 1% of Penicillin-Streptomycin (ThermoFisher, Cat. No. 15140122), 1 M N-acetyl L-cystine (Fisher, Cat. No. ICN19460325) diluted in phosphate buffered saline (PBS) and normalized to pH 7, supplemented with 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. No. 200-02), 5 ng/mL recombinant human interleukin-7 (Peprotech, Cat. No. 200-07) and 5 ng/mL recombinant human interleukin-15 (Peprotech, Cat. No. 200-15). T cells were activated with TransAct™ (1:100 dilution, Miltenyi Biotec, Cat. No. 130-111-160). Cells were expanded for 72 hours at 37°C prior to mRNA electroporation.

Example 20.2 T cell editing with RNA electroporation

[00636] Solutions containing mRNA encoding Cas9 protein, BC22n, or UGI were prepared in sterile water. 50 μM CD38 targeting sgRNAs were removed from their storage plates and denatured for 2 minutes at 95°C and incubated at room temperature for 5 minutes. Seventy-two hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). Messenger RNAs were prepared as described in Example 1. For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of Cas9 or BC22n mRNAs, 200 ng of UGI mRNA and 20 pmols of sgRNA as described in Table 54 in a final volume of 20 μL of P3 electroporation buffer. This mix was transferred in duplicate to a 96-well

Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 µL of X-VIVO 15 media without cytokines for 15 minutes before being transferred to new flat-bottom 96-well plates containing an additional 90 µL of X-VIVO 15 media supplemented with 2X cytokines. The resulting plates were incubated at 37°C for 10 days. To promote expansion, T cells were split at the ratios of 1:4 and 1:3 on days 3 and 6 post-electroporation, respectively, using fresh X-VIVO 15 media with 1X cytokines. On day 9 post-electroporation, cells were split 1:2 in 2 U-bottom plates and one plate was collected for NGS sequencing, while the other plate was used for flow cytometry on Day 10.

Example 20.3 Flow cytometry and NGS sequencing

[00637] On day 10 post-editing, T cells were phenotyped by flow cytometry to determine CD38 receptor expression. Briefly, T cells were incubated for 30 min at 4°C with a mixture of antibodies against CD3 (BioLegend, Cat. No. 317340), CD4 (BioLegend, Cat. No. 300537), CD8 (BioLegend, Cat. No. 344706) diluted at 1:200 and CD38 (BioLegend, Cat. No. 303546), diluted at 1:100 in cell staining buffer (BioLegend, Cat. No. 420201). Cells were subsequently washed and stained with viability antibody DAPI (BioLegend, Cat. No. 422801) diluted at 1:10,000 in cell staining buffer. Cells were then processed on a Cytoflex flow cytometer (Beckman Coulter) and analyzed using the FlowJo software package. T cells were gated based on size, shape, viability, and CD38 expression.

[00638] On day 9 DNA samples were subjected to PCR and subsequent NGS analysis, as described in Example 1. **Table 54** shows CD38 gene editing and CD38 positive results for cells edited with BC22n or Cas9.

Table 54 – Percent editing and percent of CD38 positive cells following CD38 editing with Cas9 or BC22n base editor

	Cas9 %Indels			Cas9 %CD38+			BC22n %C to T			BC22 %CD38+		
Guide ID	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
G019761	94.82	0.97	2	22.35	0.49	2	87.21	1.57	2	56.45	2.05	2
G019762	98.19	0.63	2	15.45	1.06	2	85.04	0.42	2	58.55	0.07	2
G019763	91.42	1.77	2	8.17	2.38	2	87.01	0.51	2	2.53	0.21	2
G019764	87.78	1.04	2	11.80	0.71	2	87.00	2.33	2	41.40	2.12	2
G019765	96.43	0.40	2	13.15	2.33	2	81.51	5.02	2	60.40	1.27	2
G019766	97.61	0.06	2	5.31	1.15	2	80.51	2.67	2	61.80	0.85	2
G019767	97.54	0.08	2	17.45	0.21	2	81.92	0.46	2	62.55	3.46	2
G019768	96.65	0.96	2	2.19	1.06	2	89.45	0.78	2	2.49	1.18	2

Guide ID	Cas9 %Indels			Cas9 %CD38+			BC22n %C to T			BC22 %CD38+		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
G019769	95.63	2.44	2	2.58	0.35	2	69.93	5.38	2	14.40	4.81	2
G019770	97.74	1.26	2	1.15	0.59	2	85.06	5.81	2	38.80	3.54	2
G019771	96.90	0.54	2	1.18	0.18	2	87.34	0.16	2	1.39	0.26	2
G019772	96.38	0.05	2	14.50	1.98	2	88.86	0.71	2	70.25	1.63	2
G019773	95.38	2.21	2	6.70	0.57	2	79.83	0.15	2	64.15	4.45	2
G019774	97.12	0.07	2	14.55	0.92	2	71.14	2.52	2	68.60	2.83	2
G019775	95.33	3.04	2	16.75	1.63	2	92.02	0.11	2	74.90	0.28	2
G019776	98.87	0.45	2	1.95	1.04	2	84.25	0.57	2	11.45	1.63	2
G019777	97.87	0.54	2	14.00	1.27	2	86.76	1.32	2	71.75	2.62	2
G019778	95.81	1.24	2	13.75	1.48	2	86.96	1.81	2	62.30	2.26	2
G019779	97.78	0.26	2	10.87	2.17	2	88.77	0.53	2	74.35	0.07	2
G019780	95.75	1.84	2	14.00	2.55	2	89.81	0.30	2	69.05	3.18	2
G019781	93.81	2.35	2	15.90	0.85	2	86.84	4.62	2	68.30	0.71	2
G019782	93.10	4.65	2	10.21	2.82	2	82.19	1.92	2	63.15	0.49	2
G019783	97.88	0.30	2	2.58	0.40	2	77.26	4.30	2	64.10	0.14	2
G019784	97.61	0.86	2	16.80	2.69	2	76.82	0.81	2	69.00	0.71	2
G019785	98.30	0.27	2	1.58	0.66	2	80.95	1.14	2	33.90	5.80	2
G019786	86.22	3.34	2	28.10	4.67	2	88.42	2.78	2	72.35	5.02	2
G019787	98.76	0.45	2	2.06	0.86	2	80.23	2.14	2	1.89	0.83	2
G019788	96.34	1.98	2	11.05	0.92	2	90.88	0.06	2	1.88	0.21	2
G019789	97.68	1.33	2	15.80	1.27	2	67.12	n.a.	1	52.85	1.34	2
G019790	94.58	1.87	2	12.15	0.07	2	89.81	0.22	2	77.10	2.69	2
G019791	97.79	0.90	2	3.25	0.86	2	79.56	3.97	2	9.51	0.62	2
G019792	96.99	0.13	2	9.19	0.11	2	85.20	4.14	2	64.65	0.35	2
G019793	96.13	0.29	2	19.95	2.33	2	85.55	2.40	2	71.40	2.69	2
G019794	98.30	0.28	2	12.75	0.07	2	76.28	0.67	2	31.40	0.85	2
G019795	97.56	0.59	2	2.52	0.56	2	91.40	2.37	2	1.22	0.34	2
G019796	97.70	1.27	2	7.60	2.87	2	92.29	0.28	2	6.52	0.16	2
G019797	97.04	0.26	2	9.80	2.26	2	87.45	0.46	2	3.07	0.84	2
G019798	97.56	1.17	2	4.92	0.93	2	89.68	0.89	2	59.05	1.34	2
G019799	97.65	0.22	2	8.83	1.46	2	87.59	0.90	2	66.85	0.49	2
G019800	76.71	3.24	2	25.45	4.88	2	74.46	2.84	2	16.00	3.54	2
G019801	92.76	3.74	2	6.10	2.28	2	82.69	0.52	2	33.65	4.03	2
G019802	64.21	7.62	2	33.70	5.52	2	69.14	0.45	2	73.00	0.42	2
G019803	50.46	6.67	2	43.50	5.23	2	45.01	1.82	2	55.40	1.98	2
G019804	89.70	3.71	2	7.05	3.23	2	21.88	0.42	2	67.25	2.76	2
G019805	52.39	2.21	2	40.30	0.14	2	78.69	4.52	2	9.15	1.77	2
G019806	93.47	0.40	2	4.64	0.82	2	72.68	1.54	2	21.30	4.95	2
G019807	86.45	2.17	2	12.70	1.84	2	63.32	3.90	2	65.85	0.21	2
G019808	97.15	0.70	2	2.23	1.37	2	32.87	6.70	2	64.10	1.84	2
G019809	41.04	6.38	2	56.35	4.31	2	68.30	3.73	2	27.50	5.94	2
G019810	46.14	4.29	2	52.05	5.16	2	78.63	2.98	2	10.40	2.97	2

Guide ID	Cas9 %Indels			Cas9 %CD38+			BC22n %C to T			BC22 %CD38+		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
G019811	95.15	0.67	2	14.05	1.06	2	83.21	2.11	2	4.42	2.68	2
G019812	93.95	1.06	2	4.46	2.11	2	84.23	5.13	2	24.90	2.55	2
G019813	96.79	1.07	2	2.51	0.60	2	75.68	3.13	2	5.51	1.97	2
G019814	87.26	0.64	2	11.56	2.75	2	72.31	1.65	2	13.80	2.83	2
G019815	77.53	7.29	2	18.65	3.89	2	86.69	0.87	2	8.47	0.74	2
G019816	76.89	1.37	2	38.25	6.86	2	87.17	2.00	2	8.82	2.94	2
G019817	90.53	1.09	2	5.95	1.39	2	72.59	1.92	2	33.85	1.34	2
G019818	97.71	0.21	2	1.69	0.57	2	77.99	3.46	2	14.15	3.75	2
G019819	94.33	1.87	2	3.70	2.29	2	84.14	3.95	2	2.47	0.45	2
G019820	82.74	4.09	2	21.00	3.39	2	81.28	0.66	2	14.80	2.12	2
G019821	70.98	8.64	2	26.80	8.49	2	73.39	3.46	2	11.09	3.27	2
G019822	46.59	2.28	2	50.80	5.09	2	58.43	2.69	2	73.55	0.21	2
G019823	93.48	2.48	2	4.04	1.94	2	79.65	0.80	2	24.45	6.01	2
G019824	95.90	0.76	2	6.32	1.92	2	89.11	1.03	2	74.75	3.18	2
G019825	29.68	7.06	2	56.10	0.57	2	39.49	6.14	2	66.20	3.54	2
G019826	92.62	1.03	2	3.74	1.05	2	69.37	5.37	2	61.70	0.42	2
G019827	39.32	4.26	2	54.80	2.55	2	48.52	0.86	2	38.15	1.20	2
G019828	87.90	3.26	2	9.98	1.16	2	80.25	0.57	2	7.40	2.33	2
G019829	31.95	2.81	2	63.40	1.98	2	76.10	2.57	2	14.50	2.69	2
G019830	81.01	6.21	2	16.45	3.32	2	83.57	1.78	2	4.29	1.12	2
G019831	98.29	0.30	2	0.91	0.25	2	71.43	0.65	2	2.05	0.84	2
G019832	62.85	1.73	2	34.70	3.54	2	90.87	2.45	2	71.05	6.01	2
G019833	92.66	0.94	2	4.28	0.82	2	75.96	0.45	2	18.40	2.26	2
G019834	94.43	0.37	2	2.75	0.76	2	76.23	0.91	2	69.00	2.83	2
G019835	42.83	n.a.	1	49.95	5.73	2	51.27	8.38	2	43.95	6.72	2
G019836	No data			46.05	5.02	2	31.44	6.72	2	75.10	0.00	2
G019837	43.86	6.82	2	49.85	5.44	2	63.03	1.17	2	51.50	1.70	2
G019838	75.85	0.40	2	24.25	0.64	2	74.87	1.14	2	19.05	2.62	2
G019839	97.44	n.a..	1	1.82	0.21	2	69.11	4.60	2	40.55	2.05	2
G019840	59.58	4.30	2	37.90	0.85	2	47.06	4.04	2	73.05	1.06	2
G019841	96.14	n.a.	1	3.60	1.68	2	69.28	2.55	2	27.15	2.33	2
G019842	33.72	6.77	2	58.65	2.05	2	29.32	2.31	2	71.50	3.82	2
G019843	6.02	0.04	2	73.15	0.49	2	20.85	1.36	2	70.75	0.35	2
G019844	85.26	12.45	2	6.30	0.47	2	76.89	3.63	2	39.45	3.75	2
G019845	36.76	2.44	2	55.25	3.18	2	62.90	1.02	2	60.15	2.62	2
G019846	25.85	4.33	2	64.80	2.69	2	57.55	1.80	2	63.95	4.88	2
G019847	58.00	5.06	2	37.85	2.19	2	71.58	0.74	2	42.55	5.16	2
G019848	47.39	6.18	2	46.90	5.09	2	71.78	4.09	2	67.15	3.32	2

No data means sample had technical fail in all replicates. N.A. is used when only one replicate was available and thus it was not applicable to calculate a standard deviation.

Example 21. Base editing with Nme2Cas9 base editor and chemically modified sgRNA in HepG2 cells

[00639] Base editor constructs comprising an APOBEC3A deaminase domain fused to Nme2Cas9 D16A nickase were tested for base conversion efficiency with various guide designs in HepG2 cells.

[00640] HepG2 cells constitutively overexpress solute carrier family 10 member 1 (SLC10A1) (HepG2-NTCP, Seeger et al. Mol Ther Nucleic Acids. 2014 Dec; 3(12): e216) were thawed and resuspended in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) (Media Y) followed by centrifugation. The supernatant was discarded and the cells were resuspended in Media Y and plated at a density of 25,000 cells per well in a 96-well collagen coated plate (Corning, Cat. 354407) in 100uL of Media Y.

[00641] Nme2Cas9 base editor mRNAs were prepared by in vitro transcription essentially as described in Example 1 from plasmids encoding SEQ ID No: 304 (2XNLS N-terminal, 1xC-terminal NLS Nme2Cas9 base editor), SEQ ID No: 310 (2XNLS N-terminal, NLS Nme2Cas9 base editor), and SEQ ID No: 301 (1X C-term NLS Nme2Cas9 base editor). SpyCas9 mRNA and uracil glycosylase inhibitor (UGI) mRNA (SEQ ID No: 34) were transcribed from plasmids using the same method.

[00642] Chemically modified NmeCas9 sgRNAs targeted to NTCP, with different PAM sequences, (G020927, G020928) or VEGF (G020073) and SpyCas9 sgRNA targeted to NTCP (G020929) were synthesized using routine methods.

[00643] The tested NmeCas9 sgRNAs targeting NTCP include a 24 nucleotide guide sequence (as represented by N) and a guide scaffold as follows: mN*mNNNNNNNNmNNNmNNNNNNNNNNNNmGUUGmUmAmGmCUCCCmUmGmAmAmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmCmGmAmAmAmGmAmUGUGCmCGCmAmAmCmGCUCUmGmCCmUmUmCmUGmGCmAmUC*mG*mU*mU (SEQ ID NO: 522), where A, C, G, U, and N are adenine, cytosine, guanine, uracil, and any ribonucleotide, respectively, unless otherwise indicated. An m is indicative of a 2'O-methyl modification, and an * is indicative of a phosphorothioate linkage between the nucleotides. Unmodified and modified versions of the guides are provided in Table 5C. Guide RNA, editor mRNA, and UGI mRNA were mixed at a 1:1:1 weight ratio with premixed transfection reagent containing Lipid A, cholesterol, DSPC, and PEG2k-DMG in a 50:38:9:3 molar ratio, respectively. Reagents were combined at a lipid amine to RNA phosphate (N:P) molar ratio of about 6.0. RNA-lipid mixture was mixed approximately 1:1 with 10% FBS media and

incubated for 10 minutes. Post-incubation, the cells were treated with the RNA-lipid mixture in an 8-point, 2-fold serial dilution starting at 400 ng total editor RNA per well.

[00644] At 72 hours post-treatment, cells were lysed for NGS analysis as provided in Example 1.

[00645] Dose response of editing efficiency to guide concentration was performed in triplicate. **Table 55** shows mean editing percentages calculated [at each guide concentration and a calculated EC50 value. The target site in VEGFA is prone to indel formation due to high GC content. All editor mRNAs achieved the same maximum C to T editing. There were slight differences in EC50 where SEQ ID No: 310 mRNA outperformed SEQ ID No: 304 and SEQ ID No: 301.

Table 55. Mean editing percentages in HepG2-NTCP cells at VEGFA locus (G020073) and Nme2Cas9 base editor

mRNA	Nme-base editor (ng)	C-to-T %				C-to-A/G %			Indel %		
		Mean	SD	n	EC50	Mean	SD	n	Mean	SD	n
SEQ ID No: 304	0	0.20	0.14	2	6.86	0.80	0.00	2	0.70	0.14	2
	6.25	36.05	1.63	2		2.30	0.28	2	14.25	0.21	2
	12.5	49.70	1.13	2		2.95	0.49	2	16.35	1.06	2
	25	57.60	0.42	2		2.75	0.07	2	17.45	0.21	2
	50	67.15	1.34	2		2.65	0.07	2	17.90	0.14	2
	100	70.50	0.99	2		2.65	0.07	2	18.80	0.28	2
	200	72.75	0.49	2		2.90	0.57	2	18.75	0.49	2
	400	72.00	0.14	2		3.10	0.14	2	20.40	0.99	2
SEQ ID No: 310	0	0.50	0.00	1	3.86	1.00	0.00	1	0.60	0.00	1
	6.25	49.60	3.25	2		2.10	0.14	2	14.05	0.21	2
	12.5	57.05	0.78	2		2.50	0.42	2	16.50	0.14	2
	25	66.95	0.49	2		2.40	0.00	2	16.25	0.21	2
	50	71.45	0.35	2		2.40	0.00	2	17.25	0.07	2
	100	71.60	0.85	2		2.55	0.35	2	17.90	0.14	2
	200	73.70	0.28	2		2.65	0.07	2	19.00	0.28	2
	400	73.80	0.14	2		3.05	0.35	2	18.70	0.42	2
SEQ ID No:	0	0.55	0.49	2	4.84	1.05	0.21	2	0.55	0.07	2

301	6.25	45.55	0.21	2		2.60	0.28	2	17.15	0.64	2
	12.5	56.40	1.84	2		2.80	0.00	2	19.40	0.14	2
	25	64.95	1.06	2		3.15	0.35	2	19.40	1.70	2
	50	70.70	0.28	2		3.25	0.21	2	20.00	0.71	2
	100	72.20	2.83	2		2.75	0.21	2	19.80	1.56	2
	200	70.90	0.99	2		3.15	0.35	2	19.70	0.85	2
	400	71.80	0.14	2		3.45	0.21	2	20.95	0.35	2

Example 22. Base editing with chemically modified sgRNA in PMH

[00646] Base editor constructs comprising an APOBEC3A deaminase domain fused to Nme2Cas9 nickase were tested for base conversion efficiency with various guide designs in primary mouse hepatocytes (PMH).

[00647] PMH (In Vitro ADMET Laboratories, cat# MC148) were thawed and resuspended in 50 mL Cryopreserved Hepatocyte Recovery Media (CHRM) (Invitrogen, CM7000) followed by centrifugation. Cells were resuspended in hepatocyte medium with plating supplements: Williams' E Medium Plating Supplements with FBS content (Gibco, Cat. A13450). Cells were pelleted by centrifugation, resuspended in media and plated at a density of 20,000 cells/well on Bio-coat collagen I coated 96-well plates (Corning # 354407). Plated cells were allowed to settle and adhere for 4-6 hours in a tissue culture incubator at 37°C and 5% CO2 atmosphere. After incubation cells were checked for monolayer formation and were washed once and plated with 100 uL hepatocyte maintenance medium: Williams' E Medium (Gibco, Cat. A12176-01) plus supplement pack (Gibco, Cat. CM3000).

[00648] Nme2Cas9 base editor mRNAs, SEQ ID No: 304, SEQ ID No: 310, and SEQ ID No: 301; and uracil glycosylase inhibitor (UGI) mRNA (SEQ ID No: 34) were prepared as described in Example 1 and paired with a series of chemically modified sgRNA targeted to mouse TTR and screened at a single dose of 128 ng of base editor mRNA. At 72 hours post-treatment, cells were lysed for NGS analysis as provided in Example 1. The mean editing of representative guides (ratio of edit types) are shown in Table 56.

Table 56 – Mean editing percentages in PMH cells using modified gRNAs targeting the TTR locus and an Nme2Cas9 base editor. "n/a" means SD is not applicable when only 1 replicate is reported.

mRNA	Guide	C-to-T %			C-to-A/G %			Indel %		
		Mean	SD	N	Mean	SD	N	Mean	SD	N
SEQ ID No: 304	G021237	69.00	5.94	2	3.65	1.63	2	9.00	3.39	2
	G021249	47.15	2.33	2	1.20	0.14	2	1.45	1.34	2
	G021321	7.25	0.92	2	0.25	0.21	2	91.25	0.92	2
SEQ ID No: 310	G021237	76.75	1.34	2	4.35	1.48	2	7.95	1.34	2
	G021249	54.05	4.17	2	1.20	0.28	2	1.60	0.42	2
	G021321	0.50	n/a	1	0.10	n/a	1	99.00	n/a	1
SEQ ID No: 301	G021237	73.55	5.44	2	5.15	2.47	2	12.50	1.70	2
	G021249	53.30	5.52	2	1.05	0.21	2	2.15	1.63	2
	G021321	7.05	1.20	2	0.55	0.21	2	90.40	2.97	2

Example 23. *In vivo* base editing with gRNA

[00649] The editing efficiency of the modified gRNAs with different mRNAs were tested with a Nme2Cas9 base editor construct in the mouse model.

[00650] LNPs were generally prepared as described in Example 1 with a single RNA species as cargo. The LNPs used were prepared with Lipid A, cholesterol, DSPC, and PEG2k-DMG in a 50:38:9:3 molar ratio, respectively. The LNPs were formulated with a lipid amine to RNA phosphate (N:P) molar ratio of about 6. The LNPs were formulated as described in Example 1, except that each component, guide RNA, or mRNA was formulated individually into an LNP, and the LNP were mixed prior to administration as described in Table 57. For Nme2Cas9 and Nme2Cas9 base editor samples, LNPs were mixed at a ratio of 2:1 by weight of gRNA to editor mRNA cargo. For SpyCas9 base editor samples, LNPs were mixed at a ratio of 1:2 by weight of gRNA to editor mRNA cargo. Dose, as indicated in Table 57 and Fig. 38, is calculated based on the combined RNA weight of gRNA and editor mRNA. Base editor samples were treated with an additional 0.03 mpk of UGI mRNA. Transport and storage solution (TSS) used in LNP preparation was dosed in the experiment as a vehicle-only negative control.

[00651] The tested NmeCas9 gRNA (G021844) including linkers has the following modification pattern of $(N)_{24}$

GUUGmUmAmGmCUCCCmUmUmC(L1)mGmAmCmCGUUmGmCUAmCAAU*AAGmGmCCmGmUmC(L1)mGmAmUGUGCmCGmCAAmCGCUCUmGmCC(L1)GGCAUCG*mU*mU (SEQ ID NO: 519),

where A, C, G, U, and N are adenine, cytosine, guanine, uracil, and any ribonucleotide,

respectively, unless otherwise indicated. An m is indicative of a 2'O-methyl modification, and an * is indicative of a phosphorothioate linkage between the nucleotides. Unmodified and modified versions of the guides are provided in Table 5C (Sequence Table) .

[00652] CD-1 female mice, ranging 6-10 weeks of age were used in each study involving mice (n = 5 per group, except TSS control n= 4). Formulations were administered intravenously via tail vein injection according to the doses listed in Table 57. Animals were periodically observed for adverse effects for at least 24 hours post-dose. Six days after treatment, animals were euthanized by cardiac puncture under isoflurane anesthesia; liver tissue were collected for downstream analysis. Liver punches weighing between 5 and 15 mg were collected for isolation of genomic DNA and total RNA. Genomic DNA was extracted using a DNA isolation kit (ZymoResearch, D3010) and samples were analyzed with NGS sequencing as described in Example 1. The editing efficiency for LNPs containing the indicated gRNAs are shown in Table 57 and illustrated in Fig. 38.

Table 57 – Mean percent editing in mouse liver.

Sample	Dose (mg/kg)	C-to-T %			C-to-A/G %			Indel %		
		Mean	SD	n	Mean	SD	n	Mean	SD	n
TSS	0	0.00	0.00	4	0.10	0.00	4	0.08	0.05	4
SEQ ID NO: 340+ G021844 (Nme2Cas9 +pgRNA)	0.03	0.00	0.00	5	0.08	0.04	5	40.88	14.16	5
	0.1	0.00	0.00	5	0.02	0.04	5	66.02	5.01	5
SEQ ID No: 310 + SEQ ID No: 34 + G021844 (Nme2Cas9 base editor + UGI + pgRNA)	0.03	25.60	5.28	5	3.50	0.76	5	11.14	2.18	5
	0.1	46.34	1.53	5	5.74	0.33	5	13.52	0.90	5
SEQ ID No: 1 + SEQ ID No: 34 + G000502 (SpyBC22n + UGI + sgRNA)	0.03	9.28	2.82	5	0.94	0.54	5	7.34	1.61	5
	0.1	30.72	8.51	5	2.86	0.23	5	15.60	2.58	5

Example 24. C to T deaminase base editing screen

[00653] Candidate deaminase-Cas9 fusion constructs were evaluated on efficiency of C to T conversion activity. All experimental deaminase C to T conversion activity were compared against construct BC27 encoding BE3 (Komor AC, Kim YB, Packer

MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016;533(7603):420-424.) In total, 56 deaminase-Cas9 fusion constructs as mentioned in Table 58 were each evaluated in triplicate against sg005883 in a single experiment. The constructs were designed and methods performed as described in Example 1A.

Table 59 and Fig. 39 shows the percent of total reads containing at least 1 cytosine to thymidine conversion across all base editor constructs tested.

Table 58- Deaminases tested in base editor activity screens

Base Editor Construct ID	UniProt for Deaminase Domain	Deaminase Gene Name	Deaminase Species	Amino acid sequence (SEQ ID NO)
BC01	A0A0N4SVL6	Apobec1	*M. musculus*	960
BC02	A0A0N4SW85	Apobec1	*M. musculus*	961
BC03	A0QT11	cdd	*M. smegmatis*	962
BC04	Q3JJN0	cdd	*B. pseudomallei*	963
BC05	B2HDU6	cdd	*M. marinum*	964
BC06	Q81LT6	cdd2	*B. anthracis*	965
BC07	P9WPH3	cdd	*M. tuberculosis*	966
BC08	A0A0H3MRL9	cdd	*M. leprae*	967
BC09	P19079	cdd	*B. subtilis*	968
BC10	Q06549	CDD1	*S. cerevisiae*	969
BC11	Q81Y61	GBAA	*B. anthracis*	970
BC12	P56389	Cda Cdd	*M. musculus*	971
BC13	P32320	CDA	*H. sapiens*	972
BC14	Q12178	FCY1	*S. cerevisiae*	973
BC15	Q22922	cdd	*C. elegans*	974
BC16	Q20628	cdd	*C. elegans*	975
BC17	Q9NRW3	APOBEC3C	*H. sapiens*	976
BC18	A0A087WX48	Apobec3DE	*H. sapiens*	977
BC19	Q82Y41	NE0047	*N. europaea*	978
BC20	Q9GZX7	AICDA AID	*H. sapiens*	979
BC21	Q9WVE0	Aicda Aid	*M. musculus*	980
BC22	P31941	APOBEC3A	*H. sapiens*	40
BC23	Q6NTF7	APOBEC3H	*H. sapiens*	981
BC24	Q9Y235	APOBEC2	*H. sapiens*	982
BC25	Q9WV35	Apobec2	*M. musculus*	983
BC26	P51908	Apobec1	*M. musculus*	984
BC27	P38483	Apobec1	*R. norvegicus*	41
BC28	Q9TUI7	APOBEC1	*M. domestica*	985
BC29	P41238	APOBEC1	*H. sapiens*	986
BC30	P47855	APOBEC1	*O. cuniculus*	987

Base Editor Construct ID	UniProt for Deaminase Domain	Deaminase Gene Name	Deaminase Species	Amino acid sequence (SEQ ID NO)
BC31	Q8TWU6	MK0935	*M. kandleri*	988
BC32	P0ABF6	cdd	*E. coli*	989
BC33	Q9KSM5	cdd	*V. cholerae*	990
BC34	Q3TI69	Apobec3	*M. musculus*	991
BC35	O65896	CDA1	*A. thalia*	992
BC36	Q7YR23	APOBEC3G	*M. mulatta*	993
BC37	Q8IUX4	APOBEC3F	*H. sapiens*	994
BC38	Q7YR25	APOBEC3G	*C. aethiops*	995
BC39	Q9UH17	APOBEC3B	*H. sapiens*	996
BC40	M1GSK9	APOBEC3G	*M. mulatta*	997
BC41	Q9HC16	APOBEC3G	*H. sapiens*	998
BC42	Q7YR24	APOBEC3G	*P. troglodytes*	999
BC43	Q96AK3	APOBEC3D	*H. sapiens*	1000
BC44	H3BJQ3	Apobec3	*M. musculus*	1001
BC45	H3BKE5	Apobec3	*M. musculus*	1002
BC46	E9QMH1	Apobec3	*M. musculus*	1003
BC47	Q99J72	Apobec3	*M. musculus*	1004
BC48	H3BJ44	Apobec3	*M. musculus*	1005
BC49	H3BJL0	Apobec3	*M. musculus*	1006
BC50	G3UW39	Cdadc1	*M. musculus*	1007
BC51	P25524	codA	*E. coli*	1008
BC52	Q19Q52	APOBEC3H	*M. mulatta*	1009
BC53	Q97MB6	CA_C0282	*C. acetobutylicum*	1010
BC54	A0A0E1CHI1	KPNJ1	*K. pneumoniae*	1011
BC55	P46087	NOP2	*H. sapiens*	1012
BC56	Q08J23	NSUN2	*H. sapiens*	1013

Table 59 Percent of total reads containing at least 1 cytosine to thymidine conversion for sg005883. (n=3)

Construct	Mean	SD	Construct	Mean	SD	Construct	Mean	SD
1XNLS Cas9	0.0	0.1	BC18	6.0	1.0	BC38	26.8	1.5
3XNLS Cas9	0.0	0.0	BC19	2.3	0.3	BC39	50.0	7.4
BC01	1.5	0.3	BC20	17.9	3.5	BC40	31.9	3.3
BC02	1.3	0.6	BC21	27.8	5.3	BC41	35.0	9.1
BC03	1.4	0.5	BC22	46.0	7.7	BC42	49.0	6.9
BC04	1.2	0.5	BC23	2.5	0.7	BC43	9.8	2.3
BC05	1.9	0.2	BC24	1.7	0.2	BC44	12.8	4.8
BC06	1.0	0.4	BC25	1.5	0.5	BC45	22.7	5.2
BC07	1.7	0.3	BC26	42.1	1.5	BC46	14.6	6.0
BC08	1.8	0.3	BC27	40.8	5.2	BC47	14.8	4.6

Construct	Mean	SD	Construct	Mean	SD	Construct	Mean	SD
BC09	1.8	0.5	BC28	42.9	5.0	BC48	31.5	4.2
BC10	2.2	0.8	BC29	20.3	1.0	BC49	13.0	4.3
BC11	1.8	0.4	BC30	23.7	5.2	BC50	1.7	0.3
BC12	2.0	0.5	BC31	1.9	0.2	BC51	1.1	0.2
BC13	1.0	0.2	BC32	1.9	0.4	BC52	8.0	0.5
BC14	2.1	0.6	BC33	2.2	0.6	BC53	1.1	0.2
BC15	0.9	0.2	BC34	2.5	0.7	BC54	1.7	0.5
BC16	1.6	0.3	BC35	1.4	0.5	BC55	2.5	0.9
BC17	10.7	1.5	BC36	53.2	5.1	BC56	1.7	0.6
BC18	6.0	1.0	BC37	14.7	3.2	GFP	0.2	0.1

Example 25. Base editing by constructs with various linkers when expressed by plasmid in U-2OS cells

[00654] A set of sixty-eight amino acid linkers listed in Table 60 encompassing various lengths and flexibilities were encoded into the region between the N-terminal cytosine deaminase and C-terminal Cas9 nickase of expression plasmid BC27 as mentioned in Example 1A. Selected base editor constructs described in Table 58 were redesigned with the substitution of the XTEN linker between the cytosine deaminase domain and the Spy Cas9 nickase domain with a linker from **Table 60**. The BC27 linker screen was conducted using sg001373 with target sequence UCCCUGGCUGAGGAUCCCCA (SEQ ID NO: 157). The other 10 deaminase domain linker screens were conducted using sg005883 with target sequence CCCCCCGCCGUGUUUGUGGG (SEQ ID NO: 159)

Table 60. Sixty-eight amino acid linkers tested in linker activity screens.

Linker ID	Linker amino acid sequence
L001 (XTEN control)	SGSETPGTSESATPES (SEQ ID NO: 46)
L002	GGS (SEQ ID NO: 211)
L003	GGGGS (SEQ ID NO: 212)
L004	EAAAK (SEQ ID NO: 213)
L005	SEGSA (SEQ ID NO: 214)
L006	SEGSAGTST (SEQ ID NO: 215)
L007	GGGGSGGGGS (SEQ ID NO: 216)
L008	GGGGSEAAAK (SEQ ID NO: 217)
L009	EAAAKGGGGS (SEQ ID NO: 218)
L010	EAAAKEAAAK (SEQ ID NO: 219)

Linker ID	Linker amino acid sequence
L011	SEGSAGTSTESEGSA (SEQ ID NO: 220)
L012	GGGGSGGGGSGGGGS(SEQ ID NO: 221)
L013	GGGGSGGGGSEAAAK(SEQ ID NO: 222)
L014	GGGGSEAAAKGGGGS(SEQ ID NO: 223)
L015	GGGGSEAAAKEAAAK(SEQ ID NO: 49)
L016	EAAAKGGGGSGGGGS(SEQ ID NO: 50)
L017	EAAAKGGGGSEAAAK(SEQ ID NO: 224)
L018	EAAAKEAAAKGGGGS(SEQ ID NO: 225)
L019	EAAAKEAAAKEAAAK (SEQ ID NO: 51)
L020	SEGSAGTSTESEGSAGTSTE (SEQ ID NO: 226)
L021	GGGGSGGGGSGGGGSGGGGS (SEQ ID NO: 52)
L022	GGGGSGGGGSGGGGSEAAAK (SEQ ID NO: 227)
L023	GGGGSGGGGSEAAAKGGGGS (SEQ ID NO: 228)
L024	GGGGSGGGGSEAAAKEAAAK (SEQ ID NO: 53)
L025	GGGGSEAAAKGGGGSGGGGS (SEQ ID NO: 54)
L026	GGGGSEAAAKGGGGSEAAAK (SEQ ID NO: 229)
L027	GGGGSEAAAKEAAAKGGGGS (SEQ ID NO: 230)
L028	GGGGSEAAAKEAAAKEAAAK (SEQ ID NO: 231)
L029	EAAAKGGGGSGGGGSGGGGS (SEQ ID NO: 232)
L030	EAAAKGGGGSGGGGSEAAAK (SEQ ID NO: 233)
L031	EAAAKGGGGSEAAAKGGGGS (SEQ ID NO: 234)
L032	EAAAKGGGGSEAAAKEAAAK (SEQ ID NO: 235)
L033	EAAAKEAAAKGGGGSGGGGS (SEQ ID NO: 236)
L034	EAAAKEAAAKGGGGSEAAAK (SEQ ID NO: 237)
L035	EAAAKEAAAKEAAAKGGGGS (SEQ ID NO: 238)
L036	EAAAKEAAAKEAAAKEAAAK (SEQ ID NO: 56)
L037	SEGSAGTSTESEGSAGTSTESEGSA (SEQ ID NO: 239)
L038	GGGGSGGGGSGGGGSGGGGSGGGGS(SEQ ID NO: 240)
L039	GGGGSGGGGSGGGGSGGGGSEAAAK (SEQ ID NO: 241)
L040	GGGGSGGGGSGGGGSEAAAKGGGGS (SEQ ID NO: 242)
L041	GGGGSGGGGSGGGGSEAAAKEAAAK (SEQ ID NO: 243)
L042	GGGGSGGGGSEAAAKGGGGSGGGGS (SEQ ID NO: 244)
L043	GGGGSGGGGSEAAAKGGGGSEAAAK (SEQ ID NO: 245)
L044	GGGGSGGGGSEAAAKEAAAKGGGGS (SEQ ID NO: 246)
L045	GGGGSGGGGSEAAAKEAAAKEAAAK (SEQ ID NO: 247)
L046	GGGGSEAAAKGGGGSGGGGSGGGGS (SEQ ID NO: 248)
L047	GGGGSEAAAKGGGGSGGGGSEAAAK (SEQ ID NO: 249)
L048	GGGGSEAAAKGGGGSEAAAKGGGGS (SEQ ID NO: 250)

Linker ID	Linker amino acid sequence
L049	GGGGSEAAAKGGGGSEAAAKEAAAK (SEQ ID NO: 251)
L050	GGGGSEAAAKEAAAKGGGGSGGGGS (SEQ ID NO: 252)
L051	GGGGSEAAAKEAAAKGGGGSEAAAK (SEQ ID NO: 57)
L052	GGGGSEAAAKEAAAKEAAAKGGGGS (SEQ ID NO: 253)
L053	GGGGSEAAAKEAAAKEAAAKEAAAK (SEQ ID NO: 254)
L054	EAAAKGGGGSGGGGSGGGGSGGGGS (SEQ ID NO: 255)
L055	EAAAKGGGGSGGGGSGGGGSEAAAK (SEQ ID NO: 256)
L056	EAAAKGGGGSGGGGSEAAAKGGGGS (SEQ ID NO: 257)
L057	EAAAKGGGGSGGGGSEAAAKEAAAK (SEQ ID NO: 258)
L058	EAAAKGGGGSEAAAKGGGGSGGGGS (SEQ ID NO: 259)
L059	EAAAKGGGGSEAAAKGGGGSEAAAK (SEQ ID NO: 260)
L060	EAAAKGGGGSEAAAKEAAAKGGGGS (SEQ ID NO: 261)
L061	EAAAKGGGGSEAAAKEAAAKEAAAK (SEQ ID NO: 262)
L062	EAAAKEAAAKGGGGSGGGGSGGGGS (SEQ ID NO: 58)
L063	EAAAKEAAAKGGGGSGGGGSEAAAK (SEQ ID NO: 59)
L064	EAAAKEAAAKGGGGSEAAAKGGGGS (SEQ ID NO: 263)
L065	EAAAKEAAAKGGGGSEAAAKEAAAK (SEQ ID NO: 264)
L066	EAAAKEAAAKEAAAKGGGGSGGGGS (SEQ ID NO: 55)
L067	EAAAKEAAAKEAAAKGGGGSEAAAK (SEQ ID NO: 265)
L068	EAAAKEAAAKEAAAKEAAAKGGGGS (SEQ ID NO: 266)
L069	EAAAKEAAAKEAAAKEAAAKEAAAK (SEQ ID NO: 267)

[00655] Plasmid A, uses a pUC19 backbone (GenBank® Accession Number U47119), expresses an S. pyogenes single-guide RNA (sgRNA) from a U6 promoter. Plasmid B, called pCI, expresses base editor constructs from a CMV promoter which are comprised of the candidate deaminase fused to S. pyogenes-D10A-Cas9 by an experimental linker which is subsequently fused to one copy of UGI and one copy of SV40 NLS. U-2OS cells growing in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) in 96-well plates were transfected using Mirus *Trans*IT-X2® with 100 ng each of plasmid A and plasmid B. 100 uL of fresh media was added 24 hours after initial transfection. After an additional 48 hours, the media was removed and the cells were lysed with QuickExtract™ DNA Extraction Solution (Lucigen, Cat. QE09050). Table 61A and 61Blists the percent of total reads containing at least 1 cytosine to thymidine conversion across all deaminase domains tested and representative data is shown in Figs. 40A-40K.

Table 61A. Percent of total reads containing at least 1 cytosine to thymidine (%C to T) conversion

Linker	BC18			BC22			BC27			BC28			BC29			BC30		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
L001 (XTEN control)	2.07	0.90	3	50.50	2.36	3	13.47	2.42	3	43.15	8.56	2	9.70	3.97	3	12.50	2.94	3
L002	0.53	0.19	3	57.43	9.03	3	5.70	1.31	3	19.30	0.42	2	5.03	2.14	3	7.25	1.10	3
L003	1.10	0.57	3	54.37	12.74	3	6.87	2.15	3	6.80	2.12	2	4.20	1.51	3	3.30	0.87	3
L004	0.40	0.08	3	30.97	0.65	3	6.07	0.64	3	34.65	9.40	2	3.93	1.43	3	11.55	1.18	3
L005	0.93	0.54	3	29.30	2.41	3	8.07	1.10	3	17.50	3.25	2	3.60	1.47	3	10.45	2.14	3
L006	1.50	0.96	3	57.70	6.70	3	8.03	3.06	3	28.20	9.90	2	9.10	1.74	3	4.20	1.73	3
L007	1.00	0.54	3	57.10	4.87	3	5.40	1.40	3	23.50	0.42	2	7.23	2.31	3	11.05	2.36	3
L008	1.87	1.04	3	46.07	0.74	3	6.23	1.86	3	22.85	6.01	2	11.00	2.42	3	9.55	1.34	3
L009	1.50	1.20	3	58.83	5.76	3	12.73	1.93	3	22.95	7.71	2	5.87	1.27	3	2.45	0.95	3
L010	0.70	0.41	3	60.83	6.66	3	10.03	1.63	3	22.60	10.18	2	5.97	1.80	3	5.00	1.47	3
L011	1.23	0.78	3	53.87	1.56	3	10.70	2.61	3	40.70	3.82	2	12.37	4.82	3	3.15	0.85	3
L012	0.73	0.25	3	49.30	4.60	3	7.97	0.84	3	41.85	1.34	2	6.87	2.48	3	7.45	1.75	3
L013	1.53	1.11	3	53.50	4.95	3	8.40	1.85	3	32.70	9.62	2	13.63	7.80	3	2.95	0.45	3
L014	1.60	1.06	3	61.13	2.70	3	6.83	1.99	3	35.85	6.72	2	10.80	2.43	3	7.65	1.18	3
L015	0.60	0.28	3	57.80	3.47	3	6.23	1.17	3	37.30	1.27	2	14.40	1.82	3	6.70	2.08	3
L016	2.03	1.30	3	59.23	4.37	3	9.97	0.81	3	30.95	7.14	2	4.00	1.15	3	7.85	1.97	3
L017	1.23	0.49	3	52.60	3.40	3	9.83	0.59	3	37.15	1.77	2	15.23	1.79	3	8.75	2.68	3
L018	2.77	1.58	3	53.93	3.61	3	4.90	0.82	3	32.65	2.62	2	12.70	5.78	3	11.80	2.86	3
L019	0.73	0.05	3	58.03	2.42	3	9.33	1.61	3	36.35	6.72	2	13.93	5.31	3	15.65	1.59	3
L020	2.23	1.41	3	56.40	3.21	3	10.33	0.25	3	26.95	11.10	2	14.77	1.90	3	8.80	1.78	3
L021	0.83	0.09	3	54.17	10.11	3	8.50	2.01	3	25.45	8.13	2	9.10	1.15	3	7.95	2.65	3
L022	1.80	0.78	3	53.87	4.14	3	6.43	4.02	3	22.20	8.91	2	10.43	2.23	3	7.85	1.60	3

Linker	BC18			BC22			BC27			BC28			BC29			BC30		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
L023	1.50	1.07	3	42.73	5.58	3	6.27	1.36	3	33.75	1.63	2	13.00	4.80	3	3.45	1.42	3
L024	3.07	1.52	3	57.00	2.35	3	9.83	3.35	3	38.30	2.26	2	14.60	1.99	3	7.10	2.48	3
L025	2.03	1.24	3	54.70	1.92	3	4.90	0.82	3	39.45	0.21	2	2.87	1.70	3	3.75	1.70	3
L026	1.00	0.83	3	51.07	2.79	3	5.67	2.28	3	32.95	2.19	2	8.07	4.16	3	3.35	0.32	3
L027	2.00	1.44	3	58.23	5.15	3	9.07	0.31	3	37.80	4.81	2	6.53	4.43	3	6.80	0.74	3
L028	2.60	1.70	3	24.20	5.10	3	7.90	1.41	3	19.25	4.60	2	11.53	0.90	3	2.30	1.31	3
L029	1.93	0.95	3	59.35	3.82	3	8.57	0.80	3	18.90	1.41	2	12.47	2.51	3	9.15	1.84	3
L030	2.67	0.39	3	52.50	6.90	3	8.77	1.15	3	35.30	0.57	2	18.23	2.30	3	9.80	1.50	3
L031	0.67	0.29	3	56.57	2.35	3	10.47	0.59	3	23.65	0.64	2	8.03	1.91	3	3.40	1.25	3
L032	1.40	0.78	3	27.07	6.12	3	8.90	1.22	3	15.75	2.90	2	19.57	3.56	3	10.80	1.88	3
L033	0.63	0.26	3	57.75	4.76	3	8.93	2.55	3	13.15	5.87	2	12.40	5.44	3	3.55	1.23	3
L034	0.23	0.12	3	50.70	6.94	3	11.70	4.20	3	22.45	3.46	2	0.80	0.26	3	6.50	1.22	3
L035	2.00	1.13	3	55.00	4.45	3	8.80	3.99	3	44.80	7.21	2	10.77	8.43	3	10.60	-	1
L036	2.17	0.50	3	51.90	9.44	3	11.47	2.01	3	27.75	25.95	2	12.60	4.39	3	11.50	0.10	3
L037	1.17	0.45	3	53.30	2.92	3	10.23	1.65	3	34.70	8.06	2	9.20	4.12	3	7.90	0.83	3
L038	0.83	0.54	3	23.97	2.82	3	8.23	1.38	3	37.30	2.26	2	6.80	2.01	3	1.50	0.23	3
L039	1.07	0.69	3	58.70	6.50	3	10.67	2.49	3	15.40	0.99	2	11.40	3.21	3	6.35	0.71	3
L040	1.80	1.14	3	54.37	3.81	3	7.20	1.83	3	13.35	1.63	2	4.50	3.22	3	4.95	2.54	3
L041	0.33	0.21	3	47.27	9.92	3	11.57	3.50	3	33.15	2.19	2	15.67	1.39	3	10.70	0.10	3
L042	0.60	0.08	3	43.97	2.45	3	7.60	0.95	3	19.60	8.91	2	6.83	3.21	3	11.00	1.99	3
L043	1.17	0.42	3	45.07	1.69	3	2.93	0.81	3	31.55	8.56	2	11.37	8.06	3	6.95	2.52	3
L044	1.00	0.64	3	66.35	4.26	3	8.00	0.52	3	25.95	9.12	2	10.37	2.06	3	2.60	0.42	3
L045	0.73	0.45	3	64.20	3.24	3	10.27	2.05	3	34.60	6.22	2	12.30	2.90	3	5.95	1.36	3
L046	1.77	1.11	3	47.70	11.53	3	6.70	1.97	3	20.05	2.33	2	8.60	4.10	3	6.40	0.78	3

Linker	BC18			BC22			BC27			BC28			BC29			BC30		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
L047	2.00	0.57	3	45.37	12.50	3	9.43	1.02	3	40.15	4.60	2	11.33	7.96	3	1.65	0.75	3
L048	0.67	0.41	3	50.53	11.07	3	6.37	1.37	3	42.25	3.32	2	13.30	6.26	3	2.70	0.45	3
L049	1.47	0.50	3	22.23	0.48	3	6.57	0.95	3	18.05	1.91	2	14.97	4.96	3	5.35	0.98	3
L050	1.47	1.07	3	43.87	5.89	3	4.80	0.40	3	22.65	10.68	2	13.23	2.67	3	4.10	0.70	3
L051	1.20	0.59	3	57.13	3.24	3	7.53	2.51	3	30.80	8.63	2	10.97	8.58	3	8.85	2.08	3
L052	2.00	1.28	3	49.30	1.31	3	10.80	3.12	3	43.40	0.14	2	11.27	1.78	3	3.95	2.18	3
L053	2.00	1.35	3	42.30	3.48	3	9.27	1.05	3	37.05	6.86	2	10.13	6.62	3	13.80	1.23	3
L054	1.60	0.99	3	56.00	1.80	3	6.23	3.70	3	32.10	7.07	2	12.03	6.46	3	7.05	1.00	3
L055	0.50	0.22	3	55.37	1.92	3	4.43	0.78	3	14.85	5.59	2	12.77	6.56	3	6.80	2.86	3
L056	2.47	1.40	3	56.43	2.27	3	8.67	5.59	3	29.00	5.37	2	3.30	1.73	3	6.20	0.99	3
L057	3.17	0.97	3	53.63	6.42	3	9.57	2.49	3	27.65	6.86	2	15.07	8.36	3	3.00	0.50	3
L058	1.03	0.42	3	52.20	9.24	3	4.67	0.76	3	23.05	5.59	2	9.90	2.18	3	1.15	1.27	3
L059	1.47	0.52	3	41.27	17.48	3	11.10	3.01	3	44.20	4.67	2	8.80	4.41	3	5.95	1.59	3
L060	2.03	1.25	3	42.53	1.09	3	8.33	0.55	3	26.80	6.08	2	13.00	6.80	3	1.60	0.92	3
L061	1.53	0.87	3	33.53	4.62	3	11.10	1.65	3	34.35	9.83	2	15.10	11.23	3	3.20	0.12	3
L062	2.33	0.94	3	56.47	5.91	3	12.63	5.36	3	42.20	8.20	2	13.10	7.41	3	9.00	1.93	3
L063	2.17	1.25	3	60.77	3.09	3	11.03	0.91	3	32.20	8.34	2	8.77	4.32	3	9.00	1.45	3
L064	1.97	1.14	3	55.70	10.99	3	10.40	1.00	3	26.35	3.04	2	9.50	3.74	3	8.40	2.60	3
L065	2.10	1.31	3	61.57	4.12	3	7.93	2.32	3	41.55	1.77	2	9.27	4.77	3	4.95	2.04	3
L066	1.47	0.94	3	35.83	3.98	3	8.40	2.82	3	26.45	7.14	2	9.67	5.69	3	10.75	3.61	3
L067	1.70	1.15	3	52.07	3.77	3	10.13	2.87	3	30.20	1.27	2	13.90	6.69	3	12.00	3.89	3
L068	1.57	0.62	3	54.97	0.16	3	13.40	1.71	3	45.60	0.99	2	8.43	1.76	3	11.30	2.78	3
L069	1.50	0.45	3	45.43	0.45	3	12.10	5.22	3	34.25	0.49	2	8.00	3.39	3	7.55	2.70	3
No treatment	0.30	0.22	3	0.40	0.22	3	0.23	0.06	3	0.05	0.07	2	0.40	0.10	3	0.00	0.00	3

Linker	BC18			BC22			BC27			BC28			BC29			BC30		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
1XNLS Cas9	0.07	0.09	3	0.50	4.77	3	0.07	0.06	0	0.00	0.00	2	0.10	0.00	3	0.00	0.06	3
GFP	0.13	0.12	3			3	0.30	0.20	1	0.10	0.14	2	0.13	0.12	3	0.10	0.10	3

Table 61B. Mean percent of total reads containing at least 1 cytosine to thymidine (%C to T) conversion

Linker	BC36			BC38			BC39			BC41			BC42		
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
L001 (XTEN control)	33.95	15.20	2	22.90	0.42	2	29.47	3.11	3	0.33	0.09	3	55.80	4.60	3
L002	26.15	6.72	2	1.70	0.42	2	40.03	3.27	3	22.60	1.63	3	57.15	3.55	3
L003	34.35	7.28	2	2.05	0.07	2	39.30	5.48	3	27.73	2.58	3	39.90	8.40	3
L004	38.85	3.18	2	5.35	2.05	2	36.73	3.75	3	30.97	2.17	3	47.45	0.85	3
L005	28.70	6.08	2	6.45	4.60	2	42.53	3.93	3	17.90	3.16	3	41.15	2.85	3
L006	33.50	5.94	2	20.20	1.13	2	43.60	2.02	3	17.17	0.88	3	39.80	2.20	3
L007	38.35	1.63	2	5.75	2.62	2	23.37	2.80	3	20.03	2.01	3	32.55	8.15	3
L008	19.30	1.84	2	12.30	2.26	2	37.15	4.88	3	17.70	0.71	3	30.95	14.55	3
L009	26.70	5.23	2	6.70	0.85	2	25.40	0.96	3	15.27	4.74	3	40.85	3.95	3
L010	24.70	2.12	2	12.10	6.08	2	29.85	1.06	3	12.20	2.13	3	39.05	10.55	3
L011	22.40	1.13	2	8.00	0.00	2	33.67	5.40	3	13.07	4.90	3	52.95	0.95	3
L012	31.75	5.59	2	14.00	15.41	2	39.67	2.90	3	14.00	4.01	3	28.40	4.90	3
L013	53.15	4.60	2	10.20	1.41	2	38.80	5.02	3	12.40	2.57	3	43.25	3.55	3
L014	48.25	4.31	2	13.80	3.39	2	41.50	5.51	3	10.60	4.83	3	42.65	3.85	3
L015	46.45	5.87	2	15.05	1.77	2	41.23	2.85	3	11.97	3.06	3	37.25	6.45	3
L016	48.15	3.61	2	4.60	0.85	2	45.47	2.84	3	9.57	2.47	3	38.70	2.90	3
L017	49.50	10.47	2	9.00	0.42	2	40.40	2.10	3	20.07	1.80	3	47.10	3.90	3
L018	50.75	9.12	2	6.45	2.62	2	42.60	5.54	3	17.50	3.15	3	41.80	12.30	3
L019	44.05	0.07	2	26.45	4.17	2	43.20	5.78	3	8.13	1.72	3	47.35	0.75	3

	BC36			BC38			BC39			BC41			BC42		
Linker	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**
L020	39.40	5.09	2	14.40	6.08	2	41.90	3.65	3	9.63	3.84	3	39.95	3.55	3
L021	51.10	2.83	2	16.80	2.69	2	38.83	1.74	3	10.30	3.89	3	26.40	6.20	3
L022	35.20	2.12	2	3.30	1.27	2	29.05	1.63	3	16.77	2.65	3	34.10	3.10	3
L023	48.70	10.18	2	5.40	3.82	2	35.23	4.81	3	8.23	3.21	3	56.10	1.50	3
L024	45.80	0.42	2	15.45	5.59	2	33.87	2.02	3	14.87	3.20	3	22.10	7.20	3
L025	57.40	7.21	2	18.35	4.03	2	38.27	3.41	3	24.83	2.50	3	45.90	1.10	3
L026	51.50	11.74	2	10.70	0.42	2	43.47	3.01	3	13.80	5.37	3	35.30	0.60	3
L027	51.00	2.69	2	18.00	4.95	2	41.73	3.43	3	10.90	3.24	3	49.90	0.10	3
L028	50.25	2.90	2	20.40	2.69	2	41.53	3.00	3	15.67	5.03	3	40.90	6.70	3
L029	33.80	1.56	2	16.85	1.06	2	37.70	3.80	3	23.47	2.16	3	25.75	9.15	3
L030	25.80	3.68	2	19.10	3.54	2	43.77	3.59	3	15.10	7.97	3	46.80	3.60	3
L031	53.35	5.02	2	7.60	2.12	2	40.80	4.72	3	9.73	5.61	3	47.75	1.65	3
L032	61.65	8.98	2	11.35	3.04	2	23.70	4.80	3	29.97	3.97	3	30.70	4.60	3
L033	19.90	3.68	2	11.75	0.49	2	32.37	1.42	3	12.37	3.82	3	49.95	1.65	3
L034	48.00	0.85	2	7.55	3.32	2	19.60	0.99	3	14.63	2.67	3	32.05	1.25	3
L035	40.85	0.64	2	5.75	1.48	2	27.43	4.48	3	9.87	4.11	3	42.10	1.30	3
L036	30.50	5.23	2	9.80	4.10	2	27.27	3.21	3	16.03	2.52	3	41.60	1.30	3
L037	41.80	11.17	2	27.50	3.39	2	26.73	3.25	3	12.90	7.05	3	39.75	2.65	3
L038	41.20	1.98	2	4.55	0.78	2	42.27	0.76	3	10.07	2.83	3	41.85	0.95	3
L039	39.20	0.99	2	12.85	7.85	2	32.83	3.56	3	16.00	7.47	3	29.30	3.70	3
L040	58.90	3.54	2	7.20	2.55	2	41.53	5.04	3	15.37	4.58	3	30.25	3.85	3
L041	25.00	1.84	2	10.30	4.10	2	31.27	4.11	3	15.80	6.83	3	46.85	1.65	3
L042	59.35	9.26	2	14.80	0.28	2	27.70	2.63	3	17.97	4.63	3	37.35	8.25	3
L043	34.45	6.58	2	0.20	0.28	2	25.53	5.01	3	8.53	2.82	3	39.95	6.45	3
L044	50.70	4.10	2	13.20	1.27	2	34.77	3.35	3	16.63	4.64	3	24.85	4.15	3
L045	36.40	6.36	2	7.45	3.75	2	36.67	2.66	3	18.67	2.53	3	43.45	1.65	3
L046	49.05	5.16	2	9.35	0.64	2	26.90	0.28	3	14.33	1.87	3	40.00	0.10	3

	BC36			BC38			BC39			BC41			BC42		
Linker	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**
L047	49.05	8.84	2	10.50	1.13	2	33.10	5.60	3	15.03	2.81	3	48.10	8.50	3
L048	56.50	0.14	2	7.45	3.46	2	40.77	2.81	3	17.60	2.84	3	46.95	5.45	3
L049	29.50	3.82	2	13.35	0.78	2	38.67	2.25	3	11.30	4.30	3	19.50	5.00	3
L050	36.70	0.28	2	13.30	2.26	2	29.10	2.08	3	9.80	6.38	3	39.85	5.15	3
L051	31.30	0.99	2	24.45	7.14	2	34.07	3.72	3	11.33	3.11	3	30.50	3.20	3
L052	50.55	5.30	2	19.45	0.21	2	33.63	3.65	3	13.07	5.70	3	48.20	2.60	3
L053	44.35	4.60	2	6.55	3.75	2	35.83	2.54	3	12.83	4.09	3	49.65	2.25	3
L054	33.65	4.31	2	25.50	2.69	2	39.97	7.47	3	21.33	8.77	3	53.95	2.15	3
L055	47.80	2.40	2	7.95	0.78	2	28.17	2.50	3	12.17	7.25	3	49.60	2.50	3
L056	31.10	3.39	2	9.55	0.92	2	45.03	6.52	3	12.37	3.70	3	51.35	2.25	3
L057	52.45	7.85	2	12.90	9.90	2	26.93	1.20	3	14.63	5.10	3	47.65	4.55	3
L058	44.95	3.61	2	7.90	0.42	2	26.85	2.05	3	11.17	0.57	3	38.80	3.50	3
L059	54.65	7.57	2	8.45	4.31	2	33.60	4.75	3	11.03	4.98	3	44.90	7.60	3
L060	60.95	1.34	2	16.25	2.62	2	31.47	1.54	3	11.63	5.54	3	1.35	0.75	3
L061	38.25	3.75	2	8.25	2.33	2	42.17	1.14	3	18.40	5.80	3	30.95	5.15	3
L062	30.95	3.18	2	10.40	5.52	2	33.70	6.12	3	22.27	7.07	3	31.60	10.50	3
L063	41.15	1.20	2	24.15	4.45	2	38.67	4.02	3	14.50	3.11	3	32.35	3.05	3
L064	60.90	3.11	2	22.20	2.83	2	40.80	2.30	3	17.30	3.49	3	43.90	0.90	3
L065	55.45	6.01	2	18.75	2.19	2	34.97	4.17	3	21.40	1.28	3	36.45	11.15	3
L066	34.10	3.82	2	25.05	3.32	2	24.77	1.02	3	18.30	2.55	3	40.95	5.95	3
L067	28.15	0.78	2	8.60	6.65	2	38.77	0.93	3	22.00	9.53	3	37.25	5.45	3
L068	16.55	1.77	2	26.75	2.62	2	35.30	1.87	3	5.00	1.80	3	42.00	8.00	3
L069	51.70	3.25	2	20.40	2.40	2	40.30	2.82	3	9.93	3.16	3	36.30	10.40	3
No treatment	0.15	0.07	2	0.15	0.07	2	1.27	1.36	3	0.17	0.09	3	0.20	0.00	3
1XNLS Cas9	0.15	0.21	2	0.00	0.00	2	0.00	0.00	3	0.03	0.05	3	0.10	0.10	3
GFP	0.40	0.28	2	0.00	0.00	2	0.33	0.12	3	0.23	0.05	3	0.25	0.05	3

Example 26. Base editing by constructs with various linkers in Huh-7 cells

[00656] Select base editor constructs derived from BC22 but with linkers from Table 60 substituted between the cytosine deaminase and Cas9 nickase were assayed for C to T base editing activity using Lipofectamine transfection of mRNA and gRNA. Tested base editor mRNAs include SEQ ID NOs: 19 and 347-357. Constructs were screened in a dilution series ranging from 150 ng to 1.17 ng of base editor mRNA in Huh-7 cells, co-delivered with a dilution series of SERPINA1 sgRNA (G000255) ranging from 20 to 0.15 nM and UGI mRNA (SEQ ID No: 25) at a dose of 25 to 0.20 ng in 100 uL of media. C to T base editing of each construct was compared to BC22 (SEQ ID No: 19).

Example 26.1. Cell preparation and transfection

[00657] Huh-7 cells growing in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) in 96-well plates were transfected using Lipofectamine® RNAiMAX with an 8-point, two-fold dilution series starting with a maximum dose of 150 ng base editor mRNA, 25 ng UGI mRNA and 20 nM sgRNA in 100 uL media.

Example 26.2. Evaluation of C to T editing purity by NGS

[00658] Seventy-two hours post transfection, Huh-7 cells were subjected to lysis, PCR amplification of the SERPINA1 locus and subsequent NGS analysis, as described in Example 1. Table 62 and Fig. 41 show SERPINA1 editing levels and the C to T editing purity in samples treated with either 2.3 or 75 ng of different base editor mRNAs and corresponding levels of diluted UGI mRNA and G000255 (SERPINA1). Mean percent C to T conversion for base editor mRNAs with different linkers are shown in Tables 63A-63B. Table 64 and Fig. 42 show EC90 (mass of base editor mRNA required to edit 90% of maximum C to T edits) ranges from 9.7 to 23.1 ng. All the tested base editor mRNAs similar levels of maximal editing at high doses

Table 62 – Mean percent editing at SERPINA1 in Huh-7 Indels = insertions or deletions.

Base editor mRNA (ng)	Linker	C to T (%)		C to A/G (%)		Indels (%)	
		Mean	SD	Mean	SD	Mean	SD
2.3	L001 (XTEN control)	49.6	3.0	3.0	0.4	10.7	0.6
	L015	43.7	0.8	2.7	0.1	9.0	1.6
	L016	43.4	2.8	3.0	0.4	10.7	0.5

Base editor mRNA (ng)	Linker	C to T (%)		C to A/G (%)		Indels (%)	
		Mean	SD	Mean	SD	Mean	SD
	L019	50.2	5.7	3.3	0.4	11.3	0.6
	L021	45.3	7.5	3.1	0.6	10.5	0.1
	L024	46.3	2.2	3.3	0.0	10.7	0.1
	L025	33.2	14.3	2.4	0.7	9.9	0.9
	L036	46.5	1.4	3.0	0.6	8.0	1.2
	L051	38.4	6.2	2.5	0.1	7.2	0.4
	L062	40.8	7.7	1.2	1.7	10.0	2.6
	L063	37.8	1.8	2.0	0.5	6.9	0.3
	L066	33.4	9.2	2.7	0.4	7.2	1.6
75	L001 (XTEN control)	89.8	0.8	2.3	0.4	4.4	0.2
	L015	83.6	0.8	2.9	0.8	5.3	0.1
	L016	87.9	1.1	2.9	0.0	5.2	0.3
	L019	89.6	0.8	2.7	0.2	4.3	0.5
	L021	91.1	0.3	2.0	0.3	3.3	0.3
	L024	89.9	2.1	3.0	0.5	3.8	0.9
	L025	89.8	0.8	3.4	0.1	3.5	0.3
	L036	90.4	1.0	2.8	0.0	3.4	0.1
	L051	91.3	0.9	2.0	0.1	3.3	0.4
	L062	91.1	0.4	2.6	0.2	2.7	0.2
	L063	87.2	3.5	2.6	0.2	3.6	0.4
	L066	91.9	0.4	2.4	0.1	3.1	0.1

Table 63A– Mean percent C to T editing at SERPINA1 in Huh-7 cells.

	L001 (XTEN control)		L015		L016		L019		L021		L024	
Base editor mRNA (ng)	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
150.00	83.4	0.9	81.3	0.4	79.9	0.3	80.4	1.0	81.6	3.4	84.4	-
75.00	89.8	0.8	83.6	0.8	87.9	1.1	89.6	0.8	91.1	0.3	89.9	2.1
37.50	87.9	1.0	84.2	6.2	88.6	0.7	88.1	0.6	87.9	0.5	88.7	1.3
18.75	85.6	1.5	74.7	4.7	83.8	1.0	85.4	2.3	83.6	1.9	84.9	1.6
9.38	73.2	0.1	68.7	0.1	74.5	3.6	72.2	4.5	71.6	5.3	76.5	1.8
4.69	66.1	0.3	54.7	4.0	61.0	3.7	66.9	1.9	62.3	1.2	58.2	2.0
2.34	49.6	3.0	43.7	0.8	43.4	2.8	50.2	5.7	45.3	7.5	46.3	2.2
1.17	29.0	0.2	25.8	0.8	25.0	0.9	28.3	3.7	24.9	3.0	25.1	2.6

Table 63B - Mean percent C to T editing at SERPINA1 in Huh-7 cells.

	L025		L036		L051		L062		L063		L066	
Base editor mRNA (ng)	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
150.00	81.5	3.4	82.2	3.7	84.0	0.5	84.3	1.6	78.7	2.4	78.8	3.5
75.00	89.8	0.8	90.4	1.0	91.3	0.9	91.1	0.4	87.2	3.5	91.9	0.4
37.50	85.4	4.0	90.0	2.2	89.6	2.6	89.1	1.8	81.7	1.6	89.5	0.7
18.75	83.1	1.0	87.0	1.3	84.1	2.9	86.2	1.0	74.9	1.8	84.8	0.6
9.38	71.7	0.3	78.9	0.5	76.8	0.7	73.9	7.4	67.2	2.5	67.9	1.0
4.69	59.3	2.1	68.3	2.2	58.7	5.1	62.7	4.1	48.6	2.0	55.7	8.5
2.34	33.2	14.3	46.5	1.4	38.4	6.2	40.8	7.7	37.8	1.8	33.4	9.2
1.17	23.1	0.7	30.8	1.4	25.1	1.1	26.8	1.2	23.5	3.6	28.1	2.8

Table 64 - Mass of BC22 mRNAs that leads to 90% of the maximum C to T editing at SERPINA1. The 95% confidence interval for each EC90 value is also shown.

Linker	EC90 (ng)	95% confidence interval	
		Max (ng)	Min (ng)
L001 (XTEN control)	11.23	16.33	8.199
L015	18.84	32.71	12.28
L016	11.57	16.91	8.421
L019	9.974	17.4	6.444
L021	13.07	22.77	8.449
L024	14.67	21.5	10.700
L025	13.99	25.53	8.893
L036	9.744	13.98	7.213
L051	14.25	21.44	10.21
L062	13.84	22.82	9.324
L063	23.14	41.8	14.77
L066	19.38	48.34	10.54

Example 27. Base editing by constructs with various linkers in primary human hepatocytes (PHH)

[00659] Select base editor constructs derived from BC22n but with linkers from **Table 65** substituted between the cytosine deaminase and Cas9 nickase were assayed for C to T base editing activity in PHH. Constructs (SEQ ID NOs: 341-346) were screened in a 12-point dilution series of base editor mRNA in PHH cells, co-delivered with a fixed mass of ANAPC5 sgRNA (G019427) and UGI mRNA (SEQ ID NO: 34). The C to T editing efficiency of each tested based editor construct was compared to BC22n (SEQ ID NO: 1).

Table 65 – Additional linker peptides.

Linker ID	Linker amino acid sequence
L001 (XTEN control)	SGSETPGTSESATPES (SEQ ID NO: 46)
L070	GTKDSTKDIPETPSKD (SEQ ID NO: 268)
L071	GRDVRQPEVKEEKPES (SEQ ID NO: 269)
L072	EGKSSGSGSESKSTAG (SEQ ID NO: 270)
L073	TPGSPAGSPTSTEEGT (SEQ ID NO: 271)
L074	GSEPATSGSETPGTST (SEQ ID NO: 272)
L019	EAAAKEAAAKEAAAK (SEQ ID NO: 51)

Example 27.1. Cell preparation and transfection

[00660] PHH cells are thawed and recovered in CHRMs media (Gibco, cat#CM7000). They are then resuspended in Primary Hepatocyte Plating Media (Consisting of William's E media (Gibco, Cat#A1217601) and primary hepatocyte plating supplements (Gibco, Cat#CM3000))) to be plated in collagen-coated 96-well plates at a density of 33,000 cells /well for twenty-four hours. Post twenty-four hours, cells are washed, and fresh primary hepatocyte maintenance media is added on cells. Cells are then transfected simultaneously with separate lipoplexes formed individually with either UGI mRNA (SEQ ID No: 34), ANAPC5 sgRNA G019427 or base editor mRNA.

[00661] Lipofection reagent was prepared as mixture of lipids at a ratio of 50/9/38/3 Lipid A, , DSPC, cholesterol and PEG2k-DMG as described in Example 1. Lipofection reagent was combined by bulk mixing with separately with each RNA species: base editor mRNA, UGI mRNA or gRNA G019427. The materials were combined at a lipid amine to RNA phosphate (N:P) molar ratio of about 6. The resulting bulk-mixed lipoplex material (lipid kits) was pre-incubated with 10% FBS (Gibco; A3160501) in Primary Hepatocyte Maintenance Media (Consisting of William's E media (Gibco, Cat#A1217601) and primary hepatocyte maintenance supplements (Gibco, Cat#CM4000)), for 15 min before addition to hepatocytes.

[00662] Each well received three components in a final volume of 100 uL: base editor mRNAs ranging from 400 ng to 0 ng of mRNAs, 30 ng of UGI mRNA and 5 pmols of G019427 (ANAPC5).

Example 27.2 Evaluation of C to T editing purity by next generation sequencing (NGS)

[00663] Seventy-two hours post transfection, PHH cells were subjected to lysis, PCR amplification of the ANAPC5 locus and subsequent NGS analysis, as described in Example 1. **Table 66** and **Fig. 43** show ANAPC5 total editing levels and the C to T editing purity in samples treated with either 400 or 6.25 ng of different base editor mRNAs in addition to 30 ng of UGI mRNA (SEQ ID No: 34) and 5 pmols of G019427 (ANAPC5).

[00664] C to T base editing for a range of base editor mRNAs doses is shown in **Table 67**. The EC95 (mass of BC22n mRNA required to edit 95% of maximum C to T edits) was calculated as shown in **Table 68** and **Fig. 44**. All of the base editor mRNAs were able to achieve the similar levels of maximal editing at high doses.

Table 66– Mean percent editing at ANAPC5 in PHH cells using base editing constructs designed with various peptide linkers. Indels = insertions or deletions.

Base editor mRNA (ng)	Linker	C TO T (%)		C to A/G (%)		Indels (%)	
		Mean	SD	Mean	SD	Mean	SD
6.25	L001 (XTEN control)	74.9	1.6	0.6	0.1	0.6	0.2
	L070	58.0	3.4	0.5	0.0	0.5	0.1
	L071	57.8	2.0	0.6	0.1	0.9	0.2
	L072	63.7	2.8	0.4	0.1	0.7	0.6
	L073	59.7	1.3	0.5	0.1	0.6	0.1
	L074	53.1	2.5	0.5	0.3	0.7	0.6
	L019	59.9	3.5	0.4	0.1	0.8	0.2
400	L001 (XTEN control)	86.3	0.1	0.6	0.4	1.2	0.4
	L070	84.3	3.3	0.6	0.2	0.7	0.1
	L071	87.7	2.4	0.7	0.2	0.5	0.4
	L072	86.0	0.1	0.9	0.4	0.7	0.4
	L073	87.9	1.8	0.9	0.0	0.8	0.3
	L074	86.6	1.8	0.6	0.1	0.7	0.0
	L019	86.7	4.9	0.9	0.0	1.2	0.1

Table 67 – Mean percent C to T editing at the ANAPC5 locus in PHH for base editing constructs designed with various peptide linkers.

mRNA (ng)	L001		L070		L071		L072		L073		L074		L019	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
400.00	86.3	0.1	84.3	3.3	87.7	2.4	86.0	0.1	87.9	1.8	86.6	1.8	86.7	4.9
200.00	87.8	2.4	83.9	1.1	88.1	0.5	86.9	5.4	87.6	1.8	86.6	2.5	87.4	3.2
100.00	91.1	0.2	84.7	2.3	84.6	4.5	87.0	2.5	83.6	0.1	85.1	4.7	85.9	2.5
50.00	89.0	1.8	78.0	5.4	82.4	5.0	85.8	0.6	83.9	1.8	84.3	0.8	85.0	3.7
25.00	86.9	1.8	79.1	0.2	78.6	4.6	81.3	3.3	79.4	2.3	79.4	0.5	82.7	1.1
12.50	83.2	0.8	71.3	2.5	69.3	6.0	71.8	1.6	72.6	4.4	69.2	0.9	72.8	2.5
6.25	74.9	1.6	58.0	3.4	57.8	2.0	63.7	2.8	59.7	1.3	53.1	2.5	59.9	3.5
3.13	64.3	2.9	40.5	3.0	42.8	2.9	47.8	1.4	42.4	0.6	37.7	2.1	45.6	1.6
1.56	50.8	2.7	24.6	6.6	30.8	0.2	28.3	0.7	27.8	0.7	27.1	0.7	31.6	1.1
0.78	35.7	2.3	18.8	0.0	20.1	0.0	20.8	0.7	17.4	0.8	16.0	0.8	22.8	3.7
0.39	23.6	3.9	12.1	1.7	12.1	0.7	13.2	0.6	12.5	1.1	10.0	0.6	14.6	0.8
0.00	0.4	0.2	0.6	0.1	0.5	0.0	0.6	0.2	0.4	0.1	0.4	0.3	0.6	0.1

Table 68 – Mass of base editor mRNAs that leads to 95% of the maximum C to T editing. The 95% confidence interval for each EC95 value is also shown.

Linker	EC95 (ng)	95% confidence interval	
		Max (ng)	Min (ng)
L001 (XTEN control)	22.73	31.91	16.75
L070	48.99	84.83	30.52
L071	83.46	134.60	54.97
L072	45.72	67.33	32.28
L073	53.62	73.54	40.17
L074	69.89	98.83	51.13
L019	62.06	98.67	41.39

Example 28. Base editing by constructs with various linkers in T cells

[00665] Select base editor constructs derived from BC22n but with linkers from Table 65 substituted between the cytosine deaminase and Cas9 nickase were assayed for C to T base editing activity. Constructs (SEQ ID Nos: 341-346) were screened in a 12-point dilution series of base editor mRNA in healthy human T cells, co-delivered with a fixed amount of TRAC sgRNA (G016017) and of UGI mRNA (SEQ ID NO: 34). The C to T editing efficiency of each base editor mRNA construct was compared to BC22n (SEQ ID NO: 1).

Example 28.1. T cell preparation

[00666] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed, re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. 130-070-525) and processed in a MultiMACS™ Cell 24 Separator Plus device (Miltenyi Biotec). T cells were isolated via positive selection using a Straight from Leukopak® CD4/CD8 MicroBead kit, human (Miltenyi Biotec Cat. 130-122-352). T cells were aliquoted and cryopreserved for future use in Cryostor® CS10 (StemCell Technologies Cat. 07930).

[00667] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth media (TCGM) composed of CTS OpTmizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512) 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/ml recombinant human interleukin 15

(Peprotech, Cat. 200-15). T cells were rested in this media for 24 hours, at which time they were activated with T Cell TransAct™, human reagent (Miltenyi, Cat. 130-111-160) added at a 1:100 ratio by volume. T cells were activated for 48 hours prior to use in screening experiments.

Example 28.2. T cell electroporation and expansion

[00668] A sgRNA targeting TRAC (G016017) was denatured for 2 minutes at 95°C, incubated at room temperature for 5 minutes and stored on ice. 48 hours post-activation, T cells were harvested, centrifuged at 500 g for 5 min, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with a decreasing amount of base editor mRNAs ranging from 400 ng to 0 ng, 200 ng of UGI mRNA and 40 pmols of G016017 in a final volume of 20 uL of P3 electroporation buffer.

The resulting mix was transferred in duplicate to 96-well Nucleofector™ plates (Lonza) and electroporated using the manufacturer's pulse code. Immediately after electroporation, T cells received 80 µL of TCGM and plates were incubated at 37°C for 15 minutes. After incubation, 80 µL were transferred to new flat-bottom 96-well plates containing 80 µL of TCGM. T cells were incubated at 37°C for 4 days, at which time they were mixed, split 1:4 with fresh TCGM, and incubated at for 3 additional days prior to phenotypic assessments by flow cytometry.

Example 28.3. Evaluation of C to T editing purity by next generation sequencing (NGS)

[00669] Four days post electroporation, T cells were subjected to lysis, PCR amplification of the TRAC locus and subsequent NGS analysis, as described in Example 1. Table 69 and Fig. 45 show TRAC editing levels and the C to T editing purity in samples treated with either 400 or 6.25 ng of different base editor mRNAs in addition to 200 ng of UGI mRNA (SEQ ID No: 34) and 40 pmols of G016017 (TRAC).

[00670] None of the linkers tested between the N-terminal cytosine deaminase and C-terminal Cas9 nickase limit the efficiency of cytosine base editing or impacts the levels of C to T editing purity.

Table 69 – Mean percent editing at the TRAC locus in T cells treated with base editor constructs designed with various linker peptides. Indels = insertions or deletions.

base editor	Linker	C to T (%)	C to A/G (%)	Indels (%)

mRNA (ng)		Mean	SD	Mean	SD	Mean	SD
6.25	L001 (XTEN control)	78.9	4.9	0.3	0.1	0.3	0.3
	L070	68.2	1.8	0.3	0.2	0.2	0.1
	L071	62.3	8.5	0.2	0.1	0.4	0.0
	L072	62.0	4.7	0.1	0.2	0.2	0.1
	L073	57.4	3.4	0.4	0.4	0.3	0.1
	L074	64.8	2.1	0.1	0.0	0.3	0.0
	L019	67.3	7.0	0.3	0.0	0.2	0.0
400	L001 (XTEN control)	97.0	0.9	0.3	0.1	0.7	0.3
	L070	98.1	0.1	0.2	0.1	0.4	0.2
	L071	97.7	0.5	0.3	0.1	0.5	0.1
	L072	97.6	0.6	0.2	0.0	0.3	0.0
	L073	97.5	0.7	0.4	0.0	0.6	0.0
	L074	96.9	0.5	0.4	0.3	0.6	0.1
	L019	97.4	0.3	0.3	0.2	0.2	0.2

Example 28.4. Evaluation of receptor knockout by flow cytometry

[00671] Seven days post electroporation, T cells were assayed by flow cytometry to evaluate loss of CD3 expression. T cells were incubated with a fixable viability dye (Beckman Coulter, Cat. C36628) and an antibody cocktail targeting the following molecules: CD3 (Biolegend, Cat. 317336), CD4 (Biolegend, Cat. 317434) and CD8 (Biolegend, Cat. 301046). Cells were subsequently washed, analyzed on a Cytoflex LX instrument (Beckman Coulter) using the FlowJo software package. T cells were gated on size, viability and CD8 positivity before expression of any markers was determined. The resulting data was plotted on GraphPad Prism v. 9.0.2 and analyzed using a variable slope (four parameter) non-linear regression.

[00672] As shown in Table 70, treatment with >100 ng of any of base editor mRNA tested resulted in >95% of CD8+ T cells lacking expression of CD3. The EC90s of the base editor mRNAs tested (i.e., mass of mRNA that leads to 90% of CD8+ T cells lacking CD3) with 95% confidence interval of each non-linear regression are shown in Table 71 and Fig. 46.

Table 70 – Mean percentages of CD8+ T cells that are negative for CD3 surface expression following treatment with base editor constructs designed with various linker peptides.

	L001 (XTEN control)		L070		L071		L072		L073		L074		L019	
Base editor mRNA (ng)	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
400.00	96.5	0.3	97.2	0.5	96.1	0.3	96.1	0.4	96.2	0.2	96.2	0.4	95.6	0.4
200.00	96.8	0.8	96.4	1.8	96.7	0.8	96.8	0.9	96.1	1.0	95.7	0.9	96.2	1.0
100.00	97.1	0.1	96.2	1.6	97.0	0.2	95.0	1.4	96.0	0.7	96.4	0.7	95.3	0.3
50.00	96.3	0.7	96.0	0.6	95.2	1.2	95.2	1.6	94.6	2.3	94.5	1.2	95.8	1.3
25.00	94.9	1.5	92.8	3.4	92.1	3.7	90.7	2.4	90.7	0.3	91.5	1.2	91.7	1.1
12.50	91.5	1.6	85.7	5.1	82.8	5.9	83.6	4.2	82.5	2.3	79.9	0.6	85.0	1.8
6.25	76.1	7.9	69.0	5.4	61.8	8.6	63.8	3.6	61.8	3.1	65.7	4.1	67.1	5.6
3.13	56.5	6.2	49.0	5.5	42.6	6.3	42.2	5.7	43.0	6.7	43.7	3.2	47.7	5.5
1.56	37.8	2.2	29.0	4.9	26.9	2.8	24.1	5.4	24.4	3.2	25.1	2.7	29.0	3.5
0.78	23.5	3.7	17.6	3.7	14.0	0.1	13.8	2.1	14.0	3.1	13.5	1.8	16.3	2.1
0.39	13.2	2.9	9.0	2.3	7.6	0.8	7.7	0.5	7.0	0.3	6.5	0.7	9.0	0.9
0.00	1.9	0.5	1.5	0.1	1.4	0.7	2.1	1.3	1.3	0.4	1.6	0.7	2.1	0.5

Table 71 - Mass of base editor mRNAs that leads to 90% of the maximum knockdown of CD3. The 95% confidence interval for each EC90 value is also shown.

Linker	**EC90 (ng)**	**95% confidence interval**	
		Min (ng)	**Max (ng)**
L001 (XTEN control)	13.4	10.6	17.5
L070	17.1	13.6	21.9
L071	21.5	16.6	28.7
L072	18.9	15.5	23.5
L073	20.7	17.1	25.5
L074	19.7	17.2	22.6
L019	17.6	14.6	21.5

Example 29. Multiplex base editing by constructs with various linker in T cells

[00673] Select base editor constructs derived from BC22n but with linkers from Table 65 substituted between the cytosine deaminase and Cas9 nickase were assayed for multiple simultaneous base editing efficacy via protein knockdown. Constructs were screened in a 12-point dilution series of base editor mRNA in healthy human T cells, co-delivered with a fixed mass of UGI mRNA (SEQ ID NO: 34) and 4 distinct 91-mer sgRNAs. The potency of each mRNA construct was compared to BC22n mRNA (SEQ ID NO: 1) to evaluate the effects of each linker in terms of phenotypic receptor knockout.

Example 29.1. T cell preparation

[00674] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed, re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. 130-070-525) and processed in a MultiMACS™ Cell 24 Separator Plus device (Miltenyi Biotec). T cells were isolated via positive selection using a Straight from Leukopak® CD4/CD8 MicroBead kit, human (Miltenyi Biotec Cat. 130-122-352). T cells were aliquoted and cryopreserved for future use in Cryostor® CS10 (StemCell Technologies Cat. 07930).

[00675] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth media (TCGM) composed of CTS OpTmizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512) 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/ml recombinant human interleukin 15 (Peprotech, Cat. 200-15). T cells were rested in this media for 24 hours, at which time they

were activated with T Cell TransAct™, human reagent (Miltenyi, Cat. 130-111-160) added at a 1:100 ratio by volume. T cells were activated for 48 hours prior to use in screening experiments.

Example 29.2. T cell electroporation and expansion

[00676] Four 91-mer sgRNAs targeting TRAC (G023520), TRBC1/2 (G023524), CIITA (G023521) and HLA-A (G023523) were denatured for 2 minutes at 95°C, incubated at room temperature for 5 minutes and stored on ice. 48 hours post-activation, T cells were harvested, centrifuged at 500 g for 5 min, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with a dilution series of base editor mRNAs from 400 ng to 0 ng (SEQ ID NOs: 1 and 321-326), 200 ng of UGI mRNA (SEQ ID NO: 34), 3.6 pmols of G023520 (TRAC), 6.96 pmols of G023524 (TRBC1/2), 17.12 pmols of G023521 (CIITA) and 52.24 pmols of G023523 (HLA-A) in a final volume of 20 uL of P3 electroporation buffer.

[00677] The resulting mix was transferred in duplicate to 96-well Nucleofector™ plates (Lonza) and electroporated using the manufacturer's pulse code. Immediately after electroporation, T cells received 80 µL of TCGM and plates were incubated at 37°C for 15 minutes. After incubation, 80 µL were transferred to new flat-bottom 96-well plates containing 80 µL of TCGM. T cells were incubated at 37°C for 4 days, at which time they split 1:4 with fresh TCGM, and incubated at for 3 additional days prior to phenotypic assessments by flow cytometry.

Example 29.3. Evaluation of receptor knockout by flow cytometry

[00678] Seven days post electroporation, T cells were assayed by flow cytometry to evaluate loss of CD3, HLA-A3 and/or HLA-DR, DP, DQ. T cells were incubated with a fixable viability dye (Beckman Coulter, Cat. C36628) and an antibody cocktail targeting the following molecules: CD3 (Biolegend, Cat. 317336), CD4 (Biolegend, Cat. 317434), CD8 (Biolegend, Cat. 301046), HLA-A3 (ThermoFisher, Cat. 12-5754-42) and HLA-DP, DQ, DR (Biolegend, Cat. 361714). Cells were subsequently washed, analyzed on a Cytoflex LX instrument (Beckman Coulter) using the FlowJo software package. T cells were gated on size, viability and CD8 positivity before expression of any markers was determined. The resulting data was plotted on GraphPad Prism v. 9.0.2 and analyzed using a variable slope (four parameter) non-linear regression.

[00679] As shown in Table 72 treatment with >100 ng of any of BC22n mRNA tested resulted in >97% of CD8+ T cells lacking expression of CD3, HLA-A3 and HLA-DR, DP, DQ. The EC90s of the base editor mRNAs tested (i.e., mass of mRNA that leads to 90% of CD8+ T cells lacking CD3) with 95% confidence interval of each non-linear regression are shown in Table 73 and Fig. 47.

Table 72 – Mean percentages of CD8+ T cells that are negative for CD3, HLA-A3 and HLA-DR, DP, DQ surface expression following treatment with base editor mRNAs.

	L001 (XTEN control)		L070		L071		L072		L073		L074		L019	
mRNA (ng)	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
400.00	97.4	0.5	98.3	ND	98.4	0.3	98.7	0.1	97.8	0.7	98.1	0.5	98.6	0.4
200.00	98.4	0.1	98.6	0.2	98.2	0.2	98.3	0.4	98.1	0.7	98.7	0.1	98.9	0.1
100.00	97.1	0.8	97.5	1.1	97.0	2.1	97.7	0.5	97.4	0.5	98.5	0.5	98.7	0.1
50.00	95.7	0.1	96.6	0.2	95.4	0.9	95.6	0.8	95.1	2.8	97.4	0.7	97.1	0.3
25.00	90.8	0.1	88.5	4.0	89.7	3.3	89.1	2.8	87.6	3.2	91.5	2.6	93.6	1.9
12.50	75.5	3.2	77.8	3.7	75.6	4.8	71.6	6.6	69.1	3.2	77.9	4.7	82.6	3.3
6.25	54.9	4.2	54.0	1.1	54.8	6.2	53.3	5.5	46.2	5.9	55.6	7.8	60.0	5.1
3.13	32.9	2.1	32.7	1.9	31.5	3.2	30.5	2.6	27.6	2.5	32.9	3.6	35.7	2.5
1.56	17.1	0.5	16.6	1.6	16.3	2.4	16.1	1.6	15.5	3.1	18.4	2.5	21.0	2.1
0.78	10.0	1.3	9.4	1.2	9.5	0.3	7.5	0.4	7.3	0.5	9.6	2.9	11.7	0.3
0.39	5.1	0.4	4.5	0.6	4.7	0.4	4.5	0.8	4.4	1.2	5.2	0.0	6.8	1.4
0.00	0.2	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.3	0.2	0.2	0.2	0.2	0.1

Table 73 - Mass of base editor mRNAs that leads to 90% of the maximum knockdown of CD3, HLA-A3 and HLA-DR, DP, DQ (EC90). The 95% confidence interval for each EC90 value is also shown.

Linker	EC90 (ng)	95% confidence interval	
		Min (ng)	Max (ng)
L001 (XTEN control)	26.3	22.9	30.4
L070	26.5	22.8	31.1
L071	26.4	22.0	32.1
L072	30.3	25.1	37.3
L073	34.3	28.3	42.4
L074	25.0	20.4	31.2
L019	4.709	4.314	5.130

Example 30 Editing human T cells with BC22n, UGI and 91-mer sgRNAs

[00680] The base editing efficacy of 91-mer sgRNAs as assessed by NGS and/or receptor knockout was compared to that of 100-mer sgRNA controls with the same guides sequences.

[00681] The tested 91-mer sgRNAs include a 20 nucleotide guide sequence (as represented by N) and a guide scaffold as follows:
mN*mN*mN*NNNNNNNNNNNNNNNNNNGUUUUAGAmGmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUAAGGCUAGUCCGUUAUCACGAAAGGGCACCGAGUCG GmUmGmC*mU (SEQ ID NO: 520), where A, C, G, U, and N are adenine, cytosine, guanine, uracil, and any ribonucleotide, respectively, unless otherwise indicated. An m is indicative of a 2'O-methyl modification, and an * is indicative of a phosphorothioate linkage between the nucleotides. Unmodified and modified versions of the guides are provided in Table 5C (Sequence Table).

Example 30.1. T cell preparation

[00682] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed, re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. 130-070-525) and processed in a MultiMACS™ Cell 24 Separator Plus device (Miltenyi Biotec). T cells were isolated via positive selection using a Straight from Leukopak® CD4/CD8 MicroBead kit, human (Miltenyi Biotec Cat. 130-122-352). T cells were aliquoted and cryopreserved for future use in Cryostor® CS10 (StemCell Technologies Cat. 07930).

[00683] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth media (TCGM) composed of CTS OpTmizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512) 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/ml recombinant human interleukin 15 (Peprotech, Cat. 200-15). T cells were rested in this media for 24 hours, at which time they were activated with T Cell TransAct™, human reagent (Miltenyi, Cat. 130-111-160) added at a 1:100 ratio by volume. T cells were activated for 48 hours prior to LNP treatments.

Example 30.2. T cell LNP treatment and expansion

[00684] Forty-eight hours post-activation, T cells were harvested, centrifuged at 500 g for 5 min, and resuspended at a concentration of 1 x 10^6 T cells/mL in T cell plating media (TCPM): a serum-free version of TCGM containing 400 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 10 ng/ml recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 10 ng/ml recombinant human interleukin 15 (Peprotech, Cat. 200-15). 50 µL of T cells in TCPM (5 x 10^4 T cells) were added per well to be treated in flat-bottom 96-well plates.

[00685] LNPs were prepared as described in Example 1 at a ratio of 35/47.5/15/2.5 (Lipid A/ cholesterol/DSPC/PEG2k-DMG). The LNPs were formulated with a lipid amine to RNA phosphate (N:P) molar ratio of about 6. LNPs encapsulated a single RNA species, either a sgRNA as described in Table 74, BC22n mRNA (SEQ ID No: 1), or UGI mRNA (SEQ ID No. 34).

Table 74 - 100-mer and 91-mer sgRNAs.

Gene target	100-mer	91-mer
TRAC	G016017	G023520
TRBC1/2	G016200	G023524
CIITA	G016086	G023521
B2M	G015991	G023519
CD38	G019771	G023522
HLA-A	G021209	G023523

[00686] Prior to T cell treatment, LNPs encapsulating a sgRNA were diluted to 6.64 µg/mL in T cell treatment media (TCTM): a version of TCGM containing 20 ug/mL rhApoE3 in the absence of interleukins 2, 5 or 7. These LNPs were incubated at 37°C for 15

minutes and serially diluted 1:4 using TCTM, which resulted in an 8-point dilution series ranging from 6.64 µg/mL to zero. Similarly, single-cargo LNPs with BC22n mRNA (SEQ ID NO: 1) or UGI mRNA (SEQ ID NO: 34) were diluted in TCTM to 3.32 and 1.67 µg/mL, respectively, incubated at 37°C for 15 minutes, and mixed 1:1 by volume with sgRNA LNPs serially diluted in the previous step. Last, 50 µL from the resulting mix was added to T cells in 96-well plates at a 1:1 ratio by volume. T cells were incubated at 37 °C for 24 hours, at which time they were harvested, centrifuged at 500 g for 5 min, resuspended in 200 µL of TCGM and returned to the incubator.

Example 30.3. Evaluation of editing outcomes by next generation sequencing (NGS)

[00687] Four days post-LNP treatment, T cells were subjected to lysis, PCR amplification of each targeted locus and subsequent NGS analysis, as described in Example 1. Tables 75-80 and Figs. 48, 50A, 50B, 51 show editing levels and the C to T editing purity in T cells treated with a decreasing mass of 100-mer or 91-mer sgRNAs targeting TRAC, TRBC1, TRBC2, CIITA, B2M or CD38.

[00688] When compared to their 100-mer versions, 91-mer sgRNAs resulted in higher editing frequencies when delivered at the same concentration. This result was observed for all 5 different sets of sgRNAs assayed by NGS, which target 6 different genomic loci. No differences in C to T editing purity were observed between 100-mer and 91-mer sgRNAs. The set of sgRNAs targeting the HLA-A gene were evaluated by flow cytometry instead of NGS due to the hyperpolymorphic nature of the HLA-A locus.

Table 75 – Mean percent editing at the TRAC locus in T cells treated with sgRNAs in the 100-mer (G016017) or 91-mer format (G023520).

sgRNA (ng)	TRAC											
	100-mer						91-mer					
	C to T		C to A/G		Indels		C to T		C to A/G		Indels	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
166.00	96.7	0.4	1.0	0.2	1.5	0.3	94.9	0.0	2.2	0.2	2.5	0.3
41.50	98.0	0.5	0.4	0.1	0.5	0.4	97.5	0.1	0.8	0.0	1.2	0.1
10.38	98.4	0.4	0.2	0.1	0.5	0.1	98.4	0.1	0.4	0.1	0.7	0.1
2.59	97.2	0.1	0.3	0.1	0.4	0.1	98.2	0.2	0.3	0.1	0.5	0.0
0.65	78.3	0.9	0.3	0.1	0.5	0.0	92.0	0.5	0.3	0.1	0.4	0.1
0.16	37.4	0.9	0.1	0.0	0.2	0.1	55.3	0.7	0.2	0.0	0.3	0.1
0.04	11.6	0.9	0.2	0.1	0.1	0.1	19.4	0.9	0.2	0.1	0.2	0.1
0.00	0.9	0.2	0.1	0.1	0.1	0.0	0.8	0.1	0.2	0.0	0.1	0.0

Table 76 – Mean percent editing at the TRBC1 locus in T cells treated with sgRNAs in the 100-mer (G016200) or 91-mer format (G023524). .

sgRNA (ng)	TRBC1											
	100-mer						91-mer					
	C to T		C to A/G		Indels		C to T		C to A/G		Indels	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
166.00	95.1	0.5	1.3	0.2	1.9	0.2	91.6	1.0	3.4	0.1	4.0	0.8
41.50	96.2	0.8	0.6	0.1	1.3	0.3	95.3	0.6	1.5	0.2	2.1	0.3
10.38	94.6	0.6	0.4	0.1	1.0	0.2	96.0	0.4	0.8	0.2	1.1	0.1
2.59	83.5	0.7	0.5	0.1	0.7	0.3	92.1	0.6	0.6	0.1	0.9	0.5
0.65	48.1	2.0	0.4	0.1	0.3	0.1	69.8	0.1	0.4	0.2	0.4	0.2
0.16	17.3	0.7	0.3	0.1	0.3	0.4	34.0	0.5	0.4	0.0	0.3	0.1
0.04	6.2	1.2	0.3	0.1	0.2	0.1	10.9	0.7	0.5	0.1	0.1	0.1
0.00	1.3	0.1	0.4	0.1	0.2	0.1	1.2	0.4	0.3	0.2	0.1	0.1

Table 77– Mean percent editing at the TRBC2 locus in T cells treated with sgRNAs in the 100-mer (G016200) or 91-mer format (G023524).

sgRNA (ng)	TRBC2											
	100-mer						91-mer					
	C to T		C to A/G		Indels		C to T		C to A/G		Indels	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
166.00	96.5	0.1	1.2	0.3	1.3	0.3	92.4	0.4	3.2	0.1	3.7	0.3
41.50	97.3	0.5	0.7	0.0	0.9	0.0	96.1	0.3	1.3	0.1	1.7	0.5
10.38	95.3	0.9	0.6	0.2	1.0	0.1	97.1	0.1	0.6	0.1	1.2	0.2
2.59	78.3	1.4	0.6	0.0	0.6	0.1	92.8	0.8	0.7	0.1	0.7	0.2
0.65	41.1	1.1	0.3	0.0	0.4	0.1	66.5	2.0	0.4	0.3	0.5	0.2
0.16	13.9	1.5	0.3	0.1	0.2	0.0	29.7	2.5	0.3	0.2	0.3	0.2
0.04	5.1	0.7	0.3	0.0	0.3	0.1	10.3	0.5	0.4	0.1	0.3	0.1
0.00	1.5	0.4	0.4	0.1	0.2	0.1	1.8	0.3	0.4	0.1	0.2	0.1

Table 78 - Mean percent editing at the CIITA locus in T cells treated with sgRNAs in the 100-mer (G016086) or 91-mer format (G023521).

sgRNA (ng)	CIITA											
	100-mer						91-mer					
	C to T		C to A/G		Indels		C to T		C to A/G		Indels	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
166.00	98.1	0.2	0.7	0.1	0.6	0.1	96.9	0.3	1.3	0.0	1.4	0.2
41.50	97.3	0.2	0.4	0.1	0.3	0.1	98.1	0.2	0.6	0.1	0.8	0.1
10.38	82.9	0.3	0.3	0.0	0.2	0.1	97.5	0.4	0.5	0.2	0.4	0.1
2.59	43.3	0.5	0.3	0.1	0.2	0.1	88.0	1.1	0.4	0.1	0.3	0.1
0.65	14.5	0.9	0.3	0.1	0.1	0.1	50.4	1.9	0.3	0.1	0.1	0.1
0.16	4.3	0.2	0.2	0.1	0.0	0.0	17.7	0.5	0.3	0.0	0.1	0.1
0.04	1.8	0.2	0.3	0.1	0.1	0.0	5.7	0.1	0.2	0.1	0.1	0.0
0.00	0.6	0.1	0.2	0.1	0.1	0.0	0.8	0.0	0.2	0.1	0.1	0.0

Table 79- Mean percent editing at the B2M locus in T cells treated with sgRNAs in the 100-mer (G015991) or 91-mer format (G023519).

sgRNA (ng)	B2M											
	100-mer						91-mer					
	C to T		C to A/G		Indels		C to T		C to A/G		Indels	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
166.00	92.5	0.3	4.2	0.2	2.8	0.4	89.0	0.9	6.9	0.5	3.8	0.3
41.50	97.2	0.2	1.3	0.2	1.1	0.1	96.1	0.4	2.1	0.1	1.5	0.4
10.38	97.6	0.4	0.6	0.1	0.9	0.1	97.7	0.1	0.8	0.0	1.0	0.1
2.59	96.8	0.1	0.5	0.0	0.7	0.1	97.9	0.1	0.6	0.0	0.6	0.1
0.65	79.5	0.9	0.3	0.0	0.3	0.0	90.3	0.3	0.4	0.0	0.5	0.1
0.16	39.8	0.3	0.3	0.0	0.2	0.1	54.1	0.4	0.3	0.1	0.2	0.0
0.04	13.9	0.3	0.3	0.1	0.1	0.0	20.8	0.8	0.3	0.0	0.1	0.1
0.00	2.1	0.2	0.3	0.0	0.1	0.0	2.0	0.4	0.3	0.0	0.1	0.0

Table 80 - Mean percent editing at the CD38 locus in T cells treated with sgRNAs in the 100-mer (G019771) or 91-mer format (G023522).

sgRNA (ng)	CD38											
	100-mer						91-mer					
	C to T		C to A/G		Indels		C to T		C to A/G		Indels	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
166.00	92.7	0.6	1.6	0.2	5.4	0.4	91.0	0.6	2.5	0.2	6.1	0.5
41.50	93.9	0.4	0.7	0.1	5.1	0.3	93.4	0.5	1.0	0.0	5.2	0.5
10.38	93.7	0.1	0.7	0.1	4.8	0.2	94.0	0.4	0.6	0.1	4.6	0.2
2.59	88.4	0.9	0.5	0.1	3.3	0.3	92.4	0.7	0.6	0.1	3.8	0.3
0.65	54.8	1.6	0.4	0.0	1.9	0.3	67.3	0.3	0.4	0.1	2.3	0.0
0.16	19.7	0.4	0.5	0.1	0.7	0.1	27.8	0.4	0.5	0.1	1.0	0.2
0.04	5.9	0.5	0.5	0.0	0.3	0.1	9.1	0.6	0.4	0.1	0.4	0.1
0.00	0.4	0.0	0.4	0.1	0.1	0.0	0.5	0.1	0.4	0.0	0.0	0.0

Example 30.4. Evaluation of receptor knockout by flow cytometry

[00689] Seven days post LNP treatment, T cells were assayed by flow cytometry to evaluate receptor knockout. T cells were incubated with a fixable viability dye (Beckman Coulter, Cat. C36628) and an antibody cocktail targeting the following molecules: CD3 (Biolegend, Cat. 317336), CD4 (Biolegend, Cat. 317434) and CD8 (Biolegend, Cat. 301046), B2M (Biolegend, Cat. 316306), CD38 (Biolegend, Cat. 303516), HLA-A2 (Biolegend, Cat. 343304) and HLA-DR, DP, DQ (Biolegend, Cat. 361714). Cells were subsequently washed, analyzed on a Cytoflex LX instrument (Beckman Coulter) using the FlowJo software package. T cells were gated on size, viability and CD8 positivity before

expression of any markers was determined. The resulting data was plotted on GraphPad Prism v. 9.0.2 and analyzed using a variable slope (four parameter) non-linear regression.

[00690] As shown in Tables 81-83 and Figs. 52, 53A-55B, all 91-mer sgRNAs tested outperformed their 100-mer versions. As show in **Table 84,** this increase in potency was variable, with a 4.6-fold increase observed for CIITA and a 1.14-fold increase observed for B2M. Among the 6 targets tested, those with a lower potency (i.e., higher EC50) in the 100-mer format (CIITA and HLA-A) seem to benefit the most from usage of 91-mer sgRNAs.

Table 81 - Mean percentage of CD8+ T cells that are negative for CD3 surface receptors following sgRNA targeting TRAC or TRBC1/2 in the 100-mer or 91-mer formats.

sgRNA (ng)	TRAC (CD3-)				sgRNA (ng)	TRBC1/2 (CD3-)			
	100-mer		91-mer			100-mer		91-mer	
	Mean	SD	Mean	SD		Mean	SD	Mean	SD
166.00	98.4	0.3	98.2	0.3	166.00	99.6	0.3	99.1	0.3
41.50	98.1	0.1	98.5	0.5	41.50	99.8	0.1	99.6	0.1
10.38	97.7	0.2	98.0	0.4	10.38	98.8	0.4	99.9	0.1
2.59	95.2	0.4	97.4	0.6	2.59	84.9	1.1	97.4	0.6
0.65	74.7	1.0	88.0	0.8	0.65	45.7	1.6	73.5	2.8
0.16	33.5	2.4	51.0	1.0	0.16	16.4	0.9	31.6	0.8
0.04	11.0	0.7	18.3	0.8	0.04	5.5	0.4	10.2	0.5
0.00	1.1	0.2	1.3	0.2	0.00	1.2	0.2	1.2	0.1

Table 82 – Mean percentage of CD8+ T cells that are negative for HLA-DR, DP, DQ or HLA-A2 surface receptors following treatment sgRNA targeting CIITA or HLA-A, respectively, in the 100-mer or 91-mer formats.

sgRA (ng)	CIITA (HLA-DR, DP, DQ-)				sgRNA (ng)	HLA-A (HLA-A2-)			
	100-mer		91-mer			100-mer		91-mer	
	Mean	SD	Mean	SD		Mean	SD	Mean	SD
166.00	98.3	0.2	98.7	0.2	166.00	98.8	0.1	99.6	0.2
41.50	96.9	0.7	98.4	0.3	41.50	93.6	0.8	99.2	0.4
10.38	85.2	0.7	97.7	0.3	10.38	70.2	1.0	93.8	1.4
2.59	58.7	0.2	89.4	1.3	2.59	34.0	2.1	63.2	3.0
0.65	44.8	1.1	63.4	0.5	0.65	12.1	1.3	28.5	1.2
0.16	38.2	1.6	45.2	2.6	0.16	3.3	0.2	8.3	0.6
0.04	37.8	1.0	38.2	0.5	0.04	0.9	0.3	2.6	0.5
0.00	35.1	2.5	37.6	1.7	0.00	0.1	0.0	0.3	0.2

Table 83 – Mean percentage of CD8+ T cells that are negative for B2M or CD38 surface receptors following treatment with sgRNAs targeting B2M or CD38, respectively in the 100-mer or 91-mer formats.

sgRNA (ng)	B2M (B2M-)				sgRNA (ng)	CD38 (CD38-)			
	100-mer		91-mer			100-mer		91-mer	
	Mean	SD	Mean	SD		Mean	SD	Mean	SD
166.00	98.0	0.4	98.3	0.2	166.00	99.8	0.2	99.8	0.1
41.50	98.2	0.0	98.4	0.4	41.50	99.6	0.1	99.9	0.1
10.38	94.5	0.7	97.2	0.3	10.38	98.4	0.3	99.5	0.1
2.59	68.3	3.2	81.4	3.4	2.59	87.4	1.9	94.3	0.9
0.65	8.4	6.8	11.4	2.6	0.65	47.0	4.3	60.0	1.9
0.16	0.4	0.3	0.5	0.1	0.16	20.5	4.9	23.8	1.6
0.04	0.4	0.6	0.2	0.1	0.04	16.5	10.1	12.5	0.6
0.00	0.9	0.7	1.6	0.9	0.00	17.2	4.1	13.7	0.7

Table 84 – Amount (pmol) of sgRNAs that lead to a 50% loss of receptor expression in the surface of CD8+ T cells (EC50s). The far right column shows the fold-increase in potency achieved by 91-mer sgRNAs when compared to their 100-mer with the same guide sequence.

Gene target	100-mer		91-mer		EC50 shift (100-mer/91-mer)
	sgRNA ID	EC50 (pmols)	sgRNA ID	EC50 (pmols)	
TRAC	G016017	0.008	G023520	0.005	1.63
TRBC1/2	G016200	0.023	G023524	0.010	2.22
CIITA	G016086	0.123	G023521	0.027	4.60
B2M	G015991	0.055	G023519	0.048	1.14
CD38	G019771	0.027	G023522	0.020	1.41
HLA-A	G021209	0.150	G023523	0.053	2.81

Example 31 Base editing impact of UGI mRNA dose on *in-vivo* editing

[00691] *In vivo* liver editing profiles were evaluated using fixed doses of BC22n (**SEQ ID No: 1)** and guide RNA targeting ANAPC5 (G019427) along with a serial dilution of UGI mRNA (**SEQ ID No: 34)**.

Example 31.1 *In vivo* editing as assayed by NGS

[00692] Fifteen commercially available CD-1 female mice ranging from 6-10 weeks of age (n=3 per group) were used in this study. Animals were weighed pre-dose for dosing calculations. Each RNA species was formulated separately in an LNP. LNPs were formulated generally as described in Example 1. LNPs contained ionizable Lipid A, cholesterol, DSPC, and PEG2k-DMG in a 50:38:9:3 molar ratio, respectively. The lipid nucleic acid assemblies were formulated with a lipid amine to RNA phosphate (N:P) molar ratio of about 6. LNPs encapsulated a single RNA species, either G019427, BC22n mRNA (SEQ ID No: 1), or UGI mRNA (SEQ ID No. 34).

[00693] LNPs encapsulating base editor mRNA and LNPs encapsulating sgRNA were mixed and administered simultaneously at fixed doses of 0.2 mpk and 0.1 mpk RNA by weight, respectively (editor mRNA and sgRNA) along with UGI mRNA doses of either 0.0, 0.03, 0.1, and 0.3 mpk RNA by weight. The negative control group was dosed with TSS buffer only. Formulations were administered intravenously via tail vein injection according to the doses listed in **Table 85**.

Table 85– Animal groups and the respective LNPs

Sample	LNP cargos	Editor mRNA (mg/kg)	sgRNA (mg/kg)	UGI mRNA (mg/kg)
TSS	None	0	0	0
0.0 MPK UGI	BC22n mRNA	0.2 mg/kg	0.1 mg/kg	n/a
0.03 MPK UGI	UGI mRNA			0.03 mg/kg
0.1 MPK UGI	G019427			0.1 mg/kg
0.3 MPK UGI				0.3 mg/kg

[00694] Animals were periodically observed for adverse effects for at least 24 hours post-dose. Six days after treatment, animals were euthanized by cardiac puncture under isoflurane anesthesia; liver tissue were collected for downstream analysis. Liver punches weighing between 5 and 15 mg were collected for isolation of genomic DNA. Genomic DNA samples were analyzed with NGS sequencing as described in Example 1. Editing data and percent C to T purity are shown in **Table 86 and Figs. 56-57**.

Table 86 –Mean percent editing at the ANAPC5 locus in mouse liver with increasing amounts of UGI mRNA.

UGI mRNA (mpK)	% C to T		% C to A/G		Indel		%C to T Purity	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TSS	0.10	0.00	0.71	0.03	0.11	0.10	11.01	0.38
0	25.77	4.18	21.64	1.52	24.29	26.51	34.80	4.90
0.03	68.20	2.38	2.30	0.54	4.11	4.02	91.50	1.27
0.1	67.47	0.55	2.04	0.25	3.92	3.71	92.15	0.39
0.3	66.67	0.67	2.17	0.14	3.88	3.63	91.99	0.66

Example 32. Determination of BC22n to UGI ratio *in vitro*

[00695] Protein expression levels are determined using HiBiT tagged versions of a protein. To determine the relative amounts of base editor protein and UGI protein

expressed for efficient base editing with high C to T purity, parallel editing experiments are performed one of which uses an mRNA encoding a HiBiT-tagged base editor and the other of which uses an mRNA encoding a HiBiT -tagged UGI.

[00696] Cell preparation, engineering and editing assays are performed as in Example 9. One arm of the experiment (editor tagged) transfects with an mRNA encoding a HiBiT -tagged BC22n (SEQ ID No. 4) and an mRNA UGI (SEQ ID No: 34). Another arm of the experiment (UGI tagged) transfects with an mRNA encoding base editor (SEQ ID No: 1) and an mRNA encoding a Hibit-tagged UGI.

[00697] Twenty-four hours after transfection, the number of live T cells per sample is determined using a cell titer-glo assay (Promega Cat No. G7571) and the number of HiBiT tagged proteins is measured using a Nano-Glo® HiBiT lytic detection assay (Promega Cat No. N3040). NGS sequencing is performed 96 hours post-transfection to assess total editing and C to T purity levels. BC22n and UGI protein levels is plotted against total editing levels to determine the minimum number of each proteins cell that is required to achieve saturating editing levels and C to T purity.

[00698] The optimal ratio of BC22n to UGI proteins is calculated by dividing the minimum number of BC22 proteins per cell that is required to achieve saturating editing levels by the minimum number of UGI proteins per cell that is required to achieve maximum C to T purity.

[00699] Similar methods could be used to calculate the ratio of base editor to UGI protein that use different cell types, different guides, or different transfection methods.

Example 33. *In vivo* UGI titration with lower doses

[00700] Base editing is performed in mice to assess overall editing efficiency and C to T purity with lower levels of UGI mRNA. Experimental details are as described in **Example 30** with the following exceptions. LNPs encapsulating BC22n mRNA and LNPs encapsulating sgRNA are mixed at fixed doses of 0.2 mg/kg and 0.1 mg/kg of RNA cargos respectively, and combined with UGI mRNA doses of either 0.0, 0.0001, 0.001, 0.003, 0.01 0.03, 0.1, and 0.3 mg/kg. The negative control group is a TSS treated animal. Formulations are administered intravenously via tail vein injection. Six days after treatment, animals are euthanized by cardiac puncture under isoflurane anesthesia; liver tissue are collected for downstream analysis. Liver punches weighing between 5 and 15 mg are collected for isolation of genomic DNA and total RNA. Genomic DNA samples are analyzed with NGS

sequencing as described in Example 1 to determine the minimum UGI mRNA dose required for maximum C-to-T editing purity in mouse liver.

Example 34. Cytotoxic susceptibility of engineered T cells

[00701] Engineered T cells are assayed for cytotoxic susceptibility when targeted by natural killer (NK) cells.

[00702] NK cells (Stemcell Technologies) are thawed and resuspended at a cell concentration of 1x10^6 cells/ml into T cell growth media (TCGM) composed of OpTmizer TCGM and further supplemented with 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml IL-7 (Peprotech, Cat. 200-07), 5 ng/ml IL-15 (Peprotech, Cat. 200-15). Cells are incubated at 37˚C for 24 hours.

[00703] Twenty-four hours post thaw, the NK cells are labelled with 0.5 μM Cell Trace Violet as follows: a vial of Cell Trac e Violet (CellTrace™ Violet Cell Proliferation Kit, for flow cytometry, Cat. C34571) is reconstituted in DMSO from the kit to give a 5 mM stock concentration. Two μL of CTV stock is diluted with 18 μL Phosphate-Buffered Saline (Corning, Cat. 21-040-CV) to obtain a concentration of 0.5 mM. NK cells are centrifuged at 500xg for 5 minutes, the media is aspirated, and cells are resuspended in Phosphate-Buffered Saline (PBS) at a concentration of 1x10^6 cells/mL such that the final concentration of CTV dye is 0.5 μM. The cells are mixed with CTV (Cell Trace Violet) dye solution incubated at37˚C for 20 minutes. Unbound dye is quenched by the addition of TCGM and incubated for 5 minutes. The cells are centrifuged at 500xg for 5 minutes. Cells are resuspended in TCGM supplemented with 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml IL-7 (Peprotech, Cat. 200-07), 5 ng/ml IL-15 (Peprotech, Cat. 200-15) at a concentration of 2x10^6 cells/mL. To test a range of effector:target (E:T) ratios, CTV-labelled NK cells aliquoted in 100 ul of media in a 6 point, 2 fold serial dilution with the highest number of cells being 2 x 10^5 cells Media only samples are included as negative controls.

[00704] T cells are engineered using BC22n and UGI mRNA as described in **Example 14** using the G023523 targeting HLA-A or G015991 targeting B2M. Unedited (WT) T cells, Unstained, LD heat killed and 7-AAD FMO (2x10^4 unlabelled NK cells and 2x10^4 WT T cells) and CTV+ (2 x 10^5 CTV labelled NK cells) are also included as controls. T cells are resuspended at a density of 2x10^5 cells of TCGM composed of OpTmizer TCGM and further supplemented with 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml IL-7 (Peprotech, Cat. 200-07), 5 ng/ml IL-15

(Peprotech, Cat. 200-15). Twenty thousand T cells are added each well of NK cells and media controls. Cells are incubated at 37˚C for 24 hours.

[00705] At twenty-four hours, half of the volume of the cells from the LD Heat Killed well are heat killed and transferred back to the same well in the assay plate. Cells are centrifuged and resuspended in in 80 µL of a 1:200 v/v solution of 7-AAD (BD Biosciences, Cat. 559925) in FACS buffer (PBS + 2% FBS (Gibco, Cat. A31605-02) + 2mM EDTA (Invitrogen, Cat. 15-575-020). Data for specific lysis of T cells is acquired by flow cytometry using a Cytoflex LX instrument (Beckman Coulter) and analyzed using the FlowJo software package. Gates are first drawn on the CTV negative population to gate out the NK cells, followed by gating on singlets after which a gate is drawn on the 7-AAD negative population to gate for the live T cells. The percent lysis of T cells is calculated by subtracting the live cell percentage from 100.

Example 35. this paragraph is intentionally left blank.

Example 36. Protein expression of base editor and UGI in Primary Human Hepatocytes

[00706] Base editing was performed with mRNAs encoding proteins with fused HiBiT tags in primary human hepatocytes to determine the relative amounts of base editor and UGI protein enzymatic units appropriate for high efficiency editing with high C to T purity.

[00707] Messenger RNAs encoding BC22n with a C-terminal HiBiT tag (BC22n-HiBIT, SEQ ID NO: 4), BC22-2XUGI (2 tandem copies of UGI *in cis*) with a C-terminal HiBiT tag (BC22-2XUGI-HibIT, SEQ ID NO: 314), and UGI with a C-terminal HiBiT tag (UGI-HiBiT, SEQ ID NO: 316), were transfected into primary human hepatocytes (PHH) in dose response with a fixed concentration of a guide RNA targeting B2M (G015991; SEQ ID NO: 179) to determine the minimum number of intracellular base editor and UGI protein copies to perform base editing with high activity (total editing % = 2X EC90) and high C-to-T purity (C to T purity % = 2X EC90). BC22n-HiBiT was titrated across fixed concentrations of B2M guide and UGI (SEQ ID NO: 34), UGI-HiBiT was titrated across fixed concentrations of B2M guide and BC22n (SEQ ID NO: 1), and BC22-2XUGI-HiBiT was titrated across a fixed concentration of B2M guide without any additional UGI mRNA *in trans*. A HiBiT lytic assay (Promega cat # N3040) was performed twenty hours post transfection to determine intracellular protein levels at an early timepoint. Those protein levels were then related to endpoint editing at the B2M locus, derived from NGS data, ninety-six hours post transfection.

Example 36.1 Cell preparation and transfection

[00708] PHH cells (ThermoFisher, Lot HU8284) were thawed and recovered in CHRMs media (Gibco, cat#CM7000). Cells were resuspended in Primary Hepatocyte Plating Media (Consisting of William's E media (Gibco, Cat # A1217601) and primary hepatocyte plating supplements (Gibco, Cat # CM3000)) before being plated in collagen-coated 96-well plates at a density of 30,000 cells/well for twenty-four hours. Cells were washed, and fresh primary hepatocyte maintenance media was added.

[00709] Lipofection reagent was prepared as described in Example 1, with a molar ratio of 50/9/38/3 of Lipid A, DSPC, cholesterol, and PEG2k-DMG respectively. Each RNA species (base editor mRNA, UGI mRNA, or gRNA G015991 (SEQ ID NO: 179)) was individually bulk-mixed with lipofection reagent at a lipid amine to RNA phosphate (N:P) molar ratio of about 6. The resulting bulk-mixed lipoplex material was pre-incubated with 10% FBS (Gibco, Cat # A3160501) in Primary Hepatocyte Maintenance Media (Consisting of William's E media (Gibco, Cat # A1217601) and primary hepatocyte maintenance supplements (Gibco, Cat # CM4000)), for 15 minutes before addition to hepatocytes.

[00710] For the BC22n-HiBiT titration, each well received three components in a final volume of 100 uL: BC22n-HiBiT mRNA ranging from 100 ng to 0 ng, 11 ng of UGI mRNA and 2 pmols of G015991 (B2M) as described in **Table 87**. For the UGI-HiBiT titration each well received three components in a final volume of 100 uL: UGI-HiBiT mRNA ranging from 100 ng to 0 ng, 33 ng of BC22n mRNA and 2 pmols of G015991 (B2M) as described in **Table 88**. For the BC22-2XUGI-HiBiT titration each well received three components in a final volume of 100 uL: BC22-2XUGI-HiBiT mRNA ranging from 100 ng to 0 ng, 11 ng of UGI mRNA and 2 pmols of G015991 (B2M) as described in **Table 89**.

[00711] Ninety-six hours post transfection, separate PHH replicate plates that had been transfected at the same time as the HiBiT plates were subjected to lysis, PCR amplification of the B2M locus, and subsequent NGS analysis, as described in Example 1. All experiments were performed in biological triplicate. Background C-to-T, C-to-A/G and indel edits (all < 2%) were subtracted from all wells by calculating the average background rates from a set of untreated wells. Editing results for different BC22n-HiBiT mRNA concentrations are shown in **Table 87**, and total editing at different BC22n-HiBiT mRNA concentrations are shown in **Fig. 58A**. Editing results for different UGI-HiBiT mRNA

concentrations are shown in **Table 88** and C-to-T purity at different UGI-HiBiT mRNA concentrations are shown in **Fig. 58B**. Editing results for different BC22-2XUGI-HiBiT mRNA concentrations are shown in **Table 89**. Total editing and C-to-T purity with different BC22-2XUGI-HiBiT mRNA concentrations are shown in **Fig. 58C** and **Fig. 58D**, respectively.

[00712] To facilitate comparisons between mRNAs of different sizes, all figures display mRNA doses in terms of molarity. For these purposes, the molecular weight of BC22n mRNA was considered as 1,724.25 kDa, BC22-2X-UGI mRNA as 1915.12 kDa, UGI mRNA as 287.58 kDa and sgRNAs as 29.66 kDa. The addition of a HiBiT tag to any mRNA increased its molecular weight by 18.24 kDa.

Table 87 – Mean editing in PHH with increasing doses of BC22n-HiBiT mRNA.

Dose (ng)	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
100.00	0.2857	3	84.87	1.37	0.95	0.28	0.80	0.61	86.74	1.83	97.99	0.93
33.00	0.0952	3	78.63	0.70	0.73	0.19	0.93	0.29	80.08	0.94	97.93	0.50
11.00	0.0317	3	66.33	3.72	0.46	0.24	0.72	0.39	67.19	3.67	98.27	0.39
3.70	0.0106	3	45.33	1.25	0.19	0.21	0.40	0.21	46.43	1.70	98.72	0.25
1.23	0.0035	3	26.57	0.99	0.32	0.29	0.90	0.58	26.92	1.29	95.69	2.67
0.41	0.0012	3	11.77	0.61	0.12	0.21	0.03	0.05	11.96	0.72	98.75	1.55
0.14	0.0004	3	5.43	0.85	0.00	0.00	0.08	0.14	5.47	0.82	98.58	2.45
0.00	0.0000	3	0.09	0.13	0.07	0.13	0.03	0.03	1.25	0.55	NA	NA

Table 88 – Mean editing in PHH with increasing doses of UGI-Hibit mRNA.

Dose (ng)	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
100.00	1.6400	3	81.10	3.33	0.80	0.13	1.09	0.42	83.32	3.34	97.74	0.52
25.00	0.4100	1	80.73	2.17	0.52	0.32	1.43	0.11	82.40	1.81	97.64	0.49
6.25	0.1025	3	77.10	3.71	0.64	0.55	1.15	0.61	79.19	4.40	97.74	0.39
1.56	0.0256	3	78.63	2.72	1.92	0.22	1.45	0.53	83.53	3.42	95.91	0.61
0.39	0.0064	3	74.87	1.88	3.15	0.04	2.98	0.71	82.23	1.91	92.45	0.70
0.10	0.0016	3	73.60	2.70	5.62	1.25	4.21	0.53	85.06	2.51	88.22	2.18
0.02	0.0004	3	67.33	0.46	7.30	1.64	5.84	0.61	80.15	1.21	83.67	1.18
0.00	0.0000	3	71.47	2.72	6.99	1.28	5.51	0.63	79.80	3.99	85.15	1.16

Table 89 – Mean editing in PHH with increasing doses of BC22-2XUGI-Hibit mRNA.

Dose (ng)	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD

Dose (ng)	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
100.00	0.2575	3	81.90	2.37	1.47	0.87	1.34	0.20	84.29	1.23	96.67	0.85
33.00	0.0858	3	79.70	0.49	1.41	0.81	0.92	0.37	82.58	1.13	97.17	1.15
11.00	0.0286	3	65.93	0.75	1.70	0.30	1.47	0.19	68.52	0.96	95.41	0.20
3.70	0.0095	3	47.90	1.01	1.45	0.72	0.89	0.29	50.21	0.91	95.34	1.32
1.23	0.0032	3	27.77	2.64	1.45	0.74	0.86	0.60	29.52	2.34	92.39	3.92
0.41	0.0011	3	14.90	3.07	0.79	0.35	0.30	0.35	15.71	3.41	93.63	2.84
0.14	0.0004	3	7.30	0.74	0.19	0.15	0.02	0.02	7.51	0.86	97.33	1.89
0.00	0.0000	3	0.09	0.13	0.06	0.10	0.02	0.03	0.15	0.10	NA	NA

Example 36.2 Evaluation of intracellular protein levels

[00713] Twenty hours post transfection, protein levels were detected using the Nano-Glo® HiBiT Lytic Detection System (Promega cat # N3040) according to the manufacture's protocol. Promega #N3010, a commercially available control protein with a HiBiT tag (of known concentration and known molecular weight) was serially diluted (1:5) in PBS and spiked into wells containing PHH that had not been transfected as a reference control. One hundred µL of reconstituted HiBiT lytic reagent was added to the standard wells and all other experiment wells. Lysates were moved to white-walled plates, and relative luminescence units (RLU) were read out by the CLARIstar plus (BMG Labtech) plate reader with gain set at 3,600. Background signal was subtracted from all wells, and the number of protein copies per well for each sample was calculated using the linear regression equation derived by the HiBiT standard. HiBiT quantitation is presented in **Table 90** and normalized to the expression level of the 3.7 ng BC22-2XUGI-HiBiT mRNA sample**.**

Table 90 – Mean protein expression (arbitrary units) for Hibit-tagged proteins.

Hibit mRNA	Dose (ng)	Dose (nM)	Mean	SD
BC22n-HiBiT	100.00	0.2857	22.59	2.05
	33.00	0.0952	10.07	1.04
	11.00	0.0317	4.12	0.59
	3.70	0.0106	1.09	0.02
UGI-HiBiT	100.00	1.6400	4063.15	532.85
	25.00	0.4100	993.10	76.48
	6.25	0.1025	254.04	26.29
	1.56	0.0256	67.16	7.73
	0.39	0.0064	21.66	1.83
BC22-2XUGI-HiBiT	100.00	0.2575	25.82	2.08

Hibit mRNA	Dose (ng)	Dose (nM)	Mean	SD
	33.00	0.0858	9.13	1.15
	11.00	0.0286	2.87	0.72
	3.70	0.0095	1.00	0.27

[00714] The data in **Table 90** were used to generate hyperbolic curves and interpolate the protein units expressed at the minimum dose required to achieve saturating editing or C-to-T purity levels. The minimum dose as saturating levels is defined here as twice the EC90 (2X EC90), the concentration required to achieve 90% total editing or C-to-T purity, respectively, for each mRNA tested. **Table 91** shows the 2X EC90 doses (in ng and nM) of the mRNAs tested in this experiment and the BC22 and UGI peptide units and their ratios at these doses. For base editors and UGI-HiBiT, protein units are equal to peptide units. UGI peptide units for BC22-2XUGI were calculated by multiplying the BC22-2XUGI protein units by a factor of 2. For the samples edited with BC22n and UGI expressed as separate, unfused proteins, the ratio of BC22n peptide units at total editing 2x EC90 to the UGI peptide units at C to T purity 2X EC90 was approximately 1:10. For the samples edited with BC22-2XUGI, the ratio of base editor peptide units at total editing 2x EC90 to the UGI peptide units at C to T purity 2X EC90 was approximately 1:5.

Table 91 – 2X EC90 for mRNA and protein at relevant editing readouts.

mRNA	Editing Readout	mRNA (ng)	mRNA (nM)	Peptide units
BC22n-HiBiT	total editing	53.97	0.16	14.94
UGI-HiBiT	C-to-T purity	3.85	0.06	156.38
BC22-2XUGI-HiBiT	total editing (base editor)	49.89	0.13	13.39
	C-to-T purity (UGI)	139.23	0.36	70.14

Example 37. Protein expression of base editor and UGI in T cells

[00715] Base editing was performed with mRNAs encoding proteins with fused HiBiT tags in isolated human T cells to determine the relative amounts of base editor and UGI protein enzymatic units appropriate for high efficiency editing with high C to T purity.

Example 37.1. T cell preparation

[00716] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed, re-suspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. 130-070-525) and processed in a MultiMACS™ Cell 24 Separator Plus device (Miltenyi Biotec). T cells were isolated via positive selection using a Straight from Leukopak® CD4/CD8 MicroBead kit, human (Miltenyi Biotec Cat. 130-122-352). T cells were aliquoted and cryopreserved for future use in Cryostor® CS10 (StemCell Technologies Cat. 07930).

[00717] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth media (TCGM) composed of CTS OpTmizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512) 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/ml recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/ml recombinant human interleukin 15 (Peprotech, Cat. 200-15). After 24 hours in this media. T cells were activated with T Cell TransAct™, human reagent (Miltenyi, Cat. 130-111-160) added at a 1:100 ratio by volume.

Example 37.2. T cell LNP treatment and expansion

[00718] Forty-eight hours post-activation, T cells were harvested, centrifuged at 500 g for 5 min, and resuspended at a concentration of 1 x 10^6 T cells/mL in T cell plating media (TCPM): a serum-free version of TCGM containing 400 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 10 ng/ml recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 10 ng/ml recombinant human interleukin 15 (Peprotech, Cat. 200-15). Fifty thousand T cells in 50 ul TCPM were added per well to be treated in flat-bottom 96-well plates.

[00719] LNPs were generated as described in Example 1 at a molar ratio of 35/47.5/15/2.5 (Lipid A/ cholesterol/DSPC/PEG2k-DMG). Prior to T cell treatment, two separate LNP mixes (referred below as mixes “A” and “B”) were prepared in T cell treatment media (TCTM): a version of TCGM containing 20 ug/mL rhApoE3 (Peprotech, Cat. 350-02) in the absence of interleukins 2, 5 or 7.

[00720] To determine the minimum levels of BC22n-HiBiT mRNA necessary for saturating total editing, mix “A” consisted of an LNP with BC22n-HiBiT mRNA (SEQ ID NO: 4) diluted to 6.68 µg/mL (3.83 nM), while mix “B” consisted of an LNP with UGI mRNA (SEQ ID NO: 34) diluted to 1.67 µg/mL (5.8 nM) and a multi-cargo LNP containing

sgRNAs G023520 (TRAC), G023524 (TRBC1/2), G023521 (CIITA) and G023523 (HLA-A) at the ratio of 4.5/8.7/21.4/65.4 diluted to 6.68 µg/mL (225.19 nM). LNP mixes "A" and "B" were individually incubated at 37°C for 15 minutes. Mix "A" was serially diluted 1:4 in TCTM, and mixed 1:1 by volume with mix "B". The resulting solution was added to T cells in 96-well plates at a 1:1 ratio by volume (50 µL per well).

[00721] To determine the minimum levels of UGI-HiBiT mRNA necessary for saturating C to T purity,, mix "A" consisted of an LNP with UGI-HiBiT mRNA (SEQ ID NO: 316) diluted to 6.68 µg/mL (21.84 nM), while mix "B" consisted of an LNP with BC22n mRNA (SEQ ID NO: 1) diluted to 3.34 µg/mL (1.94 nM) and a multi-cargo LNP containing sgRNAs G023520 (TRAC), G023524 (TRBC1/2), G023521 (CIITA) and G023523 (HLA-A) at the ratio of 4.5/8.7/21.4/65.4 diluted to 6.68 µg/mL (225.19 nM). LNP mixes "A" and "B" were individually incubated at 37°C for 15 minutes. Mix "A" was serially diluted 1:4 in TCTM, and mixed 1:1 by volume with mix "B". The resulting solution was added to T cells in 96-well plates at a 1:1 ratio by volume (50 µL per well).

[00722] To determine the minimum levels of BC22-2XUGI-HiBiT mRNA necessary for saturating total editing and C to T purity, mix "A" consisted of an LNP with BC22-2XUGI-HiBiT mRNA (SEQ ID NO: 314) diluted to 6.68 µg/mL (3.46 nM), while mix "B" consisted of a multi-cargo LNP containing sgRNAs G023520 (TRAC), G023524 (TRBC1/2), G023521 (CIITA) and G023523 (HLA-A) at the ratio of 4.5/8.7/21.4/65.4 diluted to 6.68 µg/mL (225.19 nM). LNP mixes "A" and "B" were incubated individually at 37°C for 15 minutes. Mix "A" was serially diluted 1:4 in TCTM, and mixed 1:1 by volume with mix "B". The resulting solution was added to T cells in 96-well plates at a 1:1 ratio by volume (50 µL per well).

[00723] Following the addition of LNPs, T cells were incubated at 37 °C for 24 hours, at which time half of the cells were harvested for cell viability and protein expression assays, and the remaining cells were centrifuged at 500 g for 5 min, resuspended in 200 µL of TCGM and returned to the incubator.

Example 37.3. Evaluation of editing outcomes by next generation sequencing (NGS)

[00724] On day 4 post-LNP treatment, T cells were centrifuged at 500 g for 5 min, subjected to lysis, PCR amplification of each targeted locus and subsequent NGS analysis, as described in Example 1. Editing results at the TRAC, TRBC1, TRBC2, and CIITA loci with different concentrations of HiBiT mRNAs are shown in Tables 92-95,

respectively. Fig. 59A shows total editing at different BC22n-HiBiT mRNA concentrations. Fig. 59B shows C-to-T purity at different UGI-HiBiT mRNA concentrations. Total editing and C-to-T purity with different BC22-2XUGI-HiBiT mRNA concentrations are shown in Fig. 59C and Fig. 59D, respectively.

[00725] Tables 92-95 show editing data for 4 different genomic loci (TRAC, TRBC1, TRBC2 and CIITA).

Table 92. Mean editing at TRAC locus in T cells.

mRNA	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
BC22n-HiBiT	0.9584	3	97.29	0.35	0.84	0.29	0.60	0.36	98.72	0.16	98.55	0.22
	0.2396	3	97.82	0.31	0.58	0.27	0.67	0.20	99.07	0.32	98.74	0.07
	0.0599	3	96.50	0.50	0.79	0.22	0.68	0.15	97.97	0.31	98.50	0.36
	0.0150	3	87.53	0.41	0.68	0.09	0.45	0.31	88.66	0.14	98.72	0.39
	0.0037	3	50.82	2.39	0.57	0.13	0.32	0.15	51.71	2.22	98.27	0.49
	0.0009	3	17.33	1.00	0.46	0.17	0.17	0.13	17.97	1.18	96.50	1.09
	0.0002	3	5.31	0.28	0.37	0.10	0.19	0.07	5.88	0.21	90.36	1.67
	0.0000	3	0.67	0.08	0.44	0.13	0.13	0.04	1.24	0.18	NA	NA
UGI-HiBiT	5.4608	3	97.00	0.65	0.81	0.37	0.37	0.23	98.18	0.38	98.80	0.29
	1.3652	1	97.22	ND	0.28	ND	1.39	ND	98.89	ND	98.31	ND
	0.3413	3	95.90	0.49	1.43	0.16	1.01	0.21	98.35	0.58	97.51	0.18
	0.0853	3	82.34	3.24	7.45	2.31	8.06	0.86	97.85	0.15	84.15	3.22
	0.0213	3	55.53	1.88	17.35	3.45	23.98	2.08	96.86	0.39	57.34	2.17
	0.0053	3	32.80	2.32	24.36	3.54	38.14	1.67	95.30	1.10	34.44	2.78
	0.0013	3	26.28	0.77	28.61	2.18	41.47	1.09	96.36	0.43	27.28	0.91
	0.0000	3	22.46	1.08	28.38	3.67	44.56	2.27	95.40	0.37	23.55	1.22
BC22-2XUGI-HiBiT	0.8638	3	81.31	2.75	2.41	0.03	2.10	0.85	85.82	1.90	94.73	1.12
	0.2159	3	67.03	1.79	5.49	1.44	5.52	0.11	78.03	1.36	85.90	1.92
	0.0540	3	28.40	1.39	9.46	0.96	14.74	0.45	52.60	1.77	53.99	1.31
	0.0135	3	12.89	1.33	6.69	0.29	11.11	0.84	30.69	1.91	41.93	2.10
	0.0034	3	4.87	0.42	2.68	0.10	5.99	0.61	13.54	0.80	35.97	1.75
	0.0008	3	1.86	0.23	0.85	0.19	1.51	0.30	4.22	0.47	44.33	6.34
	0.0002	3	0.82	0.21	0.48	0.01	0.77	0.18	2.07	0.04	39.42	9.51
	0.0000	3	0.63	0.05	0.37	0.08	0.13	0.09	1.13	0.13	NA	NA

Table 93. Mean editing at TRBC1 locus in T cells.

mRNA	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
BC22n-HiBiT	0.9584	3	94.22	1.08	1.41	0.30	1.71	0.25	97.33	0.57	96.80	0.56
	0.2396	3	93.82	0.51	1.35	0.28	1.25	0.12	96.41	0.22	97.31	0.31
	0.0599	3	94.12	0.66	1.24	0.15	1.43	0.38	96.80	0.53	97.24	0.24
	0.0150	3	88.80	1.47	1.30	0.28	0.86	0.19	90.96	1.38	97.62	0.14
	0.0037	3	61.60	1.71	1.13	0.21	0.75	0.32	63.48	2.11	97.05	0.55

mRNA	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	0.0009	3	24.05	2.20	1.08	0.09	0.34	0.20	25.47	2.36	94.43	0.19
	0.0002	3	8.54	0.72	1.11	0.08	0.30	0.05	9.95	0.71	85.79	1.62
	0.0000	3	2.52	0.08	1.09	0.09	0.30	0.05	3.92	0.07	NA	NA
UGI-HiBiT	5.4608	3	93.60	1.03	1.45	0.08	1.38	0.26	96.44	0.90	97.06	0.26
	1.3652	3	93.90	0.77	1.59	0.16	1.57	0.49	97.05	0.33	96.75	0.47
	0.3413	3	89.15	1.94	4.27	1.30	2.99	0.98	96.41	0.34	92.48	2.33
	0.0853	3	72.64	2.56	12.64	0.23	11.89	3.35	97.16	0.88	74.78	3.31
	0.0213	3	40.67	3.56	24.32	0.89	31.25	2.94	96.25	1.16	42.23	3.22
	0.0053	3	17.56	3.20	32.38	1.05	42.82	2.69	92.76	1.49	18.90	3.14
	0.0013	3	11.59	2.11	34.92	3.79	48.02	5.70	94.53	1.67	12.27	2.31
	0.0000	3	8.51	1.36	33.48	1.78	51.60	2.15	93.60	1.31	9.08	1.34
BC22-2XUGI-HiBiT	0.8638	3	83.49	0.86	4.26	0.27	3.64	0.71	91.39	1.58	91.36	0.93
	0.2159	3	68.57	2.32	10.23	1.51	7.39	1.54	86.19	1.60	79.54	1.23
	0.0540	3	35.21	1.71	20.43	1.18	20.76	0.94	76.41	2.69	46.07	0.61
	0.0135	3	20.21	1.22	18.44	1.47	18.28	1.63	56.92	2.62	35.56	2.76
	0.0034	3	10.88	0.49	9.98	1.82	8.13	3.43	28.99	5.28	38.36	6.83
	0.0008	3	4.48	1.11	2.35	0.47	3.64	1.34	10.46	2.87	43.01	2.64
	0.0002	3	3.39	0.37	1.71	0.23	0.97	0.69	6.07	0.83	56.10	4.73
	0.0000	3	2.81	0.22	1.15	0.09	0.34	0.07	4.30	0.26	NA	NA

Table 94. Mean editing at TRBC2 locus in T cells

mRNA	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	0.9584	3	94.58	0.49	1.89	0.43	1.90	0.15	98.37	0.20	96.14	0.31
	0.2396	3	95.25	0.63	1.68	0.04	1.20	0.26	98.13	0.42	97.06	0.26
	0.0599	3	94.01	0.63	1.36	0.09	1.63	0.44	97.00	0.58	96.92	0.37
	0.0150	3	88.99	0.35	1.39	0.27	1.04	0.19	91.41	0.08	97.35	0.45
	0.0037	3	56.12	2.38	1.29	0.12	0.74	0.30	58.15	2.64	96.51	0.61
	0.0009	3	20.90	0.73	1.20	0.25	0.40	0.09	22.50	0.59	92.85	1.31
BC22n-HiBiT	0.0002	3	8.08	1.04	1.11	0.15	0.32	0.09	9.51	0.94	84.81	3.36
	0.0000	3	2.97	0.24	1.08	0.10	0.49	0.24	4.55	0.51	NA	NA
UGI-HiBiT	5.4608	3	94.87	0.14	1.43	0.21	1.06	0.20	97.35	0.12	97.45	0.02
	1.3652	1	94.76	0.11	1.67	0.13	1.51	0.09	97.94	0.14	96.76	0.20
	0.3413	3	91.81	0.45	3.34	0.40	3.16	0.47	98.31	0.20	93.39	0.27
	0.0853	3	73.09	0.73	12.14	0.96	12.56	0.15	97.79	0.41	74.74	0.95
	0.0213	3	39.77	0.07	23.73	1.69	32.65	1.19	96.15	1.22	41.37	0.53
	0.0053	3	19.66	0.29	29.74	0.81	46.54	0.63	95.93	0.45	20.49	0.20
	0.0013	3	11.79	1.55	31.00	1.01	52.42	3.12	95.21	0.94	12.40	1.73
	0.0000	3	9.04	1.42	32.73	1.37	53.94	1.88	95.71	1.35	9.45	1.57
BC22-2XUGI-HiBiT	0.8638	3	85.85	2.78	4.34	0.88	3.57	0.69	93.76	1.23	91.55	1.76
	0.2159	3	68.48	2.35	10.64	0.47	8.73	0.23	87.85	1.95	77.93	0.95
	0.0540	3	34.87	2.26	19.85	1.23	22.41	1.15	77.12	0.42	45.21	2.75
	0.0135	3	17.86	0.57	16.46	0.44	18.42	1.68	52.74	2.57	33.89	0.95

mRNA	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	0.0034	3	9.52	1.12	8.01	0.14	8.16	0.27	25.69	1.10	37.01	2.90
	0.0008	3	5.17	0.59	3.15	0.38	3.19	0.19	11.50	0.65	44.83	3.28
	0.0002	3	3.06	0.16	1.38	0.14	1.26	0.35	5.71	0.54	53.91	3.87
	0.0000	3	2.79	0.05	1.11	0.22	0.35	0.12	4.25	0.28	NA	NA

Table 95. Mean editing at CIITA locus in T cells

mRNA	Dose (nM)	n	C-to-T		C-to-A/G		Indels		Total editing		C-to-T purity	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
BC22n-HiBiT	0.9584	3	96.75	0.28	1.68	0.18	0.57	0.07	99.00	0.05	97.72	0.23
	0.2396	3	96.80	0.15	1.51	0.25	0.53	0.11	98.83	0.24	97.94	0.14
	0.0599	3	96.31	0.64	1.39	0.17	0.42	0.17	98.12	0.43	98.15	0.31
	0.0150	3	89.04	1.09	1.19	0.07	0.41	0.07	90.63	0.99	98.24	0.16
	0.0037	3	56.77	1.94	1.24	0.18	0.18	0.02	58.18	1.79	97.56	0.38
	0.0009	3	22.38	0.53	1.29	0.07	0.12	0.07	23.78	0.60	94.09	0.13
	0.0002	3	7.16	0.27	1.19	0.14	0.16	0.05	8.52	0.35	84.06	1.27
	0.0000	3	0.88	0.06	1.16	0.27	0.05	0.02	2.09	0.22	NA	NA
UGI-HiBiT	5.4608	3	96.10	0.31	1.49	0.20	0.45	0.29	98.04	0.13	98.02	0.37
	1.3652	1	96.09	0.59	1.69	0.25	0.92	0.25	98.69	0.13	97.36	0.51
	0.3413	3	94.04	0.58	3.10	0.29	1.70	0.58	98.83	0.12	95.15	0.65
	0.0853	3	80.13	1.16	8.99	0.79	9.57	0.60	98.69	0.06	81.20	1.17
	0.0213	3	51.75	1.05	20.78	0.75	25.48	0.97	98.02	0.40	52.80	1.28
	0.0053	3	32.69	1.17	27.04	1.89	37.48	0.98	97.21	0.34	33.63	1.14
	0.0013	3	26.59	0.34	26.78	1.72	43.51	1.10	96.88	1.06	27.45	0.37
	0.0000	3	25.47	1.76	27.78	1.86	43.47	3.32	96.72	0.63	26.33	1.72
BC22-2XUGI-HiBiT	0.8638	3	86.31	2.68	3.78	1.26	2.13	0.51	92.22	1.69	93.59	1.81
	0.2159	3	67.79	6.14	9.39	1.85	9.25	3.55	86.43	1.80	78.37	5.68
	0.0540	3	41.54	1.66	17.05	0.26	21.16	1.07	79.75	2.54	52.08	0.85
	0.0135	3	26.63	0.45	14.36	1.00	19.33	2.27	60.32	3.17	44.23	2.37
	0.0034	3	12.13	1.48	6.22	1.17	9.19	0.82	27.54	2.50	44.09	4.45
	0.0008	3	3.99	0.46	2.94	0.39	2.94	0.45	9.87	1.00	40.47	3.90
	0.0002	3	1.40	0.38	1.56	0.34	0.83	0.23	3.79	0.51	36.41	5.18
	0.0000	3	0.67	0.10	1.10	0.16	0.09	0.08	1.86	0.17	NA	NA

Example 37.4. Evaluation of intracellular protein levels

[00726] Twenty-four hours post-LNP treatments, all samples were mixed thoroughly and two aliquots of 25 µL were transferred to white-walled 96-well plates containing 75 µL of TCGM. One plate was subjected to a CellTiter-Glo® 2.0 Cell Viability Assay (Promega Cat No. G9242) and the other plate was subjected to a Nano-Glo® HiBiT Lytic Detection Assay (Promega Cat No. N3040). Both assays were performed following the

manufacturer's protocol. Standards curves were prepared using known numbers of donor-matched T cells or a commercially available protein with a HiBiT tag (Promega Cat No. N3010). HiBiT quantitation is presented in **Table 96** normalized to the expression level of the 0.0037 nM BC22n-HiBiT mRNA sample**.**

Table 96. Mean protein expression (arbitrary units) for HiBiT-tagged proteins.

mRNA	Dose (nM)	n	Average	SD
BC22n-HiBiT	0.9584	3	51.93	2.20
	0.2396	3	33.39	4.42
	0.0599	3	15.40	1.34
	0.0150	3	4.05	0.07
	0.0037	3	1.00	0.05
UGI-HiBiT	5.4608	3	1499.52	106.66
	1.3652	3	583.54	53.95
	0.3413	3	180.50	16.40
	0.0853	3	55.38	3.55
	0.0213	3	16.05	1.26
	0.0053	3	4.32	0.32
BC22-2XUGI-HiBiT	0.8638	3	80.23	4.45
	0.2159	3	43.95	10.80
	0.0540	3	14.74	1.31
	0.0135	3	4.17	0.66
	0.0034	3	1.32	0.15

[00727] Data from Table 96 was used to generate hyperbolic curves and interpolate the the protein units expressed at the minimum doses required to achieve saturating editing or C-to-T purity levels. The minimum dose as saturating levels is defined here as twice the EC90 (2X EC90), the concentration required to achieve 90% total editing or C-to-T purity, respectively, for each mRNA tested.

[00728] Table 97 shows the 2X EC90 doses of the mRNAs tested in this experiment and the BC22 and UGI peptide units and their ratios at these doses. For base editors and UGI-HiBiT, protein units are equal to peptide units. UGI peptide units for BC22-2XUGI were calculated by multiplying the BC22-2XUGI protein units by a factor of 2. Table 98 shows the ratio of base editor peptide units at total editing 2x EC90 to the UGI peptide units at C to T purity 2X EC90.

Table 97. 2X EC90 for mRNA and protein at relevant editing readouts.

HiBiT mRNA	Editing Readout	Locus	nM	Peptide units
BC22n-HiBiT	Total Editing	TRAC	0.038	10.07
		TRBC1	0.031	8.41
		TRBC2	0.033	8.86
		CIITA	0.035	9.32
		Average	0.034	9.16
		SD	0.003	0.71
UGI-HiBiT	C-to-T purity	TRAC	0.354	177.22
		TRBC1	0.560	271.33
		TRBC2	0.557	269.91
		CIITA	0.461	227.02
		Average	0.483	236.37
		SD	0.098	44.48
BC22-2XUGI-HiBiT	Total Editing	TRAC	1.979	95.56
		TRBC1	0.480	65.60
		TRBC2	0.434	62.84
		CIITA	0.334	55.54
		Average	0.807	69.88
		SD	0.784	17.63
	C-to-T purity	TRAC	0.859	160.35
		TRBC1	1.024	168.00
		TRBC2	1.103	171.04
		CIITA	1.660	185.66
		Average	1.161	171.26
		SD	0.348	10.60

Table 98. Base editor to UGI enzyme ratios.

Constructs	Locus	Base Editor-to-UGI enzyme ratio
BC22n-HiBiT / UGI-HiBiT	TRAC	1 to 18
	TRBC1	1 to 32
	TRBC2	1 to 30
	CIITA	1 to 24
	Average	1 to 26
	SD	7
BC22-2XUGI-HiBiT	TRAC	1 to 1.7
	TRBC1	1 to 2.6
	TRBC2	1 to 2.7

Constructs	Locus	Base Editor-to-UGI enzyme ratio
	CIITA	1 to 3.3
	Average	1 to 2.5
	SD	0.6

Example 38. In vivo editing with UGI in trans

[00729] *In vivo* editing profiles of deaminase containing constructs were compared to Cas9 when UGI was delivered *in trans* (as a separate mRNA). The constructs used encoded a fusion protein including D10A Cas9 with a deaminase.

[00730] Twenty-four commercially available CD-1 female mice ranging from 6-10 weeks of age (n=3 per group) were used in this study. Animals were weighed pre-dose for dosing calculations. Each RNA species was formulated separately in an LNP. Formulations containing editor mRNA, UGI mRNA and G019427 sgRNA were mixed in a w/w ratio of RNA cargos. The formulation mixture for Group 2 contained only editor mRNA and sgRNA and these were mixed in a w/w ratio of 2:1 (editor mRNA:sgRNA). Groups 3-8 contained mRNA:sgRNA at a w/w ratio of 2:1, and UGI mRNA mixed in at w/w ratios of 1:3, 1:10, 1:30, 1:100, 1:300, and 1:3000 (editor mRNA+sgRNA:UGI mRNA). Apart from the negative control group, which was dosed with TSS buffer only, all groups resulted in editing at the ANACP5 locus targeted with G019427. Formulations were administered intravenously via tail vein injection according to the doses listed in Table 99. Animals were periodically observed for adverse effects for at least 24 hours post-dose. Six days after treatment, animals were euthanized by cardiac puncture under isoflurane anesthesia; liver tissue were collected for downstream analysis. Liver punches weighing between 5 and 15 mg were collected for isolation of genomic DNA. Genomic DNA samples were analyzed with NGS sequencing as described in Example 1. Editing data are shown in Table 100 and Fig. 60. Fig. 60 shows C to T purity.

Table 99 – Dosage of RNA species for each experimental group

Group	Samples	LNP cargos	Editor mRNA (mg/kg)	sgRNA dose (mg/kg)	UGI mRNA (mg/kg)
1	TSS	None	0	0	0
2	0.0 MPK UGI	BC22n (SEQ ID NO: 2)	0.2	0.1	0.0
3	0.1 MPK UGI	BC22n (SEQ ID NO: 2)			0.1

Group	Samples	LNP cargos	Editor mRNA (mg/kg)	sgRNA dose (mg/kg)	UGI mRNA (mg/kg)
4	0.03 MPK UGI	G019427 UGI (SEQ ID NO: 26)			0.03
5	0.01 MPK UGI				0.01
6	0.003 MPK UGI				0.003
7	0.001 MPK UGI				0.001
8	0.0001 MPK UGI				0.0001

Table 100 – Mean editing in mouse livers

Group	C-to-T		C-to-A/G		Indel		C to T Purity	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	0.13	0.06	1.42	0.07	0.09	0.02	7.99	2.88
2	27.30	1.40	21.13	1.99	29.55	1.16	35.00	1.39
3	67.47	0.35	2.99	0.31	5.19	1.05	89.21	1.57
4	65.50	0.87	3.56	0.40	5.17	1.01	88.25	1.68
5	64.73	1.37	4.54	0.49	6.65	0.64	85.25	1.39
6	59.53	1.50	7.07	0.52	9.79	0.84	77.93	1.24
7	53.20	4.94	10.27	1.99	14.08	4.01	68.67	7.23
8	33.90	2.88	18.37	1.82	25.81	2.21	43.44	3.90

Example 39. Cytotoxic susceptibility of engineered T cells

[00731] Engineered T cells were assayed for cytotoxic susceptibility when targeted by natural killer (NK) cells.

[00732] NK cells (Stemcell Technologies) were thawed and resuspended at a cell concentration of 1x10^6 cells/ml into T cell growth media (TCGM) composed of OpTmizer TCGM and further supplemented with 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/mL IL-7 (Peprotech, Cat. 200-07), 5 ng/mL IL-15 (Peprotech, Cat. 200-15). Cells were incubated at 37 ˚C for 24 hours.

[00733] Twenty-four hours post thaw, the NK cells were labelled with 0.5 μM Cell Trace Violet (CTV) as follows: a vial of CTV (CellTrace™ Violet Cell Proliferation Kit, for flow cytometry, Cat. C34571) was reconstituted in DMSO from the kit to give a 5 mM stock concentration. Two μL of CTV stock was diluted with 18 μL Phosphate-Buffered Saline (PBS) (Corning, Cat. 21-040-CV) to obtain a concentration of 0.5 mM. NK cells were centrifuged at 500 x g for 5 minutes, the media was aspirated, and cells were resuspended in PBS at a concentration of 1 x 10^6 cells/mL such that the final concentration of CTV dye was 0.5 μM. The cells were mixed with CTV dye solution incubated at 37 ˚C for 20 minutes. Unbound dye was quenched by the addition of TCGM and incubated for 5 minutes. The cells

were centrifuged at 500 x g for 5 minutes. Cells are resuspended in TCGM supplemented with 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/mL IL-7 (Peprotech, Cat. 200-07), 5 ng/mL IL-15 (Peprotech, Cat. 200-15) at a concentration of 2 x 10^6 cells/mL. To test a range of effector:target (E:T) ratios, CTV-labelled NK cells were aliquoted in 100 µL of media in a 6-point, 2-fold serial dilution with the highest number of cells being 2 x 10^5 cells. Media-only samples were included as negative controls.

[00734] T cells were engineered using BC22n and UGI mRNA as described in Example 14 using G023523 (SEQ ID NO: 501) targeting HLA-A as a test sample and with G023519 (SEQ ID NO: 498) targeting B2M was as a positive control for NK killing. Unedited T cells were assayed as a negative control for NK killing. Other controls for flow cytometry included CTV-labelled NK cells without T cells; a "unstained" sample combining unlabelled NK cells and T cells; and a 1:1 mix of unlabeled heat killed and non-heat killed NK cells and T cells stained with 7AAD. T cells were resuspended at a density of 2 x 10^5 cells in TCGM composed of OpTmizer TCGM and further supplemented with 100 U/mL of recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/mL IL-7 (Peprotech, Cat. 200-07), and 5 ng/mL IL-15 (Peprotech, Cat. 200-15). Twenty thousand T cells were added to each well of NK cells and media controls. Cells were incubated at 37 ˚C for 24 hours.

[00735] At 24 hours, half of the volume of the cells from the LD heat killed well were heat killed and transferred back to the same well in the assay plate. Cells were centrifuged and resuspended in 80 µL of a 1:200 v/v solution of 7-AAD (BD Biosciences, Cat. 559925) in FACS buffer (PBS + 2% FBS (Gibco, Cat. A31605-02) + 2mM EDTA (Invitrogen, Cat. 15-575-020)). Data for specific lysis of T cells were acquired by flow cytometry using a Cytoflex LX instrument (Beckman Coulter) and analyzed using the FlowJo software package. Gates were first drawn on the CTV negative population to gate out the NK cells, followed by gating on singlets after which a gate was drawn on the 7-AAD negative population to gate for the live T cells. The percent lysis of T cells was calculated by subtracting the live cell percentage from 100. T cells edited using BC22n and HLA-A guide G023523 (SEQ ID NO: 501) were protected from NK cell mediated cytotoxicity as shown in Table 101 and Fig. 62.

Table 101 - Mean percentage lysis of engineered T cells exposed to HLA-B and C matched NK cells

E:T	Unedited			G023519 B2M			G023523 HLA-A		
	Mean	SD	n	Mean	SD	n	Mean	SD	n
10	19.65	2.33	2	69.60	4.81	2	22.23	1.10	3
5	18.80	1.59	3	61.10	0.85	2	21.35	0.49	2
2.5	22.27	6.62	3	47.95	0.49	2	22.10	1.27	2
1.25	18.47	1.27	3	39.20	2.98	3	21.00	0.81	3
0.63	19.30	0.66	3	30.20	NA	1	19.75	0.35	2
0.31	20.70	5.02	3	40.60	NA	1	20.27	1.67	3
0	19.77	2.01	3	26.57	2.73	3	18.30	1.41	3

Example 40 –Editing window for SpyCas9 and Nme2Cas9 base editors

[00736] The range of guide positions available for deamination by base editors designed with Nme2Cas9 nickase or SpyCas9 nickase and APOBEC3a was assayed by evaluating C to T conversion efficacy on a position by position basis across a panel of target sites with cytosine residues in the guide region.

Example 40.1 T cell Preparation

[00737] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and resuspended in in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. 130-070-525) and processed in a MultiMACS™ Cell 24 Separator Plus device (Miltenyi Biotec). T cells were isolated via positive selection using a Straight from Leukopak® CD4/CD8 MicroBead kit, human (Miltenyi Biotec Cat. 130-122-352). T cells were aliquoted and cryopreserved for future use in Cryostor® CS10 (StemCell Technologies Cat. 07930).

[00738] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth media (TCGM) composed of CTS OpTmizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512) 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/mL recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/ml recombinant human interleukin 15 (Peprotech, Cat. 200-15). T cells were rested in this media for 24 hours, at which time they were activated with T Cell TransAct™, human reagent (Miltenyi, Cat. 130-111-

160) added at a ratio of 1:100 by volume. T cells were activated for 48 hours prior to electroporation.

Example 40.2 T cell editing with RNA electroporation

[00739] Solutions containing mRNA encoding Spy BC22n (SEQ ID NO: 1) or Nme2 BC22n (SEQ ID NO: 315) and UGI (SEQ ID NO: 34) were prepared in P3 buffer. Guide RNAs targeting either the SCAP, LINC01588, LSP1, SEC61B, VEGFA, FancF, AAVS1, or ARHGEF9 locus was removed from the storage and denatured for 2 minutes at 95°C and incubated at room temperature for 5 minutes.

[00740] Forty-eight hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3 electroporation buffer (Lonza). For each well to be electroporated, 1 x 10^5 T cells were mixed with 200 ng of BC22n or Nme2 base editor mRNA, 200 ng of UGI mRNA and 4 μM of sgRNAs in a final volume of 20 μL of P3 electroporation buffer. This mix was transferred in duplicate to 96-well Nucleofector™ plates and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 μL of CTS Optimizer T cell growth media without cytokines for 15 minutes before being transferred to new flat-bottom 96-well plates containing an additional 80 μL of CTS OpTmizer T cell growth media supplemented with 2X cytokines. The resulting plates were incubated at 37°C for 4 days. On day 4 post-electroporation, 100 μL of cells was harvested for DNA extraction. DNA samples were subjected to PCR and subsequent NGS analysis with two separate primer sets (technical replicates), as described in Example 1.

Example 40.3 Base editor editing window analysis

[00741] Position values were assigned to each DNA base within the protospacer region and its 5' and 3' adjacent nucleotides. or SpyCas9, Position 1-20 represent the 20 bases bound by a SpyCas9 guide wherein position 1 is the PAM distal base and position 20 is the PAM proximal base. For Nme2Cas9, positions 1-24 represent the 24 bases bound by a Nme2Cas9 guide wherein position 1 is the PAM distal base and position 24 is the PAM proximal base. The 5' and 3' ends outside of protospacer region were assigned to be negative and positive positions, respectively; relative to Position 1.

[00742] **Define conversion frequency:** A guide can be designed to target the reference strand (strand +) or reverse complementary strand (strand -) of a gene. For strand + guides,

each wild-type cytosine position in the target region was recorded and conversion frequency (cytosine to thymine) at the position was calculated as ratio of sequencing reads with thymine at the position versus reads with either cytosine or thymine at the position; For strand – guide, each wild-type guanine position in target region was recorded and a conversion frequency (guanine to adenine) at the position is calculated as ratio of sequencing reads with adenine at the position versus reads with adenine and guanine at the position.

[00743] **Classify guide activity:** Four replicates were measured for each guide: two technical replicates for each of two biological replicates. Only positions in a replicate with >500 sequencing reads would be further analysed. For each guide, the mean conversion frequency of a position is the average of the replicates, and the highest mean conversion frequency of all positions for a guide is used to classify this guide's conversion activity: Low, medium, and high activity guides have a highest mean conversion frequency of <50%, 50% & <70%, >70%, respectively. High activity guides for BC22n (n=38) and Nme2 base editor (n=14) were chosen for further window analysis as shown in Table 102.

Table 102- High activity guide sequences included in editing window analysis

guide	Editor species	guide	Editor species	guide	Editor species
G021514	Nme2	G009272	Spy	G021549	Spy
G021519	Nme2	G009295	Spy	G021573	Spy
G021524	Nme2	G009303	Spy	G021574	Spy
G021529	Nme2	G009330	Spy	G021575	Spy
G021550	Nme2	G009333	Spy	G021576	Spy
G021551	Nme2	G012243	Spy	G021577	Spy
G021552	Nme2	G013549	Spy	G021578	Spy
G021553	Nme2	G013578	Spy	G021579	Spy
G021554	Nme2	G021537	Spy	G021582	Spy
G021555	Nme2	G021538	Spy	G021584	Spy
G021556	Nme2	G021540	Spy	G021585	Spy
G021560	Nme2	G021541	Spy	G021586	Spy
G021562	Nme2	G021542	Spy	G021587	Spy
G021570	Nme2	G021543	Spy	G021588	Spy
G009252	Spy	G021544	Spy	G021589	Spy
G009255	Spy	G021545	Spy	G021591	Spy
G009260	Spy	G021548	Spy	G021593	Spy
G009269	Spy				

[00744] **Define conversion rate for each position across all guides**: If a guide has the position recorded, each of its replicates with >500 sequencing reads at the position is considered as a "case". Total cases were summed for each position of all guides and served as a denominator for conversion rate calculation. In each case, if the conversion frequency is greater than 50%, it is considered as a high conversion "event", and total events is the sum of high conversion events for each position across all guides and replicates. The "events" serve as numerator for conversion rate calculation. Conversion Rate of a position equals the percentage of high conversion events in all cases for each position.

[00745] Table 103 displays the position, cases, events, and conversion rate for the spyBC22n C-to-T editing window. Fig. 63 shows conversion rate by position BC22n. Fig. 64 shows conversion rate by position for Nme2 base editor. Positions 1-11 for spyBC22n and positions 4-20 for Nme2BC22 show conversion rates over 25%.

Table 103 – Conversion rate for Spy base editor and Nme2 base editor constructs.

Editor	Pos.	Events	Cases	Conversion Rate	Editor	Pos.	Events	Cases	Conversion Rate
Nme2	-20	0	14	0.0%	Spy	-20	0	48	0.0%
Nme2	-19	0	9	0.0%	Spy	-19	0	29	0.0%
Nme2	-18	0	24	0.0%	Spy	-18	0	26	0.0%
Nme2	-17	0	24	0.0%	Spy	-17	0	32	0.0%
Nme2	-16	0	24	0.0%	Spy	-16	0	30	0.0%
Nme2	-15	0	14	0.0%	Spy	-15	0	43	0.0%
Nme2	-14	0	15	0.0%	Spy	-14	0	39	0.0%
Nme2	-13	0	9	0.0%	Spy	-13	0	54	0.0%
Nme2	-12	0	17	0.0%	Spy	-12	0	60	0.0%
Nme2	-11	0	17	0.0%	Spy	-11	0	35	0.0%
Nme2	-10	0	10	0.0%	Spy	-10	4	47	8.5%
Nme2	-9	0	24	0.0%	Spy	-9	0	38	0.0%
Nme2	-8	0	12	0.0%	Spy	-8	0	33	0.0%
Nme2	-7	0	14	0.0%	Spy	-7	0	24	0.0%
Nme2	-6	0	30	0.0%	Spy	-6	4	41	9.8%
Nme2	-5	0	19	0.0%	Spy	-5	4	37	10.8%
Nme2	-4	0	12	0.0%	Spy	-4	0	25	0.0%
Nme2	-3	0	4	0.0%	Spy	-3	0	42	0.0%
Nme2	-2	0	24	0.0%	Spy	-2	0	38	0.0%
Nme2	-1	0	18	0.0%	Spy	-1	11	56	19.6%
Nme2	1	0	14	0.0%	Spy	1	26	43	60.5%
Nme2	2	1	17	5.9%	Spy	2	31	51	60.8%
Nme2	3	3	17	17.6%	Spy	3	34	38	89.5%

Editor	Pos.	Events	Cases	Conversion Rate	Editor	Pos.	Events	Cases	Conversion Rate
Nme2	4	6	24	25.0%	Spy	4	41	41	100.0%
Nme2	5	8	14	57.1%	Spy	5	52	52	100.0%
Nme2	6	8	14	57.1%	Spy	6	39	39	100.0%
Nme2	7	8	18	44.4%	Spy	7	37	37	100.0%
Nme2	8	5	18	27.8%	Spy	8	42	42	100.0%
Nme2	9	7	28	25.0%	Spy	9	38	38	100.0%
Nme2	10	14	17	82.4%	Spy	10	33	45	73.3%
Nme2	11	15	21	71.4%	Spy	11	28	48	58.3%
Nme2	12	12	16	75.0%	Spy	12	8	57	14.0%
Nme2	13	6	10	60.0%	Spy	13	0	63	0.0%
Nme2	14	11	11	100.0%	Spy	14	1	55	1.8%
Nme2	15	21	23	91.3%	Spy	15	0	45	0.0%
Nme2	16	17	19	89.5%	Spy	16	0	50	0.0%
Nme2	17	11	16	68.8%	Spy	17	0	42	0.0%
Nme2	18	2	16	12.5%	Spy	18	0	59	0.0%
Nme2	19	2	12	16.7%	Spy	19	0	62	0.0%
Nme2	20	6	16	37.5%	Spy	20	0	54	0.0%
Nme2	21	4	20	20.0%	Spy	21	n/a	0	n/a
Nme2	22	0	19	0.0%	Spy	22	n/a	0	n/a
Nme2	23	0	16	0.0%	Spy	23	0	23	0.0%
Nme2	24	0	20	0.0%	Spy	24	0	46	0.0%
Nme2	25	n/a	0	n/a	Spy	25	0	123	0.0%
Nme2	26	n/a	0	n/a	Spy	26	0	127	0.0%
Nme2	27	0	10	0.0%	Spy	27	0	47	0.0%
Nme2	28	0	8	0.0%	Spy	28	0	38	0.0%
Nme2	29	0	45	0.0%	Spy	29	0	37	0.0%
Nme2	30	0	45	0.0%	Spy	30	0	51	0.0%
Nme2	31	0	9	0.0%	Spy	31	0	44	0.0%
Nme2	32	0	12	0.0%	Spy	32	0	42	0.0%
Nme2	33	0	22	0.0%	Spy	33	0	34	0.0%
Nme2	34	0	24	0.0%	Spy	34	0	39	0.0%
Nme2	35	0	18	0.0%	Spy	35	0	43	0.0%
Nme2	36	0	13	0.0%	Spy	36	0	37	0.0%
Nme2	37	0	13	0.0%	Spy	37	0	32	0.0%
Nme2	38	0	9	0.0%	Spy	38	0	33	0.0%
Nme2	39	0	16	0.0%	Spy	39	0	50	0.0%
Nme2	40	0	6	0.0%	Spy	40	0	33	0.0%

Example 41 – CIITA Guide RNA screening in T cells with BC22n

[00746] Different sgRNAs were screened for their potency in knocking out the CIITA gene in human T cells using C to T base editing. The percentage of T cells negative for MHC class II and/or CD74 protein expression was assayed following CIITA editing following electroporation with mRNA and different sgRNAs.

Example 41.1 T cell Preparation

[00747] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and resuspended in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. 130-070-525) and processed in a MultiMACS™ Cell 24 Separator Plus device (Miltenyi Biotec). T cells were isolated via positive selection using a Straight from Leukopak® CD4/CD8 MicroBead kit, human (Miltenyi Biotec Cat. 130-122-352). T cells were aliquoted and cryopreserved for future use in Cryostor® CS10 (StemCell Technologies Cat. 07930).

[00748] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth media (TCGM) composed of CTS OpTimizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512) 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/mL recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/mL recombinant human interleukin 15 (Peprotech, Cat. 200-15). T cells were rested in this media for 24 hours, at which time they were activated with T Cell TransAct™, human reagent (Miltenyi, Cat. 130-111-160) added at a 1:100 ratio by volume. T cells were activated for 48 hours prior to electroporation.

Example 41.2 T cell editing with RNA electroporation

[00749] Solutions containing mRNA encoding BC22n (SEQ ID NO: 1) and UGI (SEQ ID NO: 34) were prepared in P3 buffer. One hundred μM of CIITA-targeting sgRNAs were removed from their storage plates and denatured for 2 minutes at 95 °C and incubated at room temperature for 5 minutes. Forty-eight hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^6 T cells/mL in P3

electroporation buffer (Lonza). For electroporation, 1 x 10^5 T cells were mixed with 20 ng/μL of BC22n mRNAs, 20 ng/μL of UGI mRNA, and 20 pmols of sgRNA as described in Table 1 in a final volume of 20 μL of P3 electroporation buffer. This mix was transferred in duplicate to a 96-well Nucleofector™ plate and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 μL of CTS Optimizer T cell growth media without cytokines for 15 minutes before being transferred to new flat-bottom 96-well plates containing an additional 80 μL of CTS Optimizer T cell growth media supplemented with 2X cytokines. The resulting plates were incubated at 37 °C for 10 days. On day 4 post-electroporation, cells were split 1:2 in 2 U-bottom plates. One plate was collected for NGS sequencing, while the other plate was replenished with CTS Optimizer fresh media with 1X cytokines. This plate was used for flow cytometry on Day 7.

Example 41.3 Flow cytometry and NGS sequencing

[00750] On day 7 post-editing, T cells were assayed by flow cytometry to determine the surface expression of CD74 and HLA-DR, DP, DQ. Briefly, T cells were incubated for 30 minutes at 4 °C with a mixture of antibodies diluted in cell staining buffer (BioLegend, Cat. No. 420201). Antibodies against CD3 (BioLegend, Cat. No. 317336), CD4 (BioLegend, Cat. No. 317434), CD8 (BioLegend, Cat. No. 301046), and Viakrome (Beckman Coulter, Cat. No. C36628) were diluted at 1:100, and antibodies against HLA II-DR (BioLegend, Cat. No. 327018), HLA II-DP (BD Biosciences Cat No. 750872), HLA II-DQ (BioLegend, Cat. No. 561504), and CD74 (BioLegend, Cat. No. 326808) were diluted at 1:50. Cells were subsequently washed, resuspended in 100 μL of cell staining buffer and processed on a Cytoflex flow cytometer (Beckman Coulter). Flow cytometry data was analyzed using the FlowJo software package. T cells were gated based on size, shape, viability, CD8, HLA II-DP, HLA II-DQ, HLA II-DR, and CD74 expression.

Table 104. Percentage of cells negative for surface protein following genomic editing of *CIITA* with BC22n. (n=2)

Guide ID	%HLA II-DP-		%HLA II-DQ-		%HLA II-DR-		%CD74-	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
G000502	68.70	3.54	76.30	4.24	76.70	3.96	66.25	5.87
G016788	62.95	1.91	73.40	3.11	73.50	4.38	61.35	4.60
G016053	70.25	6.72	71.50	4.95	73.90	4.10	60.20	6.08
G016103	63.05	6.29	73.95	1.20	75.30	0.71	63.25	1.63

Guide ID	%HLA II-DP-		%HLA II-DQ-		%HLA II-DR-		%CD74-	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
G016114	65.80	1.98	74.70	4.38	75.85	5.30	65.25	5.87
G016117	63.70	0.57	74.60	3.25	76.00	3.11	63.45	5.30
G016034	85.55	2.19	86.30	1.13	87.95	0.07	80.30	0.14
G016035	85.35	3.75	83.55	1.06	84.05	0.35	75.25	1.20
G016039	74.95	0.35	78.20	0.71	78.50	0.28	68.90	1.27
G016040	61.90	0.85	76.30	2.26	77.80	1.56	64.10	2.55
G016041	68.60	1.84	77.30	0.85	76.75	0.35	64.20	0.00
G016043	79.95	1.48	82.55	0.35	82.50	0.71	73.80	0.00
G016044	87.84	3.17	87.80	1.56	88.65	1.34	83.75	1.91
G016045	82.25	2.90	88.70	0.99	88.35	0.78	83.40	0.99
G016047	76.85	0.21	85.40	0.28	85.05	0.21	79.20	0.00
G016050	59.05	1.91	86.80	0.71	83.40	0.57	79.00	0.99
G016052	76.85	0.21	79.25	0.78	80.55	0.92	70.85	1.34
G016054	70.30	1.70	79.85	1.20	79.30	0.42	70.35	0.35
G016055	73.35	1.34	82.15	2.19	82.15	1.77	73.10	1.56
G016056	75.80	1.41	86.05	1.20	86.35	1.34	79.05	1.91
G016057	77.90	0.71	83.95	0.35	84.45	0.21	75.30	0.99
G016058	83.65	2.19	87.25	1.06	88.20	0.99	81.20	1.41
G016060	72.55	0.78	82.70	1.84	83.05	2.47	73.55	2.76
G016061	73.15	6.29	83.10	0.57	82.55	0.21	74.50	0.42
G016063	74.60	7.35	83.75	0.64	83.50	0.71	75.45	0.35
G016064	97.98	0.17	97.80	0.18	96.83	0.13	98.58	0.04
G016065	77.80	0.28	77.70	1.98	80.00	1.56	69.35	3.18
G016068	97.73	0.26	98.07	0.81	97.59	0.66	98.55	0.66
G016071	87.05	4.31	88.55	0.49	89.45	0.49	84.80	0.85
G016074	96.88	0.19	96.34	0.04	95.85	0.57	96.56	0.23
G016075	86.05	0.92	88.20	0.85	88.50	0.14	83.50	1.27
G016076	96.69	0.64	96.67	0.01	96.38	0.02	96.34	0.18
G016077	92.27	2.50	91.20	1.22	91.40	1.44	89.39	1.68
G016078	71.20	0.28	79.55	1.48	80.65	1.48	70.40	0.85
G016079	89.31	1.57	91.24	0.55	90.24	0.00	88.40	0.14
G016081	74.35	0.07	83.05	1.06	83.00	0.42	73.20	1.41
G016082	82.30	5.09	87.50	0.71	88.80	0.14	82.25	0.07
G016083	74.95	4.88	82.90	0.57	83.95	0.21	74.65	1.06
G016085	79.90	2.97	85.40	0.71	87.45	0.07	79.45	0.49
G016086	97.71	0.33	98.06	0.44	96.63	0.08	98.89	0.18
G016087	70.25	9.12	78.55	3.61	78.30	3.39	69.30	4.24
G016088	82.25	5.16	87.40	0.57	88.30	0.42	81.75	1.06
G016089	69.00	1.27	76.65	1.34	79.00	1.70	67.35	1.06
G016091	95.35	0.95	96.81	0.24	96.64	0.25	97.15	0.28

Guide ID	%HLA II-DP-		%HLA II-DQ-		%HLA II-DR-		%CD74-	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
G016092	94.89	0.61	94.87	0.45	95.35	0.83	94.65	0.37
G016093	91.33	0.31	92.35	0.26	93.94	0.23	90.51	0.37
G016094	79.70	4.24	84.85	2.05	85.70	2.69	78.75	2.90
G016095	83.00	2.12	90.55	0.23	90.54	0.24	85.25	0.49
G016097	71.85	6.86	82.80	1.56	82.40	1.27	73.00	1.70
G016098	74.40	5.66	82.35	3.32	83.50	2.69	74.55	4.45
G016099	86.30	2.69	89.71	1.00	90.57	0.35	86.25	1.06
G016100	76.65	1.34	83.90	2.69	86.25	2.19	77.25	3.04
G016101	69.30	0.71	77.55	1.48	78.55	1.77	68.65	1.77
G016102	71.00	1.84	80.70	0.99	80.60	1.56	70.10	0.99
G016106	82.00	1.27	87.55	0.64	88.80	0.71	81.30	1.41
G016108	88.24	3.73	91.49	0.02	91.42	0.66	87.00	0.85
G016109	88.05	0.92	90.20	1.27	90.41	1.85	86.85	2.05
G016110	88.50	3.25	91.12	0.99	90.14	1.05	87.35	1.20
G016111	78.15	0.92	84.45	0.78	85.80	0.85	78.45	1.20
G016112	72.20	3.82	79.85	2.33	81.70	2.26	73.05	2.47
G016115	95.58	1.10	98.36	0.73	97.69	0.48	98.54	0.40
G016116	88.95	0.35	91.15	2.47	92.29	1.79	88.08	2.94
G016066	68.90	0.57	73.20	1.70	74.20	0.99	62.25	2.76
G016113	93.60	1.15	93.02	0.98	93.75	0.66	92.13	1.20
G016084	96.42	0.83	98.38	0.33	97.32	0.69	98.77	0.17
G016104	84.95	1.77	89.69	1.53	91.20	0.97	86.00	1.41
G016070	90.52	0.16	92.10	0.50	92.49	0.64	89.75	0.91
G016090	96.41	1.27	98.06	0.08	97.43	0.15	98.66	0.10
G016048	76.80	1.56	80.60	0.99	81.70	1.41	72.25	2.19
G016051	65.85	20.44	83.15	1.63	84.45	1.34	75.20	2.40
G016073	82.00	1.98	82.65	1.20	83.15	2.76	75.90	0.14
G016037	80.55	1.91	78.05	0.35	80.15	0.07	70.20	0.57
G016038	85.45	0.49	82.45	0.35	84.90	0.28	76.10	1.13
G016046	90.10	0.05	90.75	0.47	91.09	0.86	87.45	0.78
G016049	84.50	0.00	84.90	0.42	86.75	0.07	78.75	0.21
G016036	85.05	1.06	83.15	1.48	85.00	0.99	76.45	2.47
G016080	91.13	2.02	92.11	1.07	92.99	0.97	89.72	1.16
G016096	75.00	4.38	82.90	3.82	83.45	3.61	75.65	4.31
G016032	97.50	0.09	97.42	0.65	96.30	0.33	98.19	0.36
G016033	91.32	0.88	87.75	0.49	88.20	0.28	84.10	0.14
G016030	93.42	0.08	88.10	0.71	88.40	0.42	85.10	0.28
G016067	96.69	0.25	95.82	0.69	95.14	0.64	96.30	0.48
G016031	80.45	1.20	82.10	0.99	83.95	0.07	74.80	0.57
G016062	74.80	11.17	86.60	0.28	86.75	0.35	80.15	0.35

Guide ID	%HLA II-DP-		%HLA II-DQ-		%HLA II-DR-		%CD74-	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
G016059	77.05	3.75	83.05	0.92	83.85	0.92	75.00	1.41
G016105	66.55	5.87	74.90	3.68	76.50	3.68	65.25	3.75
G016107	71.80	4.10	80.70	0.28	80.75	0.49	71.00	0.57
G016042	95.06	0.02	95.90	0.30	94.17	0.35	94.89	1.02
G016069	52.65	3.18	76.35	0.21	75.90	0.28	62.35	0.21
G016072	46.45	4.45	73.50	0.57	71.35	0.92	56.55	0.64
G018021	76.60	0.99	84.00	0.57	84.65	0.49	76.85	0.07
G018022	60.00	3.39	72.20	1.13	73.25	1.34	60.35	2.62
G018023	64.80	2.97	74.35	0.07	75.85	0.35	63.85	2.33
G018024	64.10	1.70	74.80	1.84	74.90	1.98	63.75	4.03
G018025	63.85	6.15	73.50	6.08	74.90	6.36	63.75	6.72
G018026	72.15	4.17	81.35	0.92	82.55	1.48	73.00	0.71
G018027	82.40	0.99	81.45	1.91	82.40	2.26	72.35	3.32
G018028	78.25	0.49	78.40	0.00	80.05	0.49	67.90	2.40
G018029	75.60	3.11	74.80	0.14	76.95	0.07	64.15	2.05
G018030	76.70	1.56	76.35	3.46	77.80	3.54	66.65	5.44
G018031	81.40	3.96	84.00	1.13	82.40	0.42	75.55	0.49
G018032	77.00	2.97	79.70	0.14	79.30	0.42	70.35	0.92
G018033	81.15	3.04	82.05	0.35	81.60	0.57	73.20	1.27
G018034	96.87	0.43	93.49	0.88	92.67	0.99	92.96	1.63
G018035	94.39	0.64	88.20	1.13	88.64	2.03	83.70	2.97
G018036	73.80	0.28	76.45	2.19	77.55	2.90	65.30	3.82
G018037	71.80	0.14	72.25	5.16	73.85	5.73	61.40	7.64
G018038	70.25	7.85	74.95	1.06	75.65	2.05	63.65	2.62
G018039	94.72	0.49	91.67	0.05	91.17	0.51	89.20	0.99
G018040	77.30	1.13	81.70	0.42	81.60	0.57	72.35	0.64
G018041	75.10	1.70	80.15	0.35	80.15	0.21	70.50	0.85
G018042	89.35	0.49	86.60	0.14	87.90	0.14	80.50	0.00
G018043	85.50	1.13	84.00	1.56	84.95	0.35	77.55	0.64
G018044	88.20	0.99	89.01	1.57	89.25	0.78	83.95	2.19
G018045	76.15	1.34	82.35	2.33	81.60	2.26	72.50	3.25
G018046	87.20	0.42	88.65	0.21	89.00	0.57	82.80	1.27
G018047	79.95	1.63	84.60	2.12	84.25	2.76	76.25	3.61
G018048	88.50	2.12	88.00	1.13	89.50	0.42	81.60	0.14
G018049	81.40	3.25	81.40	4.24	82.60	5.23	74.00	5.66
G018050	77.75	0.49	78.60	0.85	80.85	0.07	70.25	0.35
G018051	81.20	1.56	81.95	1.48	84.20	1.13	73.55	3.04
G018052	76.30	1.56	81.70	1.27	81.10	1.56	70.90	2.69
G018053	77.25	1.06	81.40	0.85	82.35	1.06	71.55	1.63
G018054	90.23	0.14	89.00	0.99	89.25	0.64	84.50	0.28

Guide ID	%HLA II-DP-		%HLA II-DQ-		%HLA II-DR-		%CD74-	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
G018055	73.80	3.39	79.20	0.99	79.85	1.06	67.80	1.41
G018056	73.40	2.97	83.15	0.35	83.15	0.78	73.15	0.78
G018057	73.95	1.63	81.40	2.97	81.45	1.91	71.30	3.11
G018058	73.40	3.25	80.90	0.99	81.05	0.07	69.15	0.35
G018059	75.15	4.17	81.70	0.14	81.55	0.07	71.10	1.13
G018060	75.90	2.69	81.30	0.28	81.55	0.78	72.05	1.91
G018061	71.65	2.05	81.60	1.41	80.65	1.06	72.00	1.41
G018062	70.75	2.33	75.05	0.64	76.30	1.56	64.65	2.33
G018063	78.45	0.92	80.95	0.92	82.20	0.85	72.50	0.14
G018064	76.15	0.78	82.50	0.00	82.90	0.28	73.90	0.28
G018065	77.65	0.21	82.60	0.42	83.85	0.35	74.05	0.07
G018066	72.60	0.00	82.30	0.28	83.05	0.64	73.60	0.14
G018067	97.57	0.81	98.24	0.19	97.73	0.28	98.83	0.04
G018068	89.05	0.64	90.79	0.52	90.81	1.02	87.20	1.27
G018069	75.90	2.55	81.55	0.92	82.15	0.35	71.85	1.20
G018070	75.30	1.27	82.05	1.20	82.15	0.64	71.20	0.42
G018071	77.40	3.39	83.35	1.77	84.75	1.91	74.65	0.92
G018072	72.05	1.63	81.45	0.78	81.95	1.91	71.95	2.62
G018073	72.35	0.07	79.30	1.13	79.25	2.19	69.40	2.26
G018074	67.25	2.05	75.25	0.49	76.40	0.57	64.35	0.35
G018075	96.86	1.46	97.33	0.14	96.08	0.31	98.35	0.04
G018076	94.26	4.87	97.96	0.21	97.57	0.15	98.55	0.06
G018077	74.75	2.76	81.65	0.35	82.25	0.21	72.25	1.34
G018078	94.30	0.48	94.87	0.31	95.03	0.06	93.43	0.24
G018079	86.00	1.56	89.55	0.35	89.80	0.28	84.25	0.78
G018080	80.50	3.96	85.50	1.70	85.45	1.63	78.35	0.78
G018081	79.85	1.06	86.05	1.20	86.00	0.85	78.70	0.57
G018082	80.10	0.14	83.10	0.42	84.05	1.20	75.65	1.91
G018083	86.75	1.34	89.50	0.28	89.35	0.21	84.55	0.21
G018084	84.60	3.25	87.35	1.34	87.95	0.92	82.00	1.27
G018085	86.95	2.76	90.35	2.05	90.03	1.03	86.30	2.26
G018086	68.15	1.34	78.60	0.42	75.95	0.07	66.25	0.92
G018087	61.40	2.26	74.35	2.33	72.65	2.62	61.85	1.63
G018088	88.80	0.85	91.11	0.28	91.88	0.11	87.60	1.41
G018089	81.75	4.17	89.00	1.83	89.28	1.53	83.75	1.77
G018090	81.65	1.20	90.11	0.29	89.98	0.25	85.35	1.34
G018091	97.02	0.49	98.63	0.03	98.15	0.12	98.98	0.08
G018092	71.40	5.23	81.90	0.71	82.40	0.85	73.45	0.92
G018093	94.15	1.20	96.36	0.21	96.36	0.02	95.70	0.18
G018094	77.50	2.12	84.55	1.63	85.25	1.48	76.50	0.14

Guide ID	%HLA II-DP-		%HLA II-DQ-		%HLA II-DR-		%CD74-	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
G018095	69.50	0.00	83.05	0.07	82.70	0.14	73.90	0.99
G018096	69.25	0.21	83.50	0.99	83.45	1.06	75.25	2.05
G018097	68.55	4.74	81.60	2.83	82.20	1.84	73.25	3.75
G018098	66.65	0.21	78.85	0.21	79.75	0.21	68.75	1.20
G018099	65.05	3.18	73.70	4.95	74.55	4.31	62.40	4.53
G018100	95.51	0.40	95.99	0.19	95.62	0.92	96.21	0.18
G018101	93.00	0.33	93.70	0.01	94.03	0.25	93.11	0.28
G018102	96.55	0.33	96.31	0.64	96.34	0.08	96.58	0.08
G018103	96.92	1.46	97.58	0.02	97.53	0.08	98.36	0.40
G018104	72.00	0.99	82.45	2.19	82.70	1.84	73.40	2.12
G018105	71.60	1.98	82.20	0.71	81.80	0.28	73.30	0.00
G018106	92.49	1.33	98.62	0.06	97.03	0.43	98.96	0.04
G018107	96.55	1.25	98.14	0.56	97.53	0.70	98.98	0.08
G018108	85.50	0.42	88.00	0.71	89.35	0.92	83.40	1.56
G018109	84.10	1.56	84.55	2.19	86.90	1.27	79.75	3.32
G018110	67.65	3.04	75.15	0.49	77.50	0.99	65.35	1.48
G018111	96.49	0.35	96.85	1.05	95.24	1.32	97.03	0.52
G018112	68.60	5.09	76.65	4.88	77.30	4.53	66.60	4.38
G018113	72.75	3.89	79.60	3.54	80.15	2.76	70.30	3.96
G018114	70.80	2.26	80.05	2.47	81.25	1.77	71.65	2.62
G018115	62.40	1.41	77.40	0.57	76.90	1.27	67.00	1.56
G018116	96.05	0.89	98.01	0.39	96.99	0.42	98.32	0.21
G018117	96.93	1.27	97.63	0.52	97.19	0.87	98.96	0.25
G018118	97.25	0.51	97.71	0.25	96.94	0.44	98.91	0.26
G018119	94.78	1.02	96.29	1.11	95.57	1.05	97.00	1.20
G018120	96.04	1.87	97.61	0.06	97.42	0.56	98.94	0.07
G018121	95.26	0.62	97.12	0.41	96.44	0.23	97.85	0.45

[00751] On day 4 post-editing, DNA samples were subjected to PCR and subsequent NGS analysis, as described in Example 1. Table 105 shows CIITA editing outcomes in T cells edited with BC22n.

Table 105– Mean percent editing at CIITA locus with BC22n. (n=2)

Guide ID	%C to T		%C to A/G		% Indels	
	Mean	SD	Mean	SD	Mean	SD
G000502	84.63	2.39	0.82	0.19	1.26	0.37
G016788	95.63	0.06	0.99	0.39	0.72	0.15
G016053	94.77	1.07	1.66	0.43	0.71	0.02
G016103	95.60	0.97	0.63	0.23	0.75	0.29

Guide ID	%C to T		%C to A/G		% Indels	
	Mean	SD	Mean	SD	Mean	SD
G016114	96.42	0.23	0.66	0.59	0.65	0.09
G016117	96.23	0.70	0.82	0.28	0.90	0.17
G016034	93.11	1.07	0.61	0.21	1.49	0.52
G016035	95.41	0.59	0.52	0.31	0.62	0.24
G016039	88.20	1.33	1.51	0.42	7.99	1.32
G016043	88.51	2.91	1.27	0.11	2.23	0.46
G016044	0.13	0.03	90.04	0.38	9.83	0.41
G016045	61.66	0.51	1.73	0.90	7.12	0.22
G016047	92.74	0.48	0.80	0.06	3.00	0.13
G016050	90.63	0.04	1.35	0.38	1.87	0.56
G016052	92.68	1.74	0.83	0.44	3.26	1.66
G016054	90.82	2.94	1.03	0.89	4.38	4.28
G016055	83.61	1.98	0.46	0.21	10.63	1.04
G016056	96.69	0.65	0.30	0.03	0.64	0.35
G016057	93.83	0.35	0.98	0.18	0.99	0.47
G016058	95.49	0.06	0.35	0.14	0.41	0.31
G016060	85.95	0.54	0.49	0.16	7.63	1.63
G016061	92.81	2.51	0.59	0.34	0.75	0.19
G016063	70.60	1.87	0.31	0.15	0.25	0.07
G016064	94.14	2.28	0.81	0.41	0.97	0.33
G016065	57.50	1.68	0.61	0.37	1.15	0.32
G016068	94.26	0.97	0.34	0.16	0.52	0.21
G016071	93.78	0.73	0.84	0.14	2.17	0.40
G016074	93.88	1.20	0.92	0.19	2.41	0.28
G016075	91.72	0.73	0.58	0.15	2.89	0.52
G016076	91.24	1.10	0.28	0.20	3.02	1.02
G016077	94.94	1.07	1.08	0.53	1.50	0.46
G016078	93.52	1.78	0.41	0.15	3.30	1.54
G016079	96.29	0.05	0.51	0.08	0.81	0.13
G016081	75.32	7.28	1.33	1.44	7.61	7.04
G016082	31.35	5.20	0.07	0.14	34.39	6.24
G016083	0.24	0.08	99.63	0.11	0.08	0.06
G016085	80.16	3.07	0.74	0.37	0.36	0.12
G016086	96.02	1.68	1.45	0.46	0.84	0.38
G016087	90.30	1.43	0.32	0.20	5.11	0.54
G016088	92.14	0.51	1.05	0.33	2.24	0.86
G016089	94.49	0.39	0.61	0.30	1.14	0.30
G016091	95.93	0.99	1.03	0.23	0.37	0.09
G016092	95.62	0.79	1.17	0.40	0.71	0.28
G016093	95.74	0.81	0.94	0.43	0.43	0.18

Guide ID	%C to T		%C to A/G		% Indels	
	Mean	SD	Mean	SD	Mean	SD
G016094	95.85	1.03	0.58	0.39	0.99	0.48
G016095	94.77	0.52	1.32	0.31	0.72	0.37
G016097	94.90	1.64	0.55	0.19	1.40	0.51
G016098	91.71	1.11	0.48	0.18	0.50	0.36
G016099	93.42	1.55	0.56	0.05	3.64	1.06
G016100	96.54	0.93	0.46	0.22	0.46	0.36
G016101	41.21	1.15	0.77	0.07	0.22	0.04
G016102	96.06	1.23	0.46	0.05	0.69	0.25
G016106	90.41	1.16	3.33	0.54	2.66	0.49
G016108	76.19	0.74	0.42	0.28	0.69	0.38
G016109	94.93	0.35	0.46	0.21	1.86	0.59
G016110	87.64	1.01	0.55	0.20	7.63	0.72
G016111	92.93	1.09	1.17	0.58	2.11	1.08
G016112	1.56	0.46	1.79	0.90	0.03	0.05
G016115	89.86	1.01	0.67	0.29	7.30	0.50
G016116	91.37	0.33	0.54	0.27	0.59	0.09
G016066	1.02	0.16	0.32	0.07	0.40	0.21
G016113	93.23	0.98	1.10	0.16	2.40	0.57
G016084	74.10	1.35	0.87	0.16	21.41	0.94
G016104	0.00	0.00	0.00	0.00	0.00	0.00
G016070	94.80	0.30	0.54	0.04	1.66	0.17
G016090	84.09	0.00	0.51	0.00	12.29	0.00
G016048	48.81	1.82	0.78	0.07	0.54	0.18
G016051	95.69	0.45	1.06	0.09	0.91	0.25
G016073	94.11	0.83	0.76	0.18	1.37	0.29
G016037	64.35	3.31	1.61	0.81	3.46	4.94
G016038	94.99	2.01	1.80	0.64	1.25	0.88
G016046	82.29	2.89	0.88	0.54	14.28	2.86
G016049	95.33	0.15	1.41	0.87	0.46	0.66
G016036	71.71	3.82	0.63	0.14	1.94	0.40
G016080	81.09	1.58	1.33	0.47	5.64	0.94
G016096	94.79	0.33	0.45	0.10	1.81	0.37
G016032	90.27	1.53	2.25	2.37	2.97	0.38
G016033	83.90	0.89	1.79	1.63	5.03	0.94
G016030	96.66	1.35	0.65	0.39	0.31	0.14
G016067	95.79	0.26	0.61	0.17	1.37	0.26
G016031	94.51	0.24	0.75	0.07	2.35	0.68
G016062	93.37	0.45	1.17	0.13	1.91	0.33
G016059	90.26	0.93	0.80	0.30	5.62	0.33
G016105	95.81	1.17	0.53	0.24	0.47	0.45

Guide ID	%C to T		%C to A/G		% Indels	
	Mean	SD	Mean	SD	Mean	SD
G016107	91.94	2.74	0.66	0.36	3.18	2.72
G016042	90.07	0.00	0.97	0.00	2.51	0.00
G016069	93.31	0.96	0.96	0.31	1.03	0.33
G016072	84.94	7.39	0.39	0.08	11.91	8.00
G018021	91.14	2.38	2.45	0.52	3.22	2.60
G018022	93.68	3.27	1.83	0.33	1.46	0.10
G018023	92.37	1.73	2.25	0.07	2.31	0.52
G018024	96.06	1.54	0.41	0.37	0.97	0.07
G018025	96.88	0.78	0.63	0.12	0.43	0.06
G018026	89.15	1.37	1.66	0.04	7.20	1.31
G018027	88.89	3.13	1.02	0.04	5.37	2.62
G018028	91.28	2.09	0.59	0.24	2.96	2.02
G018029	22.86	0.72	0.43	0.07	0.17	0.02
G018030	79.89	2.66	0.36	0.51	0.18	0.25
G018031	90.43	1.35	0.47	0.38	1.52	0.07
G018032	64.78	2.64	1.39	0.69	1.06	0.50
G018033	88.94	0.84	1.98	0.29	1.97	0.44
G018034	89.00	1.56	1.09	0.11	6.23	0.88
G018035	95.10	1.59	0.50	0.25	0.66	0.17
G018036	83.69	0.01	1.91	0.30	0.53	0.20
G018037	94.50	0.54	1.64	0.03	1.18	0.23
G018038	90.91	0.00	1.11	0.00	3.74	0.00
G018039	94.88	0.00	0.60	0.00	0.41	0.00
G018040	44.77	1.52	0.33	0.14	0.50	0.03
G018041	66.56	2.05	0.60	0.30	0.27	0.20
G018042	90.79	1.10	1.28	0.47	1.64	0.35
G018043	94.94	0.55	0.57	0.03	0.40	0.06
G018044	91.71	1.61	0.89	0.40	1.25	0.20
G018045	90.26	1.59	1.46	0.54	3.91	1.89
G018046	94.59	1.33	0.76	0.05	2.03	0.64
G018047	93.84	1.32	0.46	0.08	2.03	0.64
G018048	95.20	2.58	0.54	0.52	0.62	0.30
G018049	89.68	2.09	0.17	0.23	0.45	0.30
G018050	95.74	0.65	0.31	0.22	0.84	0.39
G018051	92.47	0.90	0.54	0.76	0.73	0.31
G018052	93.14	0.85	1.39	0.91	0.85	0.78
G018053	78.88	3.77	0.74	0.13	0.52	0.34
G018054	92.69	1.11	0.79	0.10	2.55	0.46
G018055	89.68	1.52	0.48	0.57	1.05	0.36
G018056	94.00	1.10	0.45	0.14	1.85	0.99

Guide ID	%C to T		%C to A/G		% Indels	
	Mean	SD	Mean	SD	Mean	SD
G018057	89.87	0.49	0.72	0.04	1.31	0.31
G018058	91.04	0.77	0.59	0.12	1.46	0.28
G018059	49.64	3.56	0.46	0.19	1.56	0.43
G018060	97.65	0.65	1.22	0.17	1.13	0.47
G018061	92.52	0.29	0.89	0.33	2.76	0.52
G018062	88.19	0.72	0.84	0.18	6.65	0.41
G018063	92.94	0.63	1.54	0.20	0.94	0.38
G018064	94.14	1.12	1.38	0.31	0.97	0.09
G018065	93.20	1.50	1.32	0.57	0.91	0.23
G018066	91.62	0.45	1.69	0.06	0.45	0.17
G018067	92.72	1.43	3.06	0.48	0.42	0.15
G018068	93.33	0.49	0.62	0.08	0.27	0.06
G018069	93.83	0.77	0.94	0.33	0.99	0.07
G018070	94.61	0.57	0.59	0.08	1.33	0.35
G018071	87.61	0.46	0.26	0.22	1.10	0.29
G018072	92.86	0.84	1.76	0.04	0.34	0.06
G018073	87.03	0.52	1.48	0.33	0.40	0.15
G018074	80.06	11.41	0.42	0.13	0.46	0.24
G018075	92.85	0.27	0.59	0.37	0.61	0.08
G018076	95.88	1.59	0.56	0.38	0.61	0.30
G018077	0.23	0.12	0.55	0.26	0.13	0.08
G018078	94.97	1.42	0.48	0.16	0.44	0.22
G018079	89.45	1.79	0.38	0.26	0.81	0.47
G018080	80.86	3.14	0.22	0.15	0.58	0.35
G018081	93.61	0.63	0.17	0.14	0.91	0.44
G018082	93.15	1.21	0.16	0.07	1.80	0.50
G018083	78.69	0.88	0.29	0.13	2.73	0.50
G018084	81.88	0.70	0.43	0.19	1.39	0.60
G018085	93.17	1.26	0.44	0.16	2.62	0.43
G018086	95.03	1.07	0.61	0.20	1.39	0.44
G018087	94.37	1.49	1.07	0.79	1.41	0.21
G018088	91.62	1.31	1.01	0.43	1.91	0.99
G018089	94.41	1.17	0.75	0.41	3.04	1.52
G018090	93.66	0.82	0.54	0.28	3.12	0.96
G018091	96.82	0.00	0.60	0.00	0.77	0.00
G018092	90.55	1.58	0.74	0.27	2.02	0.34
G018093	94.46	0.49	1.13	0.13	0.96	0.28
G018094	82.55	5.17	0.55	0.11	12.35	6.02
G018095	94.60	0.85	1.94	0.68	0.75	0.21
G018096	54.68	0.62	1.99	0.16	0.38	0.12

Guide ID	%C to T		%C to A/G		% Indels	
	Mean	SD	Mean	SD	Mean	SD
G018097	91.89	2.73	0.96	0.89	1.01	0.80
G018098	70.42	0.00	0.35	0.00	0.00	0.00
G018099	96.84	0.16	0.43	0.10	0.64	0.16
G018100	93.24	1.08	0.77	0.20	3.07	0.06
G018101	92.34	0.51	2.16	0.73	3.06	0.06
G018102	91.87	0.99	0.42	0.34	0.61	0.16
G018103	94.50	1.22	0.59	0.22	1.25	0.53
G018104	94.14	0.76	0.31	0.14	1.09	0.65
G018105	95.51	0.98	0.42	0.14	0.82	0.21
G018106	94.77	0.90	1.12	0.18	2.15	0.28
G018107	94.66	0.57	0.80	0.03	2.09	0.04
G018108	94.77	1.14	1.23	0.70	1.55	0.96
G018109	94.40	1.37	1.06	0.91	1.80	0.67
G018110	92.19	1.30	1.86	0.34	4.04	0.77
G018111	91.96	3.12	0.72	0.17	5.54	3.11
G018112	91.54	1.34	0.80	0.15	5.69	0.44
G018113	96.42	0.45	0.43	0.22	1.40	0.35
G018114	95.95	0.89	0.79	0.32	0.27	0.20
G018115	92.89	1.00	0.59	0.16	1.10	0.40
G018116	91.94	0.57	0.99	0.54	0.61	0.24
G018117	95.48	1.10	0.62	0.31	0.55	0.31
G018118	96.20	0.38	0.43	0.31	0.46	0.17
G018119	83.25	1.31	0.43	0.08	0.44	0.25
G018120	97.00	0.72	0.42	0.26	0.38	0.20
G018121	95.42	0.58	0.50	0.13	0.42	0.15

Example 42 –Screening CIITA sgRNAs in dose-response with BC22n in T cells

[00752] Highly efficient CIITA sgRNAs identified in Example 41 were further assayed for base editing efficacy at multiple guide concentrations in T cells. The potency of each was assayed for genome editing efficacy by NGS or by disruption of surface protein expression of HLA-DR, DP, DQ by flow cytometry.

Example 42.1 T cell Preparation

[00753] Healthy human donor apheresis was obtained commercially (Hemacare), and cells were washed and resuspended in in CliniMACS® PBS/EDTA buffer (Miltenyi Biotec Cat. 130-070-525) and processed in a MultiMACS™ Cell 24 Separator

Plus device (Miltenyi Biotec). T cells were isolated via positive selection using a Straight from Leukopak® CD4/CD8 MicroBead kit, human (Miltenyi Biotec Cat. 130-122-352). T cells were aliquoted and cryopreserved for future use in Cryostor® CS10 (StemCell Technologies Cat. 07930).

[00754] Upon thaw, T cells were plated at a density of 1.0 x 10^6 cells/mL in T cell growth media (TCGM) composed of CTS OpTmizer T Cell Expansion SFM and T Cell Expansion Supplement (ThermoFisher Cat. A1048501), 5% human AB serum (GeminiBio, Cat. 100-512), 1X Penicillin-Streptomycin, 1X Glutamax, 10 mM HEPES, 200 U/mL recombinant human interleukin-2 (Peprotech, Cat. 200-02), 5 ng/mL recombinant human interleukin 7 (Peprotech, Cat. 200-07), and 5 ng/mL recombinant human interleukin 15 (Peprotech, Cat. 200-15). T cells were rested in this media for 24 hours, at which time they were activated with T Cell TransAct™ human reagent (Miltenyi, Cat. 130-111-160) added at a 1:100 ratio by volume. T cells were activated for 48 hours prior to electroporation.

Example 42.2 T cell editing with RNA electroporation

[00755] Solutions containing mRNAs encoding BC22n (SEQ ID NO: 1) and UGI (SEQ ID NO: 34) were prepared in P3 buffer. 100 μM CIITA targeting sgRNAs were removed from their storage plates and denatured for 2 minutes at 95 °C and incubated at room temperature for 5 minutes. Forty-eight hours post activation, T cells were harvested, centrifuged, and resuspended at a concentration of 12.5 x 10^{6} T cells/mL in P3 electroporation buffer (Lonza). Each sgRNA was serially diluted in ratio of 1:2 in P3 electroporation buffer starting from 60 pmols in a 96-well PCR plate in duplicate. Following dilution, 1 x 10^{5} T cells, 20 ng/μL of BC22n mRNAs, and 20 ng/μL of UGI mRNA were mixed with sgRNA plate to make the final volume of 20 μL of P3 electroporation buffer. This mix was transferred to 4 corresponding 96-well Nucleofector™ plates and electroporated using the manufacturer's pulse code. Electroporated T cells were immediately rested in 80 μL of CTS Optimizer T cell growth media without cytokines for 15 minutes before being transferred to new flat-bottom 96-well plates containing an additional 80 μL of CTS OpTmizer T cell growth media supplemented with 2X cytokines. The resulting plates were incubated at 37 °C for 7 days. On day 4 post-electroporation, cells were

[00756] split 1:2 in two U-bottom plates, and one plate was collected for NGS sequencing, while the other plate was replenished with CTS Optimizer fresh media with 1X cytokines. This plate was used for flow cytometry on Day 7.

Example 42.3 Flow cytometry and NGS sequencing

[00757] On day 7 post-editing, T cells were assayed by flow cytometry to determine surface expression of HLA-DR, DP, DQ. Briefly, T cells were incubated for 30 minutes at 4 °C with a mixture of antibodies diluted in cell staining buffer (BioLegend, Cat. No. 420201). Antibodies against CD3 (BioLegend, Cat. No. 317336), CD4 (BioLegend, Cat. No. 317434), CD8 (BioLegend, Cat. No. 301046), and Viakrome (Beckman Coulter, Cat. No. C36628) were diluted at 1:100, and antibodies against HLA II-DR, DP, DQ (BioLegend, Cat. No. 361714) were diluted at 1:50. Cells were subsequently washed, resuspended in 100 µL of cell staining buffer and processed on a Cytoflex flow cytometer (Beckman Coulter). Flow cytometry data was analyzed using the FlowJo software package. T cells were gated based on size, shape, viability, CD8, and HLA-DR, DP, DQ.

[00758] Table 106 shows CIITA editing outcomes and the percentage of T cells negative for HLA-DR, DP, DQ in T cells following base editing with BC22n.

Table 106 – Percent editing and percent of HLA II-DP, DQ, DR negative cells following CIITA editing with BC22n base editor

		C>T			% HLA II-DR, DP, DQ			% CD74		
	sgRNA (pmols)	Ave	SD	N	Ave	SD	N	Ave	SD	N
G016064	60	96.20%	0.69%	2	99.65	0.07	2	96.2	0.14	2
	30	94.85%	0.69%	2	99.15	0.35	2	92.65	0.21	2
	15	90.26%	0.03%	2	96.70	0.28	2	90.15	2.19	2
	7.5	72.95%	1.76%	2	87.45	1.20	2	80.95	3.32	2
	3.75	49.55%	2.09%	2	73.05	0.64	2	66.35	4.31	2
	1.88	28.63%	1.34%	2	59.40	1.70	2	55.2	5.8	2
	0.94	14.98%	0.01%	2	54.60	2.40	2	46.9	1.84	2
	0	0.26%	0.02%	2	46.70	5.52	2	45.15	0.49	2
G016068	60	97.40%	0.87%	2	99.70	0.00	2	94.25	1.2	2
	30	90.31%	0.83%	2	93.15	0.07	2	85.15	1.2	2
	15	75.08%	0.17%	2	82.45	0.21	2	74.65	1.48	2
	7.5	49.64%	2.08%	2	68.65	1.91	2	57	1.27	2
	3.75	28.95%	1.51%	2	57.85	1.63	2	43.9	1.27	2
	1.88	14.90%	0.58%	2	51.10	2.26	2	35.85	3.75	2
	0.94	7.11%	0.22%	2	49.50	4.67	2	38.2	4.1	2
	0	0.21%	0.02%	2	51.50	1.13	2	44.8	1.56	2
G016074	60	96.62%		2	96.95	0.21	2	90.6	0.28	2
	30	94.89%	0.58%	2	93.50	1.27	2	85.45	1.06	2
	15	89.79%	3.18%	2	88.50	0.28	2	80.35	0.64	2
	7.5	72.00%	1.61%	2	77.75	0.35	2	68.35	1.77	2

	3.75	47.99%	0.39%	2	65.30	1.84	2	53.05	1.91	2
	1.88	27.92%	0.44%	2	57.60	0.14	2	45	0.42	2
	0.94	13.11%	1.79%	2	54.30	0.71	2	37.9	0	2
	0	0.41%	0.13%	2	49.15	2.05	2	44.05	1.91	2
G016076	60	96.03%	0.07%	2	94.30	0.14	2	90	0.71	2
	30	93.29%	ND	2	89.10	0.99	2	81.65	1.48	2
	15	80.74%	0.41%	2	83.55	0.35	2	74.1	0.85	2
	7.5	51.33%	1.82%	2	69.05	1.91	2	57.7	2.97	2
	3.75	28.40%	2.41%	2	58.90	2.26	2	43.15	0.64	2
	1.88	15.12%	1.91%	2	56.05	0.78	2	38.45	1.77	2
	0.94	7.23%	0.26%	2	52.80	0.71	2	38.2	0.28	2
	0	0.27%	0.00%	2	47.70	1.70	2	43.5	3.68	2
G016086	60	97.81%	0.16%	2	99.25	0.35	2	94.35	0.92	2
	30	ND	ND	2	98.65	0.07	2	90.25	1.2	2
	15	ND	ND	2	98.20	0.28	2	88.8	1.98	2
	7.5	97.19%	1.91%	2	95.85	0.07	2	87.5	0.14	2
	3.75	ND	ND	2	87.10	0.28	2	74.95	2.47	2
	1.88	ND	ND	2	72.45	0.07	2	60.75	2.76	2
	0.94	ND	ND	2	59.10	0.85	2	48.05	3.89	2
	0	ND	ND	2	50.65	2.76	2	46.15	2.19	2
G016091	60	ND	ND	2	98.05	0.07	2	89.85	0.21	2
	30	ND	ND	2	96.95	0.07	2	87.9	0.85	2
	15	ND	ND	2	91.40	1.84	2	83.65	2.76	2
	7.5	ND	ND	2	82.00	1.27	2	74	3.11	2
	3.75	ND	ND	2	69.70	0.99	2	62.75	2.33	2

	1.88	ND	ND	2	59.40	1.41	2	54.4	3.54	2
	0.94	ND	ND	2	54.25	2.05	2	46.3	5.09	2
	0	ND	ND	2	46.30	2.26	2	48.3	4.67	2
G016115	60	93.33%	0.70%	2	96.80	0.14	2	92.55	0.07	2
	30	94.46%	0.23%	2	92.70	1.41	2	87.8	1.98	2
	15	93.51%	0.70%	2	89.85	0.35	2	82.75	3.75	2
	7.5	90.13%	0.37%	2	83.35	0.35	2	78.35	2.47	2
	3.75	75.18%	1.51%	2	71.40	1.27	2	65.5	2.4	2
	1.88	51.65%	1.64%	2	64.20	4.95	2	56.1	1.84	2
	0.94	30.19%	2.14%	2	55.25	3.04	2	49.15	4.88	2
	0	0.26%	0.03%	2	46.85	7.57	2	47.1	3.25	2
G016084	60	73.43%	0.13%	2	97.00	0.14	2	91.35	1.06	2
	30	73.70%	0.38%	2	94.80	0.42	2	84.9	0.42	2
	15	67.30%	1.19%	2	91.60	0.71	2	80.6	0.57	2
	7.5	48.91%	0.27%	2	80.20	0.00	2	67.5	0.42	2
	3.75	28.24%	0.35%	2	67.50	0.57	2	51.85	3.18	2
	1.88	16.89%	0.30%	2	59.05	2.76	2	42.7	0.14	2
	0.94	7.42%	0.06%	2	54.65	0.92	2	40.05	0.78	2
	0	1.06%	0.06%	2	49.25	2.47	2	46.9	1.98	2
G016090	60	87.05%	2.52%	2	98.70	1.27	2	93.25	2.19	2
	30	88.93%	0.06%	2	98.90	0.14	2	91	0.99	2
	15	89.51%	0.24%	2	98.70	0.28	2	90.45	0.21	2
	7.5	83.78%	0.02%	2	93.80	0.57	2	85.05	0.07	2
	3.75	66.23%	0.40%	2	83.35	1.06	2	72.35	0.35	2
	1.88	44.49%	0.27%	2	69.50	0.57	2	57.75	1.34	2

	0.94	22.90%	1.49%	2	57.25	1.77	2	45.15	2.47	2
	0	0.61%	0.14%	2	49.55	1.63	2	47.5	4.1	2
G016032	60	91.66%	0.76%	2	99.35	0.07	2	93.65	2.05	2
	30	86.77%	2.13%	2	94.25	1.06	2	87.45	1.63	2
	15	75.23%	1.22%	2	86.10	2.26	2	74.65	0.78	2
	7.5	54.08%	0.63%	2	74.10	0.85	2	60.65	1.34	2
	3.75	33.34%	2.39%	2	63.20	3.82	2	45.95	3.75	2
	1.88	16.30%	1.21%	2	57.20	0.57	2	39.7	0.28	2
	0.94	7.33%	0.38%	2	53.10	0.85	2	37.6	1.13	2
	0	0.33%	0.00%	2	52.65	4.31	2	47.75	2.62	2
G016067	60	97.37%	0.56%	2	93.10	0.14	2	85.95	2.19	2
	30	97.47%	0.36%	2	93.00	0.28	2	81.8	3.54	2
	15	95.78%	0.13%	2	89.30	1.41	2	79.45	0.49	2
	7.5	86.53%	0.24%	2	81.05	2.19	2	71	2.26	2
	3.75	68.12%	1.27%	2	69.35	1.20	2	58.3	2.4	2
	1.88	43.82%	0.48%	2	57.70	1.98	2	48.2	0.99	2
	0.94	22.80%	1.26%	2	54.95	0.35	2	42.55	1.34	2
	0	0.14%	0.06%	2	49.55	1.63	2	43.7	1.98	2
G018067	60	95.58%	0.35%	2	99.40	0.00	2	91.45	1.34	2
	30	94.13%	0.08%	2	98.90	0.42	2	91.75	1.63	2
	15	91.49%	0.05%	2	95.95	0.21	2	89.45	0.07	2
	7.5	77.17%	1.50%	2	83.95	0.92	2	76.6	0.57	2
	3.75	56.15%	1.07%	2	71.00	0.71	2	63.65	4.03	2
	1.88	32.76%	0.32%	2	57.65	3.46	2	50.6	3.39	2
	0.94	17.47%	0.95%	2	50.85	3.04	2	46.1	3.82	2

	0	0.41%	0.01%	2	49.45	5.16	2	47.85	1.91	2
G018075	60	96.12%	1.35%	2	98.25	0.07	2	92	1.27	2
	30	94.07%	3.53%	2	97.45	1.06	2	88.65	2.76	2
	15	92.08%	0.90%	2	97.10	0.14	2	87.25	2.05	2
	7.5	83.76%	3.28%	2	90.65	0.21	2	80.1	2.69	2
	3.75	67.35%	2.13%	2	78.85	0.21	2	67.1	2.97	2
	1.88	42.66%	2.51%	2	63.65	0.07	2	52.55	2.33	2
	0.94	25.54%	0.32%	2	56.55	3.75	2	43.1	3.25	2
	0	0.23%	0.02%	2	48.70	1.84	2	42.1	0.42	2
G018091	60	97.80%	ND	2	99.65	0.21	2	92.2	0.57	2
	30	96.32%	0.91%	2	97.50	0.28	2	88.7	1.7	2
	15	84.48%	2.73%	2	91.90	0.71	2	82.1	0.71	2
	7.5	76.56%	0.49%	2	78.05	1.63	2	64.15	0.35	2
	3.75	52.08%	1.50%	2	67.00	1.56	2	46.8	1.84	2
	1.88	30.00%	2.36%	2	58.80	2.55	2	40.7	2.4	2
	0.94	15.01%	1.12%	2	50.70	2.26	2	33	2.83	2
	0	0.00%	0.00%	2	47.90	1.70	2	39	0	2
G018100	60	95.55%	0.47%	2	94.30	0.57	2	85.3	2.69	2
	30	95.22%	1.02%	2	91.75	0.78	2	79.1	0.85	2
	15	93.33%	0.49%	2	90.80	0.99	2	77.1	0.28	2
	7.5	88.77%	0.15%	2	84.40	0.85	2	69.9	1.56	2
	3.75	71.03%	4.26%	2	75.20	2.40	2	54.4	0.57	2
	1.88	46.87%	1.95%	2	63.80	0.28	2	45.7	0.28	2
	0.94	27.20%	0.21%	2	54.80	1.56	2	36.5	2.4	2
	0	0.17%	0.04%	2	52.25	1.63	2	42.8	0	2

G018102	60	93.32%	0.10%	2	93.00	0.14	2	84.65	0.64	2
	30	77.08%	2.22%	2	80.40	1.84	2	68.7	2.55	2
	15	52.42%	1.28%	2	70.50	0.28	2	55.4	0.85	2
	7.5	31.87%	0.20%	2	58.80	4.53	2	39.4	4.38	2
	3.75	16.76%	0.26%	2	56.85	1.63	2	33.65	0.07	2
	1.88	8.18%	0.06%	2	54.15	0.21	2	30.35	0.92	2
	0.94	4.12%	0.66%	2	52.25	3.32	2	32.2	0.57	2
	0	0.28%	0.12%	2	50.00	3.68	2	42.45	4.17	2
G018103	60	96.34%	0.21%	2	98.85	0.07	2	89.2	2.12	2
	30	90.93%	0.17%	2	95.15	0.92	2	84.95	0.07	2
	15	80.01%	0.34%	2	87.95	1.63	2	76.7	0.14	2
	7.5	59.92%	0.48%	2	70.90	2.97	2	56.45	3.18	2
	3.75	36.07%	1.23%	2	62.05	2.33	2	44.65	0.49	2
	1.88	19.59%	2.15%	2	58.65	2.76	2	38.2	1.41	2
	0.94	7.98%	0.08%	2	52.75	0.07	2	33.25	2.62	2
	0	0.22%	0.03%	2	50.95	0.64	2	43.1	2.69	2
G018107	60	ND	ND	2	99.50	0.14	2	89.05	0.07	2
	30	ND	ND	2	99.05	0.07	2	91.1	0.85	2
	15	ND	ND	2	97.05	0.64	2	90	3.25	2
	7.5	ND	ND	2	90.45	0.07	2	83.25	2.76	2
	3.75	ND	ND	2	79.60	0.99	2	72.3	2.4	2
	1.88	ND	ND	2	66.20	2.26	2	58.7	4.81	2
	0.94	ND	ND	2	53.40	1.41	2	45.55	3.89	2
	0	ND	ND	2	51.65	2.05	2	47.35	4.03	2
G018111	60	96.87%	0.26%	2	95.65	0.49	2	92.45	0.35	2

	30	95.44%	0.07%	2	89.40	0.71	2	86.5	0.42	2
	15	89.11%	0.91%	2	81.35	0.07	2	75.75	2.62	2
	7.5	71.93%	0.55%	2	70.05	3.32	2	61.15	0.21	2
	3.75	47.48%	0.33%	2	56.45	5.02	2	48.5	0.14	2
	1.88	25.72%	1.46%	2	59.00	3.11	2	44.95	0.07	2
	0.94	12.27%	0.79%	2	51.30	3.39	2	42.3	0.99	2
	0	0.15%	0.01%	2	51.00	3.25	2	45.1	1.41	2
G018116	60	86.65%	0.01%	2	92.95	0.35	2	87.7	0.42	2
	30	62.59%	0.13%	2	79.90	1.84	2	69.75	2.9	2
	15	40.44%	1.09%	2	69.90	0.71	2	60.15	1.34	2
	7.5	23.19%	2.46%	2	60.35	4.74	2	45.45	0.92	2
	3.75	10.65%	0.83%	2	55.45	1.91	2	40.2	1.41	2
	1.88	5.63%	0.47%	2	56.85	1.06	2	40.75	0.64	2
	0.94	1.93%	0.33%	2	57.80	3.96	2	40	0.14	2
	0	0.19%	0.10%	2	50.55	0.21	2	47	1.41	2
G018117	60	98.01%	0.17%	2	99.65	0.07	2	93.65	0.49	2
	30	97.14%	0.69%	2	99.20	0.14	2	92.35	0.07	2
	15	93.65%	0.39%	2	97.75	0.64	2	91.55	0.21	2
	7.5	81.77%	0.50%	2	91.60	0.42	2	81.85	0.35	2
	3.75	56.56%	2.70%	2	80.55	3.46	2	69.25	3.32	2
	1.88	33.70%	3.00%	2	69.30	1.13	2	56.8	2.69	2
	0.94	17.14%	0.83%	2	60.40	1.84	2	48.25	1.34	2
	0	0.18%	0.01%	2	51.95	2.76	2	46.5	1.13	2
G018118	60	98.36%	0.27%	2	99.65	0.07	2	94.35	0.49	2
	30	97.70%	0.19%	2	99.45	0.21	2	88.35	8.13	2

	15	94.59%	1.64%	2	98.30	0.14	2	90.45	0.21	2
	7.5	83.77%	2.47%	2	92.80	1.13	2	81.2	0.57	2
	3.75	64.07%	0.54%	2	81.25	1.06	2	71.5	1.56	2
	1.88	40.72%	2.16%	2	70.65	2.90	2	61.15	3.18	2
	0.94	22.33%	3.64%	2	63.90	0.85	2	51.05	2.76	2
	0	0.14%	0.02%	2	52.95	1.48	2	47.85	1.77	2
G018120	60	98.22%	0.32%	2	99.10	0.14	2	91.65	0.78	2
	30	96.96%	0.43%	2	99.15	0.07	2	89.1	0.57	2
	15	93.71%	0.19%	2	97.05	0.21	2	88.15	0.64	2
	7.5	82.19%	1.50%	2	90.25	0.78	2	80	0.99	2
	3.75	61.25%	1.50%	2	80.30	1.56	2	66.45	1.77	2
	1.88	37.98%	1.29%	2	68.85	0.78	2	52.75	3.04	2
	0.94	18.96%	0.67%	2	59.85	0.64	2	47.25	2.19	2
	0	0.17%	0.02%	2	49.60	0.57	2	46.1	0.99	2
G018121	60	97.55%	0.01%	2	97.25	0.35	2	87.45	0.35	2
	30	92.24%	0.98%	2	92.50	0.71	2	83.6	3.96	2
	15	78.58%	0.52%	2	84.90	0.28	2	76.85	2.9	2
	7.5	57.37%	0.50%	2	73.75	3.04	2	64.5	3.25	2
	3.75	35.06%	0.53%	2	60.70	0.00	2	50.5	3.25	2
	1.88	18.80%	2.81%	2	61.00	1.41	2	45.95	4.31	2
	0.94	9.25%	0.17%	2	56.95	4.31	2	42.8	1.7	2
	0	0.21%	0.00%	2	47.45	3.04	2	44.55	1.34	2

Example 43. Additional Embodiments

[00759] The following numbered embodiments provide additional support for and descriptions of the embodiments herein.

Embodiment A1. An mRNA comprising an open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

Embodiment A2. A composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, optionally wherein the composition comprises lipid nanoparticles.

Embodiment A3. The composition of embodiment A2, wherein the first polypeptide does not comprise a UGI.

Embodiment A4. The composition of embodiment A2 or A3, wherein the molar ratio of the second mRNA to the first mRNA is from 1:1 to 30:1.

Embodiment A5. A method of modifying a target gene comprising delivering to a cell a first mRNA comprising a first open reading frame encoding a first polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, and at least one guide RNA (gRNA).

Embodiment A6. The composition or method of any one of embodiments A2-A5, wherein the molar ratio of the second mRNA to the first mRNA is from 1:1 to 30:1.

Embodiment A7. The composition or method of any one of embodiments A4-A6, wherein the molar ratio is from 2:1 to 30:1.

Embodiment A8. The composition or method of any one of embodiments A4-A6, wherein the molar ratio is from 7:1 to 22:1.

Embodiment A9. A cell comprising a composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA.

Embodiment A10. An engineered cell comprising at least one base edit and/or indel, wherein the base edit and/or indel is made by contacting a cell with a composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA.

Embodiment A11. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A comprises an amino acid sequence with at least 87% identity to SEQ ID NO: 40.

Embodiment A12. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A comprises an amino acid sequence with at least 90% identity to SEQ ID NO: 40.

Embodiment A13. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A comprises an amino acid sequence with at least 95% identity to SEQ ID NO: 40.

Embodiment A14. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A comprises an amino acid sequence with at least 98% identity to SEQ ID NO: 40.

Embodiment A15. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A comprises an amino acid sequence with at least 99% identity to SEQ ID NO: 40.

Embodiment A16. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A comprises the amino acid sequence of SEQ ID NO: 40.

Embodiment A17. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A is a human A3A.

Embodiment A18. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A is a wild-type A3A.

Embodiment A19. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 80% identity to SEQ ID NO: 27.

Embodiment A20. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 90% identity to SEQ ID NO: 27.

Embodiment A21. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 95% identity to SEQ ID NO: 27.

Embodiment A22. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 98% identity to SEQ ID NO: 27.

Embodiment A23. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 99% identity to SEQ ID NO: 27.

Embodiment A24. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises the amino acid sequence of SEQ ID NO: 27.

Embodiment A25. The composition, method, or cell of any one of the preceding embodiments, further comprising at least one guide RNA (gRNA).

Embodiment A26. The composition, method, cell, or engineered cell of any one of the preceding embodiments, comprising a gRNA, wherein the gRNA is an sgRNA.

Embodiment A27. The composition, method, cell, or engineered cell of any one of the preceding embodiments, comprising a gRNA, wherein the gRNA is a dgRNA.

Embodiment A28. The composition, method, cell, or engineered cell of any one of the preceding embodiments, comprising a gRNA, wherein the gRNA is a short-single guide RNA (short-sgRNA) comprising a conserved portion of an sgRNA comprising a hairpin region, wherein the hairpin region lacks at least 5-10 nucleotides and wherein the short-sgRNA comprises a 5' end modification and/or a 3' end modification.

Embodiment A29. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a Cas nickase.

Embodiment A30. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a Class 2 Cas nickase.

Embodiment A31. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a Cas9 nickase.

Embodiment A32. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is an *S. pyogenes* Cas9 nickase.

Embodiment A33. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a D10A SpyCas9 nickase.

Embodiment A34. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a *N. meningitidis* Cas9 nickase.

Embodiment A35. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a D16A Nme2 Cas9 nickase.

Embodiment A36. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase comprises an amino acid sequence having at least 80%, 90%, 95%, 98%, or 99% identity to any one of SEQ ID NOs: 70, 73, or 76.

Embodiment A37. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase comprises the amino acid sequence of any one of SEQ ID NOs: 70, 73, or 76.

Embodiment A38. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 80%, 90%, 95%, 98%, or 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment A39. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 90% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment A40. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 95% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment A41. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided

nickase comprises a nucleotide sequence having at least 98% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment A42. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 99% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment A43. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises the nucleotide sequence of any one of SEQ ID NOs: 71, 72, 74, 75, or 77-90.

Embodiment A44. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 5' UTR with at least 90% identity to any one of SEQ ID NOs: 91-98.

Embodiment A45. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 3' UTR with at least 90% identity to any one of SEQ ID NOs: 99-106.

Embodiment A46. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 5' UTR and a 3' UTR from the same source.

Embodiment A47. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 5' cap selected from Cap0, Cap1, and Cap2.

Embodiment A48. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) comprise minimal adenine codons and/or minimal uridine codons.

Embodiment A49. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) comprise minimal adenine codons.

Embodiment A50. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a

polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) have codons that increase translation of the mRNA in a mammal.

Embodiment A51. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) have codons that increase translation of the mRNA in a mammal, wherein the mammal is a human.

Embodiment A52. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the A3A is located N-terminal to the RNA-guided nickase in the polypeptide.

Embodiment A53. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS).

Embodiment A54. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is at the C-terminus of the RNA-guided nickase.

Embodiment A55. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is at the N-terminus of the RNA-guided nickase, or wherein an NLS is fused to both the N-terminus and C-terminus of the RNA-guided nickase.

Embodiment A56. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein a linker is present between the N-terminus of the RNA-guided nickase and the NLS, optionally wherein the linker is a peptide linker.

Embodiment A57. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS comprises a sequence having at least 80%, 85%, 90%, or 95% identity to any one of SEQ ID NOs: 63 and 110-122.

Embodiment A58. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS comprises the sequence of any one of SEQ ID NOs: 63 and 110-122.

Embodiment A59. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is encoded by a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of any one of SEQ ID NOs: 123-135.

Embodiment A60. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the A3A is located N-terminal to the NLS in the polypeptide.

Embodiment A61. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the RNA-guided nickase is located N-terminal to the NLS in the polypeptide.

Embodiment A62. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of SEQ ID NO:1.

Embodiment A63. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of SEQ ID NO: 4.

Embodiment A64. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein at least 10% of the uridine in the mRNA is substituted with a modified uridine.

Embodiment A65. The mRNA, composition, method, cell or engineered cell of embodiment A64, wherein the modified uridine is one or more of N1-methyl-pseudouridine, pseudouridine, 5-methoxyuridine, or 5-iodouridine.

Embodiment A66. The mRNA, composition, method, cell, or engineered cell of embodiment A64 or A65, wherein the modified uridine is one or both of N1-methyl-pseudouridine or 5-methoxyuridine.

Embodiment A67. The mRNA, composition, method, cell, or engineered cell of any one of embodiments A64-A66, wherein the modified uridine is N1-methyl-pseudouridine.

Embodiment A68. The mRNA, composition, method, cell, or engineered cell of any one of embodiments A64-A66, wherein the modified uridine is 5-methoxyuridine.

Embodiment A69. The mRNA, composition, method, cell, or engineered cell of any one of embodiments A64-A68, wherein 15% to 45% of the uridine is substituted with the modified uridine.

Embodiment A70. The mRNA, composition, method, cell, or engineered cell of any one of embodiments A64-A69, wherein at least 20% or at least 30% of the uridine is substituted with the modified uridine.

Embodiment A71. The mRNA, composition, method, cell, or engineered cell of any one of embodiments A64-A70, wherein at least 80% or at least 90% of the uridine is substituted with the modified uridine.

Embodiment A72. The mRNA, composition, method, cell, or engineered cell of any one of embodiments A64-A71, wherein 100% uridine is substituted with the modified uridine.

Embodiment A73. The mRNA, composition, method, or cell of any one of the preceding embodiments, further encoding a peptide linker between the A3A and RNA-guided nickase, optionally wherein the peptide linker is XTEN.

Embodiment A74. The mRNA, composition, method, or cell of any one of the preceding embodiments, further encoding a peptide linker between the A3A and RNA-guided nickase, wherein the peptide linker comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, or more amino acids.

Embodiment A75. The mRNA, composition, method, or cell of any one of the preceding embodiments, further encoding a peptide linker between the A3A and RNA-guided nickase, wherein the peptide linker comprises one or more sequences selected from SEQ ID NOs: 46-59 and 211-267.

Embodiment A76. A polypeptide encoded by the mRNA of any one of the preceding embodiments A.

Embodiment A77. A vector comprising the mRNA of any one of the preceding embodiments A.

Embodiment A78. An expression construct comprising a promoter operably linked to a sequence encoding the mRNA of any one of the preceding embodiments A.

Embodiment A79. A plasmid comprising the expression construct of embodiment A78.

Embodiment A80. A host cell comprising the vector of embodiment A77, the expression construct of embodiment A78, or the plasmid of embodiment A79.

Embodiment A81. The mRNA or composition of any one of the preceding embodiments, wherein the mRNA or composition is formulated as a lipid nucleic acid assembly composition, optionally a lipid nanoparticle.

Embodiment A82. Use of the mRNA or composition according to any one of the preceding embodiments for modifying a target gene in a cell.

Embodiment A83. Use of the mRNA or composition according to any one of the preceding embodiments for the manufacture of a medicament for modifying a target gene in a cell.

Embodiment A84. A method of modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising:

(a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase;

(b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs.

Embodiment A85. The method of embodiment A84, wherein parts (a) and (b) are in separate lipid nucleic acid assembly compositions.

Embodiment A86. The method of embodiment A84, wherein parts (a) and (b) are in the same lipid nucleic acid assembly composition.

Embodiment A87. The method of embodiment A84, wherein parts (a) and (c) are in separate lipid nucleic acid assembly compositions.

Embodiment A88. The method of embodiment A84, wherein parts (a) and (c) are in the same lipid nucleic acid assembly composition.

Embodiment A89. The method of embodiment A84, wherein parts (b) and (c) are in separate lipid nucleic acid assembly compositions.

Embodiment A90. The method of embodiment A84, wherein parts (a) and (c) are in the same lipid nucleic acid assembly composition, and part (b) is in a separate lipid nucleic acid assembly composition.

Embodiment A91. The method of embodiment A84, wherein parts (a), (b), and (c) are each in separate lipid nucleic acid assembly compositions.

Embodiment A92. The method of embodiment A84, wherein parts (a), (b), and (c) are in the same lipid nucleic acid assembly composition.

Embodiment A93. The method of any one of embodiments A84-A88, A90, and A92, wherein the one or more guide RNAs are each in separate lipid nucleic acid assembly compositions.

Embodiment A94. The method of any one of embodiments A84-A93, comprising delivering to the cell a lipid nucleic acid assembly composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI) in the same lipid nucleic acid assembly composition.

Embodiment A95. The method of any one of embodiments A84-A93, comprising delivering to the cell a first lipid nucleic acid assembly composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, and a second lipid nucleic acid assembly composition comprising a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI).

Embodiment A96. The method of any one of embodiments A84-A95, further comprising delivering one or more guide RNAs in one or more lipid nucleic acid assembly compositions that are separate from the lipid nucleic acid assembly compositions comprising the A3A and UGI.

Embodiment A97. The method of any one of embodiments A84-A96, wherein at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 lipid nucleic acid assembly compositions are delivered to the cell.

Embodiment A98. The method of any one of embodiments A84-A97, wherein at least one lipid nucleic acid assembly composition comprises lipid nanoparticle (LNPs), optionally wherein all lipid nucleic acid assembly compositions comprise LNPs.

Embodiment A99. The method of any one of embodiments A84-A98, wherein at least one lipid nucleic acid assembly composition is a lipoplex composition.

Embodiment A100. The method of any one of embodiments A84-A99, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid.

Embodiment A101. The method of any one of embodiments A84-A100, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid and wherein the ionizable lipid comprises a biodegradable ionizable lipid.

Embodiment A102. The method of any one of embodiments A84-A101, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid and wherein the ionizable lipid has a PK value in the range of pKa in the range of from about 5.1 to about 7.4, such as from about 5.5 to about 6.6, from about 5.6 to about 6.4, from about 5.8 to about 6.2, or from about 5.8 to about 6.5.

Embodiment A103. The method of any one of embodiments A84-A102, wherein the lipid nucleic acid assembly composition comprises an amine lipid.

Embodiment A104. The method of any one of embodiments A84-A103, wherein the lipid nucleic acid assembly composition comprises Lipid A.

Embodiment A105. The method of any one of embodiments A84-A104, wherein the lipid nucleic acid assembly composition comprises a helper lipid.

Embodiment A106. The method of any one of embodiments A84-A105, wherein the lipid nucleic acid assembly composition comprises a helper lipid, wherein the helper lipid is cholesterol.

Embodiment A107. The method of any one of embodiments A84-A106, wherein the lipid nucleic acid assembly composition comprises a stealth lipid.

Embodiment A108. The method of any one of embodiments A84-A107, wherein the lipid nucleic acid assembly composition comprises a stealth lipid, wherein the stealth lipid is PEG2k-DMG.

Embodiment A109. The method of any one of embodiments A84-A108, wherein the lipid nucleic acid assembly composition comprises a neutral lipid.

Embodiment A110. The method of any one of embodiments A84-A109, wherein the lipid nucleic acid assembly composition comprises a neutral lipid, wherein the neutral lipid is DSPC.

Embodiment A111. The method of any one of embodiments A84-A110, wherein the lipid nucleic acid assembly composition comprises a neutral lipid, wherein the neutral lipid is present at about 9 mol-%.

Embodiment A112. The method of any one of embodiments A84-A111, wherein the lipid nucleic acid assembly composition comprises a stealth lipid, wherein the stealth lipid is present at about 3 mol-%.

Embodiment A113. The method of any one of embodiments A84-A112, wherein the lipid nucleic acid assembly composition comprises a helper lipid, wherein the helper lipid is present at about 38 mol-%.

Embodiment A114. The method of any one of embodiments A84-A113, wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6.

Embodiment A115. The method of any one of embodiments A84-A114, wherein the lipid nucleic acid assembly composition comprises about 50 mol-% amine lipid such as Lipid A; about 9 mol-% neutral lipid such as DSPC; about 3 mol-% of stealth lipid such as a PEG lipid, such as PEG2k-DMG, and the remainder of the lipid component is helper lipid such as cholesterol wherein the N/P ratio of the is about 6.

Embodiment A116. The method of any one of embodiments A84-A115, comprising one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, and /or one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and/or one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

Embodiment A117. The method of any one of embodiments A84-A115, comprising at least two gRNAs selected from: one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

Embodiment A118. The method of any one of embodiments A84-A115, comprising one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

Embodiment A119. The method of any one of embodiments A84-A115, comprising one gRNA selected from a gRNA that targets TRAC, TRBC, B2M, HLA-A, or CIITA.

Embodiment A120. The method of any one of embodiments A84-A116, and A119, wherein the gRNA targets TRBC, wherein the gRNA comprises a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous

nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

Embodiment A121. The method of any one of embodiments A84-A116, and A119, wherein the gRNA targets TRBC, wherein the gRNA comprises a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

Embodiment A122. The method of any one of embodiments A84-A115, comprising at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or B2M, wherein the two guide RNAs do not target the same gene.

Embodiment A123. The method of any one of embodiments A84-A115, comprising at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or HLA-A, wherein the two guide RNAs do not target the same gene.

Embodiment A124. The method of any one of embodiments A84-A115, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC.

Embodiment A125. The method of any one of embodiments A84-A115, comprising one guide RNA that targets B2M, and one gRNA that targets CIITA.

Embodiment A126. The method of any one of embodiments A84-A115, comprising one guide RNA that targets HLA-A, and one gRNA that targets CIITA, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

Embodiment A127. The method of any one of embodiments A84-A115, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA that targets B2M.

Embodiment A128. The method of any one of embodiments A84-A115, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA

that targets HLA-A, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

Embodiment A129. The method of any one of embodiments A84-A115, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets B2M, and one gRNA that targets CIITA.

Embodiment A130. The method of any one of embodiments A84-A115, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets HLA-A, and one gRNA that targets CIITA, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

Embodiment A131. The method of any one of the preceding embodiments, wherein the method generates a cytosine (C) to thymine (T) conversion within a target sequence.

Embodiment A132. The method of any one of the preceding embodiments, wherein the method causes at least 60% C-to-T conversion relative to the total edits in the target sequence.

Embodiment A133. The method of any one of the preceding embodiments, wherein the method causes at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% C-to-T conversion relative to the total edits in the target sequence.

Embodiment A134. The method of any one of the preceding embodiments, wherein the ratio of C-to-T conversion to unintended edits is larger than 1:1.

Embodiment A135. The method of any one of the preceding embodiments, wherein the ratio of C-to-T conversion to unintended edits is from 2:1 to 99:1.

Embodiment A136. The method of any one of the preceding embodiments, wherein the ratio of C-to-T conversion to unintended edits is 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, or 8:1.

Embodiment A137. The method of any one of the preceding embodiments, wherein the method causes the A3A to make a base edit corresponding to any one of positions -1 to 10 relative to the 5' end of the guide sequence.

Embodiment A138. The method of any one of the preceding embodiments, wherein the method causes the A3A to make a base edit at a position 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the guide sequence.

Embodiment A139. The method of any one of the preceding embodiments, wherein the first mRNA, the second mRNA, and the guide RNA if present, delivered at a ratio of about 6:2:3 (w:w:w).

Embodiment A140. The method of any one of the preceding embodiments, wherein the mRNA, composition, or LNP is administered at a total RNA amount from 5 to 600 ng.

Embodiment A141. The method of any one of the preceding embodiments, wherein the total RNA amount is from 8 to 550 ng.

Embodiment A142. The method of any one of the preceding embodiments, wherein the total RNA amount is from 35 to 550 ng.

Embodiment A143. The method of any one of the preceding embodiments, wherein the total RNA amount is from 70 to 550 ng.

Embodiment A144. The method of any one of the preceding embodiments, wherein the total RNA amount is from 138 to 550 ng.

Embodiment A145. The method of any one of the preceding embodiments, wherein the total RNA amount is from 275 to 550 ng.

Embodiment A146. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a lymphocyte.

Embodiment A147. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a B lymphocyte.

Embodiment A148. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a T lymphocyte.

Embodiment A149. The method or use of any one of the preceding embodiments, wherein the modification of the target gene is *in vivo*.

Embodiment A150. The method or use of any one of the preceding embodiments, wherein the modification of the target gene is *ex vivo*.

Embodiment A151. The method or use of any one of the preceding embodiments, wherein the modification of the target gene reduces or eliminates expression of the target gene.

Embodiment A152. The method or use of any one of the preceding embodiments, wherein the genome editing or modification of the target gene reduces expression of the target gene by at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%.

Embodiment A153. The method or use of any one of the preceding embodiments, wherein the genome editing or modification of the target gene produces a missense mutation in the gene.

Embodiment A154. A polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

Embodiment A155. A composition comprising a first polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase and a second polypeptide comprising a UGI, wherein the second polypeptide is different from the first polypeptide.

Embodiment A156. The polypeptide of embodiment A152 or A153, wherein the A3A is fused to the RNA-guided nickase via a peptide linker, optionally XTEN.

Embodiment A157. The polypeptide of embodiment A152 or A153, wherein the A3A is attached to a linker comprising an organic molecule, polymer, or chemical moiety.

Embodiment A158. A pharmaceutical composition comprising the mRNA, composition, or polypeptide of any of the preceding embodiments Aand a pharmaceutically acceptable carrier.

Embodiment A159. A kit comprising the mRNA, composition, or polypeptide of any of the preceding embodiments A.

Embodiment A160. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase includes: the A3A, a linker, and the RNA-guided nickase in amino to carboxy terminal order.

Embodiment A161. A method of altering a DNA sequence within a TRAC gene, comprising delivering to a cell:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

(b) a nucleic acid encoding a gRNA of (a.).

Embodiment A162. A method of reducing the expression of a TRAC gene, comprising delivering to a cell:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

(b) a nucleic acid encoding a gRNA of (a.).

Embodiment A163. A method of immunotherapy comprising administering a composition comprising an engineered cell to a subject,

wherein the cell comprises a genomic modification of at least one nucleotide within the genomic coordinates selected from:

chr14: 22547596-22547616; chr14: 22550570-22550590; chr14: 22547763-22547783; chr14: 22550596-22550616; chr14: 22550566-22550586; chr14: 22547753-22547773; chr14: 22550601-22550621; chr14: 22550599-22550619; chr14: 22547583-22547603; chr14: 22547671-22547691; chr14: 22547770-22547790; chr14: 22547676-22547696; chr14: 22547772-22547792; chr14: 22547771-22547791; chr14: 22547733-22547753; chr14: 22547776-22547796; or

wherein the cell is engineered by delivering to the cell:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

(b) a nucleic acid encoding a gRNA of (a.).

Embodiment A164. The method of any one of embodiments A161-163, wherein the guide sequence comprises any one of SEQ ID NOs: 706-709.

Embodiment A165. The method of any one of embodiments A161-164, wherein the guide sequence comprises any one of SEQ ID NOs: 706-708.

Embodiment A166. The method of any one of embodiments A161-165, wherein the guide sequence comprises SEQ ID NO: 706.

Embodiment A167. The method of any one of embodiments A161-165, wherein the guide sequence comprises SEQ ID NO: 707.

Embodiment A168. The method of any one of embodiments A161-165, wherein the guide sequence comprises SEQ ID NO: 708.

Embodiment A169. A method of altering a DNA sequence within a TRBC1 and/or TRBC2 gene, comprising delivering a composition to a cell, wherein the composition comprises:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

(b) a nucleic acid encoding a guide RNA of (a.).

Embodiment A170. A method of reducing the expression of a TRBC1 and/or TRBC2 gene, comprising delivering a composition to a cell, wherein the composition comprises:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

(b) a nucleic acid encoding a guide RNA of (a.).

Embodiment A171. A method of immunotherapy comprising administering a composition comprising an engineered cell to a subject,

wherein the cell comprises a modification of at least one nucleotide within the genomic coordinates selected from: chr7: 142791757-142791777; chr7: 142801104-

142801124; chr7: 142791811-142791831; chr7: 142801158-142801178; chr7: 142792728-142792748; chr7: 142791719-142791739; chr7: 142791766-142791786; chr7: 142801113-142801133; chr7: 142791928-142791948; chr7: 142801275-142801295; chr7: 142792062-142792082; chr7: 142801409-142801429; chr7: 142792713-142792733; chr7: 142802126-142802146; chr7: 142791808-142791828; chr7: 142801155-142801175; chr7: 142792003-142792023; chr7: 142801350-142801370; chr7: 142791760-142791780; chr7: 142791715-142791735; chr7: 142792781-142792801; chr7: 142792040-142792060; chr7: 142801387-142801407; chr7: 142791862-142791882; chr7: 142791716-142791736; chr7: 142791787-142791807; chr7: 142791759-142791779; chr7: 142801106-142801126; chr7: 142791807-142791827; chr7: 142801154-142801174; chr7: 142791879-142791899; chr7: 142801226-142801246; chr7: 142791805-142791825; chr7: 142791700-142791720; chr7: 142791765-142791785; chr7: 142801112-142801132; chr7: 142791820-142791840; chr7: 142791872-142791892; chr7: 142801219-142801239; chr7: 142791700-142791720; chr7: 142791806-142791826; chr7: 142801153-142801173; chr7: 142792035-142792055; chr7: 142792724-142792744; chr7: 142792754-142792774; chr7: 142791804-142791824; chr7: 142792684-142792704; chr7: 142791823-142791843; chr7: 142792728-142792748; chr7: 142792721-142792741; chr7: 142792749-142792769; chr7: 142792685-142792705; chr7: 142791816-142791836; chr7: 142801163-142801183; chr7: 142792686-142792706; chr7: 142791793-142791813; chr7: 142793110-142793130; chr7: 142791815-142791835; chr7: 142801162-142801182; chr7: 142792770-142792790; chr7: 142792047-142792067; chr7: 142801394-142801414; chr7: 142791871-142791891; chr7: 142801218-142801238; chr7: 142791894-142791914; chr7: 142792723-142792743; chr7: 142792724-142792744; chr7: 142791897-142791917; chr7: 142801244-142801264; chr7: 142792757-142792777; chr7: 142792740-142792760; chr7: 142792758-142792778; or

wherein the cell is engineered by delivering to a cell:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous

nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

(b) a nucleic acid encoding a guide RNA of (a.).

Embodiment A172. The method of any one of embodiments A169-171, wherein the guide sequence comprises any one of SEQ ID NOs: 618-627.

Embodiment A173. The method of any one of embodiments A169-172, wherein the guide sequence comprises any one of SEQ ID NOs: 618-621.

Embodiment A174. The method of any one of embodiments A169-173, wherein the guide sequence comprises SEQ ID NO: 618.

Embodiment A175. The method of any one of embodiments A169-173, wherein the guide sequence comprises SEQ ID NO: 619.

Embodiment A176. The method of any one of embodiments A169-173, wherein the guide sequence comprises SEQ ID NO: 620.

Embodiment A177. The method of any one of embodiments A169-173, wherein the guide sequence comprises SEQ ID NO: 621.

Embodiment A178. The method of any one of embodiments A161-177, wherein the composition further comprises the mRNA or composition of any one of embodiments A1-76.

Embodiment A179. A composition comprising:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); and optionally

(b) the mRNA or composition of any one of embodiments A1-A76.

Embodiment A180. The composition of embodiment A179, wherein the guide sequence comprises any one of SEQ ID NOs: 706-709.

Embodiment A181. The composition of embodiment A179 or A180, for use in altering a DNA sequence within the TRAC gene in a cell.

Embodiment A182. The composition of any one of embodiments A179-A181 for use in reducing the expression of the TRAC gene in a cell.

Embodiment A183. The composition of any one of embodiments A179-A182, for use in immunotherapy of a subject.

Embodiment A184. A composition comprising:

(a) a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); and optionally

(b) the mRNA or composition of any one of embodiments A11-A75.

Embodiment A185. The composition of embodiment A184, wherein the guide sequence comprises any one of SEQ ID NOs: 618-621.

Embodiment A186. The composition of embodiment A184 or A185, for use in altering a DNA sequence within the TRBC1 and/or TRBC2 gene in a cell.

Embodiment A187. The composition of any one of embodiments A184 to A186 for use in reducing the expression of the TRBC1 and/or TRBC2 gene in a cell.

Embodiment A188. The composition of any one of embodiments A184 to A187, for use in immunotherapy of a subject.

Embodiment A189. A cell, altered by the method of any one of embodiments A120-A121, and A161-A178 .

Embodiment A190. The cell according to embodiment A189, wherein the cell is altered ex vivo.

Embodiment A191. The cell according to embodiment A189 or 190, wherein the cell is a T cell.

Embodiment A192. The cell according to any of embodiments A189-A191, wherein the cell is a $CD4^+$ or $CD8^+$ T cell.

Embodiment A193. The cell according to any of embodiments A189-A192, wherein the cell is a mammalian, primate, or human cell.

Embodiment A194. The cell according to any of embodiments A189-A193, for use in immunotherapy of a subject.

Embodiment A195. An engineered cell which has reduced or eliminated surface expression of TRAC, comprising a genetic modification in a human TRAC gene,

wherein the genetic modification comprises a modification of at least one nucleotide within the genomic coordinates selected from:
chr14: 22547596-22547616; chr14: 22550570-22550590; chr14: 22547763-22547783; chr14: 22550596-22550616; chr14: 22550566-22550586; chr14: 22547753-22547773; chr14: 22550601-22550621; chr14: 22550599-22550619; chr14: 22547583-22547603; chr14: 22547671-22547691; chr14: 22547770-22547790; chr14: 22547676-22547696; chr14: 22547772-22547792; chr14: 22547771-22547791; chr14: 22547733-22547753; chr14: 22547776-22547796.

Embodiment A196. An engineered cell which has reduced or eliminated surface expression of TRBC1/2, comprising a genetic modification in a human TRBC1/2 gene, wherein the genetic modification comprises a modification of at least one nucleotide within the genomic coordinates selected from:
chr7: 142791757-142791777; chr7: 142801104-142801124; chr7: 142791811-142791831; chr7: 142801158-142801178; chr7: 142792728-142792748; chr7: 142791719-142791739; chr7: 142791766-142791786; chr7: 142801113-142801133; chr7: 142791928-142791948; chr7: 142801275-142801295; chr7: 142792062-142792082; chr7: 142801409-142801429; chr7: 142792713-142792733; chr7: 142802126-142802146; chr7: 142791808-142791828; chr7: 142801155-142801175; chr7: 142792003-142792023; chr7: 142801350-142801370; chr7: 142791760-142791780; chr7: 142791715-142791735; chr7: 142792781-142792801; chr7: 142792040-142792060; chr7: 142801387-142801407; chr7: 142791862-142791882; chr7: 142791716-142791736; chr7: 142791787-142791807; chr7: 142791759-142791779; chr7: 142801106-142801126; chr7: 142791807-142791827; chr7: 142801154-142801174; chr7: 142791879-142791899; chr7: 142801226-142801246; chr7: 142791805-142791825; chr7: 142791700-142791720; chr7: 142791765-142791785; chr7: 142801112-142801132; chr7: 142791820-142791840; chr7: 142791872-142791892; chr7: 142801219-142801239; chr7: 142791700-142791720; chr7: 142791806-142791826; chr7: 142801153-142801173; chr7: 142792035-142792055; chr7: 142792724-142792744; chr7: 142792754-142792774; chr7: 142791804-142791824; chr7: 142792684-142792704; chr7: 142791823-142791843; chr7: 142792728-142792748; chr7: 142792721-142792741; chr7: 142792749-142792769; chr7: 142792685-142792705; chr7: 142791816-142791836; chr7: 142801163-142801183; chr7: 142792686-142792706; chr7: 142791793-142791813; chr7: 142793110-142793130; chr7: 142791815-142791835; chr7: 142801162-

142801182; chr7: 142792770-142792790; chr7: 142792047-142792067; chr7: 142801394-142801414; chr7: 142791871-142791891; chr7: 142801218-142801238; chr7: 142791894-142791914; chr7: 142792723-142792743; chr7: 142792724-142792744; chr7: 142791897-142791917; chr7: 142801244-142801264; chr7: 142792757-142792777; chr7: 142792740-142792760; chr7: 142792758-142792778.

Embodiment A197. One or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising:

(a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising an APOBEC3A deaminase (A3A) and an RNA-guided nickase;

(b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs.

Embodiment A198. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is an immune cell.

Embodiment A199. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a lymphocyte.

Embodiment A200. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a T cell.

[00760] The following numbered embodiments provide additional support for and descriptions of the embodiments herein.

Embodiment B1. An mRNA comprising an open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

Embodiment B2. A composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, optionally wherein the composition comprises lipid nanoparticles.

Embodiment B3. The composition of embodiment B2, wherein the first open reading frame does not comprise a sequence encoding a UGI.

Embodiment B4. The composition of embodiment B2 or B3, wherein the molar ratio of the second mRNA to the first mRNA is from 1:1 to 30:1.

Embodiment B5. The composition of any one of embodiments B2-B4, wherein the composition comprises a first composition and a second composition, wherein the first

composition comprises a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase and does not comprise a uracil glycosylase inhibitor (UGI), and the second composition comprises a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, optionally wherein the compositions comprise lipid nanoparticles.

Embodiment B6. The composition of any one of embodiments B2-B5, wherein the first mRNA and the second mRNAs are in the same or separate vials.

Embodiment B7. A method of modifying a target gene comprising delivering to a cell a first mRNA comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase and an RNA-guided nickase, a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, and at least one guide RNA (gRNA), wherein if the nickase is a SpyCas9 nickase, then the gRNA is a SpyCas9 gRNA, and if the nickase is a NmeCas9 nickase, then the gRNA is a Nme gRNA.

Embodiment B8. The composition or method of any one of embodiments B2-B7, wherein the molar ratio of the second mRNA to the first mRNA is from 1:1 to 30:1.

Embodiment B9. The composition or method of any one of embodiments B2-B7, wherein the molar ratio is from 2:1 to 30:1.

Embodiment B10. The composition or method of any one of embodiments B2-B7, wherein the molar ratio is from 7:1 to 22:1.

Embodiment B11. A cell comprising a composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA.

Embodiment B12. An engineered cell comprising at least one base edit and/or indel, wherein the base edit and/or indel is made by contacting a cell with a composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA.

Embodiment B13. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the cytidine deaminase is

(i) an enzyme of APOBEC family, optionally an enzyme of APOBEC3 subgroup;
(ii) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023;
(iii) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1013;
(iv) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009; or
(v) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 976, 981, 984, 986, and 1014-1023.

Embodiment B14. The mRNA, composition, method, cell, or engineered cell of embodiment B13, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 41, and 960-1023.

Embodiment B15. The mRNA, composition, method, cell, or engineered cell of embodiment B13, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 41, and 960-1013.

Embodiment B16. The mRNA, composition, method, cell, or engineered cell of embodiment B13, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009.

Embodiment B17. The mRNA, composition, method, cell, or engineered cell of embodiment B13, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 976, 981, 984, 986, 1014-1023.

Embodiment B18. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the cytidine deaminase is an APOBEC3A deaminase (A3A).

Embodiment B19. The mRNA, composition, method, cell, or engineered cell of embodiment B18, wherein the A3A comprises an amino acid sequence with at least 87% identity to SEQ ID NO: 40.

Embodiment B20. The mRNA, composition, method, cell, or engineered cell of embodiment B18, wherein the A3A comprises an amino acid sequence with at least 90% identity to SEQ ID NO: 40.

Embodiment B21. The mRNA, composition, method, cell, or engineered cell of embodiment B18, wherein the A3A comprises an amino acid sequence with at least 95% identity to SEQ ID NO: 40.

Embodiment B22. The mRNA, composition, method, cell, or engineered cell of embodiment B18, wherein the A3A comprises an amino acid sequence with at least 98% identity to SEQ ID NO: 40.

Embodiment B23. The mRNA, composition, method, cell, or engineered cell of embodiment B18, wherein the A3A comprises an amino acid sequence with at least 99% identity to SEQ ID NO: 40.

Embodiment B24. The mRNA, composition, method, cell, or engineered cell of embodiment B18, wherein the A3A comprises the amino acid sequence of SEQ ID NO: 40.

Embodiment B25. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B18-B24, wherein the A3A is a human A3A.

Embodiment B26. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B18-B24, wherein the A3A is a wild-type A3A.

Embodiment B27. The mRNA, composition, method, cell, or engineered cell of embodiment B18, wherein the A3A comprises an amino acid sequence with at least 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 976, 977, 993-1006, and 1009.

Embodiment B28. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 80% identity to SEQ ID NO: 27.

Embodiment B29. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 90% identity to SEQ ID NO: 27.

Embodiment B30. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 95% identity to SEQ ID NO: 27.

Embodiment B31. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 98% identity to SEQ ID NO: 27.

Embodiment B32. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises an amino acid sequence with at least 99% identity to SEQ ID NO: 27.

Embodiment B33. The composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the UGI comprises the amino acid sequence of SEQ ID NO: 27.

Embodiment B34. The composition, method, or cell of any one of the preceding embodiments, further comprising at least one guide RNA (gRNA).

Embodiment B35. The composition, method, cell, or engineered cell of any one of the preceding embodiments, comprising a gRNA, wherein the gRNA is an sgRNA.

Embodiment B36. The composition, method, cell, or engineered cell of any one of the preceding embodiments, comprising a gRNA, wherein the gRNA is a dgRNA.

Embodiment B37. The composition, method, cell, or engineered cell of any one of the preceding embodiments, comprising a gRNA, wherein the gRNA is a short-single guide RNA (short-sgRNA) comprising a conserved portion of an sgRNA comprising a hairpin region, wherein the hairpin region lacks at least 5-10 nucleotides and wherein the short-sgRNA comprises a 5' end modification or a 3' end modification or both.

Embodiment B38. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a Cas nickase.

Embodiment B39. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a Class 2 Cas nickase.

Embodiment B40. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is a Cas9 nickase.

Embodiment B41. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase is an *S. pyogenes* (Spy) Cas9 nickase.

Embodiment B42. The mRNA, composition, method, cell, or engineered cell of embodiment B41, wherein the RNA-guided nickase is a D10A SpyCas9 nickase.

Embodiment B43. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase comprises an amino acid sequence having at least 80%, 90%, 95%, 98%, or 99% identity to any one of SEQ ID NOs: 70, 73, or 76.

Embodiment B44. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the RNA-guided nickase comprises the amino acid sequence of any one of SEQ ID NOs: 70, 73, or 76.

Embodiment B45. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase

comprises a nucleotide sequence having at least 80%, 90%, 95%, 98%, or 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment B46. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 90% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment B47. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 95% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment B48. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 98% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment B49. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence having at least 99% to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

Embodiment B50. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the sequence encoding the RNA-guided nickase comprises the nucleotide sequence of any one of SEQ ID NOs: 71, 72, 74, 75, or 77-90.

Embodiment B51. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B1-37, wherein the RNA-guided nickase is a *N. meningitidis* (Nme) Cas9 nickase.

Embodiment B52. The mRNA, composition, method, cell, or engineered cell of embodiment B51, wherein the RNA-guided nickase is a D16A NmeCas9 nickase, optionally a D16A Nme2Cas9 .

Embodiment B53. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B1-B37, B51, or B52, wherein the sequence encoding the RNA-guided nickase comprises the nucleotide sequence of any one of SEQ ID NOs: 380 and 387.

Embodiment B54. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 5' UTR with at least 90% identity to any one of SEQ ID NOs: 91-98.

Embodiment B55. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 3' UTR with at least 90% identity to any one of SEQ ID NOs: 99-106.

Embodiment B56. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 5' UTR and a 3' UTR from the same source.

Embodiment B57. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the mRNA comprises a 5' cap selected from Cap0, Cap1, and Cap2.

Embodiment B58. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) comprise minimal adenine codons and/or minimal uridine codons.

Embodiment B59. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) comprise minimal adenine codons.

Embodiment B60. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) have codons that increase translation of the mRNA in a mammal.

Embodiment B61. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) have codons that increase translation of the mRNA in a mammal, wherein the mammal is a human.

Embodiment B62. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the cytidine deaminase is located N-terminal to the RNA-guided nickase in the polypeptide.

Embodiment B63. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS).

Embodiment B64. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is at the C-terminus of the RNA-guided nickase.

Embodiment B65. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is at the N-terminus of the RNA-guided nickase, or wherein an NLS is fused to both the N-terminus and C-terminus of the RNA-guided nickase.

Embodiment B66. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein a linker is present between the N-terminus of the RNA-guided nickase and the NLS, optionally wherein the linker is a peptide linker.

Embodiment B67. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS comprises a sequence having at least 80%, 85%, 90%, or 95% identity to any one of SEQ ID NOs: 63 and 110-122.

Embodiment B68. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS comprises the sequence of any one of SEQ ID NOs: 63 and 110-122.

Embodiment B69. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is encoded by a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of any one of SEQ ID NOs: 123-135.

Embodiment B70. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the cytidine deaminase is located N-terminal to the NLS in the polypeptide.

Embodiment B71. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the RNA-guided nickase is located N-terminal to the NLS in the polypeptide.

Embodiment B72. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of SEQ ID NO:1.

Embodiment B73. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of SEQ ID NO: 4.

Embodiment B74. The mRNA, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein at least 10% of the uridine in the mRNA is substituted with a modified uridine.

Embodiment B75. The mRNA, composition, method, cell or engineered cell of embodiment B64, wherein the modified uridine is one or more of N1-methyl-pseudouridine, pseudouridine, 5-methoxyuridine, or 5-iodouridine.

Embodiment B76. The mRNA, composition, method, cell, or engineered cell of embodiment B64 or 65, wherein the modified uridine is one or both of N1-methyl-pseudouridine or 5-methoxyuridine.

Embodiment B77. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B74-B76, wherein the modified uridine is N1-methyl-pseudouridine.

Embodiment B78. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B74-B76, wherein the modified uridine is 5-methoxyuridine.

Embodiment B79. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B74-B78, wherein 15% to 45% of the uridine is substituted with the modified uridine.

Embodiment B80. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B74-B79, wherein at least 20% or at least 30% of the uridine is substituted with the modified uridine.

Embodiment B81. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B74-B80, wherein at least 80% or at least 90% of the uridine is substituted with the modified uridine.

Embodiment B82. The mRNA, composition, method, cell, or engineered cell of any one of embodiments B74-B81, wherein 100% uridine is substituted with the modified uridine.

Embodiment B83. The mRNA, composition, method, or cell of any one of the preceding embodiments, further encoding a peptide linker between the cytidine deaminase and RNA-guided nickase, optionally wherein the peptide linker is XTEN.

Embodiment B84. The mRNA, composition, method, or cell of any one of the preceding embodiments, further encoding a peptide linker between the cytidine deaminase and RNA-guided nickase, wherein the peptide linker comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, or more amino acids.

Embodiment B85. The mRNA, composition, method, or cell of any one of the preceding embodiments, further encoding a peptide linker between the cytidine deaminase and RNA-guided nickase, wherein the peptide linker comprises one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272.

Embodiment B86. A polypeptide encoded by the mRNA of any one of the preceding embodiments.

Embodiment B87. A ribonucleoprotein complex (RNP) comprising (i) a polypeptide encoded by any one the mRNAs of any one of the preceding embodiments; and (ii) a guide RNA.

Embodiment B88. A vector comprising the mRNA of any one of the preceding embodiments.

Embodiment B89. An expression construct comprising a promoter operably linked to a sequence encoding the mRNA of any one of the preceding embodiments.

Embodiment B90. A plasmid comprising the expression construct of embodiment B89.

Embodiment B91. A host cell comprising the vector of embodiment B88, the expression construct of embodiment B89, or the plasmid of embodiment B90.

Embodiment B92. The mRNA or composition of any one of the preceding embodiments, wherein the mRNA or composition is formulated as a lipid nucleic acid assembly composition, optionally a lipid nanoparticle.

Embodiment B93. Use of the mRNA or composition according to any one of the preceding embodiments for modifying a target gene in a cell.

Embodiment B94. Use of the mRNA or composition according to any one of the preceding embodiments for the manufacture of a medicament for modifying a target gene in a cell.

Embodiment B95. A method of modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising:

(a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase;

(b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs.

Embodiment B96. The method of embodiment B95, wherein parts (a) and (b) are in separate lipid nucleic acid assembly compositions.

Embodiment B97. The method of embodiment B95, wherein parts (a) and (b) are in the same lipid nucleic acid assembly composition.

Embodiment B98. The method of embodiment B95, wherein parts (a) and (c) are in separate lipid nucleic acid assembly compositions.

Embodiment B99. The method of embodiment B95, wherein parts (a) and (c) are in the same lipid nucleic acid assembly composition.

Embodiment B100. The method of embodiment B95, wherein parts (b) and (c) are in separate lipid nucleic acid assembly compositions.

Embodiment B101. The method of embodiment B95, wherein parts (a) and (c) are in the same lipid nucleic acid assembly composition, and part (b) is in a separate lipid nucleic acid assembly composition.

Embodiment B102. The method of embodiment B95, wherein parts (a), (b), and (c) are each in separate lipid nucleic acid assembly compositions.

Embodiment B103. The method of embodiment B95, wherein parts (a), (b), and (c) are in the same lipid nucleic acid assembly composition.

Embodiment B104. The method of any one of embodiments B95, and B98-B103, wherein the one or more guide RNAs are each in separate lipid nucleic acid assembly compositions.

Embodiment B105. The method of any one of embodiments B95-B104, comprising delivering to the cell a lipid nucleic acid assembly composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI) in the same lipid nucleic acid assembly composition.

Embodiment B106. The method of any one of embodiments B95-B104, comprising delivering to the cell a first lipid nucleic acid assembly composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a second lipid nucleic acid assembly composition comprising a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI).

Embodiment B107. The method of any one of embodiments B95-B104, further comprising delivering one or more guide RNAs in one or more lipid nucleic acid assembly compositions that are separate from the lipid nucleic acid assembly compositions comprising the cytidine deaminase and UGI.

Embodiment B108. The method of any one of embodiments B95-B107, wherein at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 lipid nucleic acid assembly compositions are delivered to the cell.

Embodiment B109. The method of any one of embodiments B95-B108, wherein at least one lipid nucleic acid assembly composition comprises lipid nanoparticle (LNPs), optionally wherein all lipid nucleic acid assembly compositions comprise LNPs.

Embodiment B110. The method of any one of embodiments B95-B109, wherein at least one lipid nucleic acid assembly composition is a lipoplex composition.

Embodiment B111. The method of any one of embodiments B95-B110, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid.

Embodiment B112. The method of any one of embodiments B95-B111, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid and wherein the ionizable lipid comprises a biodegradable ionizable lipid.

Embodiment B113. The method of any one of embodiments B95-B112, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid and wherein the ionizable lipid has a PK value in the range of pKa in the range of from about 5.1 to about 7.4, such as from about 5.5 to about 6.6, from about 5.6 to about 6.4, from about 5.8 to about 6.2, or from about 5.8 to about 6.5.

Embodiment B114. The method of any one of embodiments B95-B113, wherein the lipid nucleic acid assembly composition comprises an amine lipid.

Embodiment B115. The method of any one of embodiments B95-B114, wherein the lipid nucleic acid assembly composition comprises Lipid A.

Embodiment B116. The method of any one of embodiments B95-B115, wherein the lipid nucleic acid assembly composition comprises a helper lipid.

Embodiment B117. The method of any one of embodiments B95-B116, wherein the lipid nucleic acid assembly composition comprises a helper lipid, wherein the helper lipid is cholesterol.

Embodiment B118. The method of any one of embodiments B95-B117, wherein the lipid nucleic acid assembly composition comprises a stealth lipid.

Embodiment B119. The method of any one of embodiments B95-B118, wherein the lipid nucleic acid assembly composition comprises a stealth lipid, wherein the stealth lipid is PEG2k-DMG.

Embodiment B120. The method of any one of embodiments B95-B119, wherein the lipid nucleic acid assembly composition comprises a neutral lipid.

Embodiment B121. The method of any one of embodiments B95-B120, wherein the lipid nucleic acid assembly composition comprises a neutral lipid, wherein the neutral lipid is DSPC.

Embodiment B122. The method of any one of embodiments B95-B121, wherein the lipid nucleic acid assembly composition comprises a neutral lipid, wherein the neutral lipid is present at about 9 mol-%.

Embodiment B123. The method of any one of embodiments B95-B122, wherein the lipid nucleic acid assembly composition comprises a stealth lipid, wherein the stealth lipid is present at about 3 mol-%.

Embodiment B124. The method of any one of embodiments B95-B123, wherein the lipid nucleic acid assembly composition comprises a helper lipid, wherein the helper lipid is present at about 38 mol-%.

Embodiment B125. The method of any one of embodiments B95-B124, wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6.

Embodiment B126. The method of any one of embodiments B95-B125, wherein the lipid nucleic acid assembly composition comprises about 50 mol-% amine lipid such as Lipid A; about 9 mol-% neutral lipid such as DSPC; about 3 mol-% of stealth lipid such as a PEG lipid, such as PEG2k-DMG, and the remainder of the lipid component is helper lipid such as cholesterol wherein the N/P ratio of the is about 6.

Embodiment B127. The method of any one of embodiments B95-B126, comprising one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, and /or one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and/or one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

Embodiment B128. The method of any one of embodiments B95-B127, comprising at least two gRNAs selected from: one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

Embodiment B129. The method of any one of embodiments B95-B128, comprising one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

Embodiment B130. The method of any one of embodiments B95-129, comprising one gRNA selected from a gRNA that targets TRAC, TRBC, B2M, HLA-A, or CIITA.

Embodiment B131. The method of any one of embodiments B95-B130, wherein the gRNA targets TRBC, wherein the gRNA comprises a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

Embodiment B132. The method of any one of embodiments B95-B130, wherein the gRNA targets TRBC, wherein the gRNA comprises a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

Embodiment B133. The method of any one of embodiments B95-B130, comprising at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or B2M, wherein the two guide RNAs do not target the same gene.

Embodiment B134. The method of any one of embodiments B95-B130, comprising at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or HLA-A, wherein the two guide RNAs do not target the same gene.

Embodiment B135. The method of any one of embodiments B95-B130, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC.

Embodiment B136. The method of any one of embodiments B95-B130, comprising one guide RNA that targets B2M, and one gRNA that targets CIITA.

Embodiment B137. The method of any one of embodiments B95-B130, comprising one guide RNA that targets HLA-A, and one gRNA that targets CIITA, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

Embodiment B138. The method of any one of embodiments B95-B130, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA that targets B2M.

Embodiment B139. The method of any one of embodiments B95-B130, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA that targets HLA-A, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

Embodiment B140. The method of any one of embodiments B95-B130, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets B2M, and one gRNA that targets CIITA.

Embodiment B141. The method of any one of embodiments B95-B130, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets HLA-A, and one gRNA that targets CIITA, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

Embodiment B142. The method of any one of the preceding embodiments, wherein the method generates a cytosine (C) to thymine (T) conversion within a target sequence.

Embodiment B143. The method of any one of the preceding embodiments, wherein the method causes at least 60% C-to-T conversion relative to the total edits in the target sequence.

Embodiment B144. The method of any one of the preceding embodiments, wherein the method causes at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% C-to-T conversion relative to the total edits in the target sequence.

Embodiment B145. The method of any one of the preceding embodiments, wherein the ratio of C-to-T conversion to unintended edits is larger than 1:1.

Embodiment B146. The method of any one of the preceding embodiments, wherein the ratio of C-to-T conversion to unintended edits is from 2:1 to 99:1.

Embodiment B147. The method of any one of the preceding embodiments, wherein the ratio of C-to-T conversion to unintended edits is 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, or 8:1.

Embodiment B148. The method of any one of the preceding embodiments, wherein the method causes the cytidine deaminase to make a base edit corresponding to any one of positions -1 to 10 relative to the 5' end of the guide sequence.

Embodiment B149. The method of any one of the preceding embodiments, wherein the method causes the cytidine deaminase to make a base edit at a position 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the guide sequence.

Embodiment B150. The method of any one of the preceding embodiments, wherein the first mRNA, the second mRNA, and the guide RNA if present, delivered at a ratio of about 6:2:3 (w:w:w).

Embodiment B151. The method of any one of the preceding embodiments, wherein the mRNA, composition, or LNP is administered at a total RNA amount from 5 to 600 ng.

Embodiment B152. The method of any one of the preceding embodiments, wherein the total RNA amount is from 8 to 550 ng.

Embodiment B153. The method of any one of the preceding embodiments, wherein the total RNA amount is from 35 to 550 ng.

Embodiment B154. The method of any one of the preceding embodiments, wherein the total RNA amount is from 70 to 550 ng.

Embodiment B155. The method of any one of the preceding embodiments, wherein the total RNA amount is from 138 to 550 ng.

Embodiment B156. The method of any one of the preceding embodiments, wherein the total RNA amount is from 275 to 550 ng.

Embodiment B157. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a lymphocyte.

Embodiment B158. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a B lymphocyte.

Embodiment B159. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a T lymphocyte.

Embodiment B160. The method or use of any one of the preceding embodiments, wherein the modification of the target gene is *in vivo*.

Embodiment B161. The method or use of any one of the preceding embodiments, wherein the modification of the target gene is *ex vivo*.

Embodiment B162. The method or use of any one of the preceding embodiments, wherein the modification of the target gene reduces or eliminates expression of the target gene.

Embodiment B163. The method or use of any one of the preceding embodiments, wherein the genome editing or modification of the target gene reduces expression of the target gene by at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%.

Embodiment B164. The method or use of any one of the preceding embodiments, wherein the genome editing or modification of the target gene produces a missense mutation in the gene.

Embodiment B165. A polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

Embodiment B166. A ribonucleoprotein complex (RNP) comprising the polypeptide of embodiment B165 and a guide RNA, wherein if the RNP comprises a SpyCas9 nickase, then the guide RNA is a Spy guide RNA, and wherein if the RNP comprises a NmeCas nickase, then the guide RNA is a Nme guide RNA.

Embodiment B167. A composition comprising a first polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the first polypeptide does not comprise a uracil glycosylase inhibitor (UGI), and a second polypeptide comprising a UGI, wherein the second polypeptide is different from the first polypeptide.

Embodiment B168. The polypeptide, RNP or composition of any one of embodiments B165-B167, wherein the cytidine deaminase is fused to the RNA-guided nickase via a peptide linker, optionally XTEN.

Embodiment B169. The polypeptide, RNP or composition of any one of embodiments B165-B167, wherein the cytidine deaminase is attached to a linker comprising an organic molecule, polymer, or chemical moiety.

Embodiment B170. A pharmaceutical composition comprising the mRNA, RNP, composition, or polypeptide of any of the preceding embodiments and a pharmaceutically acceptable carrier.

Embodiment B171. A kit comprising the mRNA, RNP, composition, or polypeptide of any of the preceding embodiments.

Embodiment B172. The mRNA, RNP, composition, method, cell, or engineered cell of any one of the preceding embodiments, wherein the polypeptide comprising a cytidine deaminase and an RNA-guided nickase includes: the cytidine deaminase, a linker, and the RNA-guided nickase in amino to carboxy terminal order.

Embodiment B173. A method of altering a DNA sequence within a TRAC gene, comprising delivering to a cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

Embodiment B174. A method of reducing the expression of a TRAC gene, comprising delivering to a cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

Embodiment B175. A method of immunotherapy comprising administering a composition comprising an engineered cell to a subject,

wherein the cell comprises a genomic modification of at least one nucleotide within the genomic coordinates selected from:

chr14: 22547596-22547616; chr14: 22550570-22550590; chr14: 22547763-22547783; chr14: 22550596-22550616; chr14: 22550566-22550586; chr14: 22547753-22547773; chr14: 22550601-22550621; chr14: 22550599-22550619; chr14: 22547583-22547603; chr14: 22547671-22547691; chr14: 22547770-22547790; chr14: 22547676-22547696; chr14: 22547772-22547792; chr14: 22547771-22547791; chr14: 22547733-22547753; chr14: 22547776-22547796; or

wherein the cell is engineered by delivering to the cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

Embodiment B176. The method of any one of embodiments B173-B175, wherein the guide sequence comprises any one of SEQ ID NOs: 706-709.

Embodiment B177. The method of any one of embodiments B173-B175, wherein the guide sequence comprises any one of SEQ ID NOs: 706-708.

Embodiment B178. The method of any one of embodiments B173-B175, wherein the guide sequence comprises SEQ ID NO: 706.

Embodiment B179. The method of any one of embodiments B173-B175, wherein the guide sequence comprises SEQ ID NO: 707.

Embodiment B180. The method of any one of embodiments B173-B175, wherein the guide sequence comprises SEQ ID NO: 708.

Embodiment B181. A method of altering a DNA sequence within a TRBC1 and/or TRBC2 gene, comprising delivering a composition to a cell, wherein the composition comprises:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a guide RNA of (a.).

Embodiment B182. A method of reducing the expression of a TRBC1 and/or TRBC2 gene, comprising delivering a composition to a cell, wherein the composition comprises:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a guide RNA of (a.).

Embodiment B183. A method of immunotherapy comprising administering a composition comprising an engineered cell to a subject,

wherein the cell comprises a modification of at least one nucleotide within the genomic coordinates selected from: chr7: 142791757-142791777; chr7: 142801104-142801124; chr7: 142791811-142791831; chr7: 142801158-142801178; chr7: 142792728-142792748; chr7: 142791719-142791739; chr7: 142791766-142791786; chr7: 142801113-142801133; chr7: 142791928-142791948; chr7: 142801275-142801295; chr7: 142792062-142792082; chr7: 142801409-142801429; chr7: 142792713-142792733; chr7: 142802126-142802146; chr7: 142791808-142791828; chr7: 142801155-142801175; chr7: 142792003-142792023; chr7: 142801350-142801370; chr7: 142791760-142791780; chr7: 142791715-142791735; chr7: 142792781-142792801; chr7: 142792040-142792060; chr7: 142801387-142801407; chr7: 142791862-142791882; chr7: 142791716-142791736; chr7: 142791787-142791807; chr7: 142791759-142791779; chr7: 142801106-142801126; chr7: 142791807-142791827; chr7: 142801154-142801174; chr7: 142791879-142791899; chr7: 142801226-142801246; chr7: 142791805-142791825; chr7: 142791700-142791720; chr7: 142791765-142791785; chr7: 142801112-142801132; chr7: 142791820-142791840; chr7: 142791872-142791892; chr7: 142801219-142801239; chr7: 142791700-142791720; chr7: 142791806-142791826; chr7: 142801153-142801173; chr7: 142792035-142792055; chr7: 142792724-142792744; chr7: 142792754-142792774; chr7: 142791804-142791824; chr7: 142792684-142792704; chr7: 142791823-142791843; chr7: 142792728-142792748; chr7: 142792721-142792741; chr7: 142792749-142792769; chr7: 142792685-142792705; chr7: 142791816-142791836; chr7: 142801163-142801183; chr7: 142792686-142792706; chr7: 142791793-142791813; chr7: 142793110-142793130; chr7: 142791815-142791835; chr7: 142801162-142801182; chr7: 142792770-142792790; chr7: 142792047-

142792067; chr7: 142801394-142801414; chr7: 142791871-142791891; chr7: 142801218-142801238; chr7: 142791894-142791914; chr7: 142792723-142792743; chr7: 142792724-142792744; chr7: 142791897-142791917; chr7: 142801244-142801264; chr7: 142792757-142792777; chr7: 142792740-142792760; chr7: 142792758-142792778; or

wherein the cell is engineered by delivering to a cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a guide RNA of (a.).

Embodiment B184. The method of any one of embodiments B181-B183, wherein the guide sequence comprises any one of SEQ ID NOs: 618-627.

Embodiment B185. The method of any one of embodiments B181-B184, wherein the guide sequence comprises any one of SEQ ID NOs: 618-621.

Embodiment B186. The method of any one of embodiments B181-B185, wherein the guide sequence comprises SEQ ID NO: 618.

Embodiment B187. The method of any one of embodiments B181-B185, wherein the guide sequence comprises SEQ ID NO: 619.

Embodiment B188. The method of any one of embodiments B181-B185, wherein the guide sequence comprises SEQ ID NO: 620.

Embodiment B189. The method of any one of embodiments B181-B185, wherein the guide sequence comprises SEQ ID NO: 621.

Embodiment B190. The method of any one of embodiments B173-B189, wherein the composition further comprises the mRNA or composition of any one of the preceding embodiments relating to mRNA or compositions.

Embodiment B191. A composition comprising:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides

±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); and optionally

b. the mRNA or composition of any one of the preceding embodiments relating to mRNA or compositions.

Embodiment B192. The composition of embodiment B191, wherein the guide sequence comprises any one of SEQ ID NOs: 706-709.

Embodiment B193. The composition of embodiment B191 or B192, for use in altering a DNA sequence within the TRAC gene in a cell.

Embodiment B194. The composition of any one of embodiments B191-B193 for use in reducing the expression of the TRAC gene in a cell.

Embodiment B195. The composition of any one of embodiments B191-B194, for use in immunotherapy of a subject.

Embodiment B196. A composition comprising:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); and optionally

b. the mRNA or composition of any one of the preceding embodiments relating to mRNA or compositions.

Embodiment B197. The composition of embodiment B196, wherein the guide sequence comprises any one of SEQ ID NOs: 618-621.

Embodiment B198. The composition of embodiment B196 or B197, for use in altering a DNA sequence within the TRBC1 and/or TRBC2 gene in a cell.

Embodiment B199. The composition of any one of embodiments B196-B198 for use in reducing the expression of the TRBC1 and/or TRBC2 gene in a cell.

Embodiment B200. The composition of any one of embodiments B196-B199, for use in immunotherapy of a subject.

Embodiment B201. A cell, altered by the method of any one of embodiments B120-B121, and 161-178 .

Embodiment B202. The cell according to embodiment B201, wherein the cell is altered ex vivo.

Embodiment B203. The cell according to embodiment B201 or B202, wherein the cell is a T cell.

Embodiment B204. The cell according to any of embodiments B201-B203, wherein the cell is a $CD4^+$ or $CD8^+$ T cell.

Embodiment B205. The cell according to any of embodiments B201-B204, wherein the cell is a mammalian, primate, or human cell.

Embodiment B206. The cell according to any of embodiments B201-B205, for use in immunotherapy of a subject.

Embodiment B207. An engineered cell which has reduced or eliminated surface expression of TRAC, comprising a genetic modification in a human TRAC gene, wherein the genetic modification comprises a modification of at least one nucleotide within the genomic coordinates selected from:
chr14: 22547596-22547616; chr14: 22550570-22550590; chr14: 22547763-22547783; chr14: 22550596-22550616; chr14: 22550566-22550586; chr14: 22547753-22547773; chr14: 22550601-22550621; chr14: 22550599-22550619; chr14: 22547583-22547603; chr14: 22547671-22547691; chr14: 22547770-22547790; chr14: 22547676-22547696; chr14: 22547772-22547792; chr14: 22547771-22547791; chr14: 22547733-22547753; chr14: 22547776-22547796.

Embodiment B208. An engineered cell which has reduced or eliminated surface expression of TRBC1/2, comprising a genetic modification in a human TRBC1/2 gene, wherein the genetic modification comprises a modification of at least one nucleotide within the genomic coordinates selected from:
chr7: 142791757-142791777; chr7: 142801104-142801124; chr7: 142791811-142791831; chr7: 142801158-142801178; chr7: 142792728-142792748; chr7: 142791719-142791739; chr7: 142791766-142791786; chr7: 142801113-142801133; chr7: 142791928-142791948; chr7: 142801275-142801295; chr7: 142792062-142792082; chr7: 142801409-142801429; chr7: 142792713-142792733; chr7: 142802126-142802146; chr7: 142791808-142791828; chr7: 142801155-142801175; chr7: 142792003-142792023; chr7: 142801350-142801370; chr7: 142791760-142791780; chr7: 142791715-142791735; chr7: 142792781-142792801; chr7: 142792040-142792060; chr7: 142801387-142801407; chr7: 142791862-142791882; chr7: 142791716-142791736; chr7: 142791787-142791807; chr7: 142791759-

142791779; chr7: 142801106-142801126; chr7: 142791807-142791827; chr7: 142801154-142801174; chr7: 142791879-142791899; chr7: 142801226-142801246; chr7: 142791805-142791825; chr7: 142791700-142791720; chr7: 142791765-142791785; chr7: 142801112-142801132; chr7: 142791820-142791840; chr7: 142791872-142791892; chr7: 142801219-142801239; chr7: 142791700-142791720; chr7: 142791806-142791826; chr7: 142801153-142801173; chr7: 142792035-142792055; chr7: 142792724-142792744; chr7: 142792754-142792774; chr7: 142791804-142791824; chr7: 142792684-142792704; chr7: 142791823-142791843; chr7: 142792728-142792748; chr7: 142792721-142792741; chr7: 142792749-142792769; chr7: 142792685-142792705; chr7: 142791816-142791836; chr7: 142801163-142801183; chr7: 142792686-142792706; chr7: 142791793-142791813; chr7: 142793110-142793130; chr7: 142791815-142791835; chr7: 142801162-142801182; chr7: 142792770-142792790; chr7: 142792047-142792067; chr7: 142801394-142801414; chr7: 142791871-142791891; chr7: 142801218-142801238; chr7: 142791894-142791914; chr7: 142792723-142792743; chr7: 142792724-142792744; chr7: 142791897-142791917; chr7: 142801244-142801264; chr7: 142792757-142792777; chr7: 142792740-142792760; chr7: 142792758-142792778.

Embodiment B209. One or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising:

(a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase;

(b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs.

Embodiment B210. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is an immune cell.

Embodiment B211. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a lymphocyte.

Embodiment B212. The method, cell, or engineered cell of any one of the preceding embodiments, wherein the cell is a T cell.

We claim:

1. A composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, optionally wherein the composition comprises lipid nanoparticles.

2. The composition of claim 1, wherein the first open reading frame does not comprise a sequence encoding a UGI.

3. The composition of claim 1 or 2, wherein the composition comprises a first composition and a second composition, wherein the first composition comprises a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase and does not comprise a uracil glycosylase inhibitor (UGI), and the second composition comprises a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, optionally wherein the compositions comprise lipid nanoparticles.

4. The composition of any one of claims 1-3, wherein the first mRNA and the second mRNAs are in the same or separate vials.

5. A method of modifying a target gene comprising delivering to a cell a first mRNA comprising a first open reading frame encoding a first polypeptide comprising a cytidine deaminase and an RNA-guided nickase, a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA, and at least one guide RNA (gRNA).

6. The method of claim 5, wherein if the nickase is a SpyCas9 nickase, then the gRNA is a SpyCas9 gRNA, and if the nickase is a NmeCas9 nickase, then the gRNA is a Nme gRNA.

7. The method of claim 5 or 6, wherein the first open reading frame does not comprise a sequence encoding a UGI.

8. The composition or method of any one of claims 1-7, wherein the molar ratio of the second mRNA to the first mRNA is from 1:1 to 30:1.

9. The composition or method of any one of claims 1-7, wherein the molar ratio is from 2:1 to 30:1.

10. The composition or method of any one of claims 1-7, wherein the molar ratio is from 7:1 to 22:1.

11. An mRNA comprising an open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

12. A method of modifying at least one cytidine within a target gene in a cell, comprising expressing in the cell or contacting the cell with: (i) a first polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the first polypeptide does not comprise a uracil glycosylase inhibitor (UGI); (ii) a UGI polypeptide; and (iii) at least one guide RNA (gRNA) wherein the first polypeptide and gRNA form a complex with the target gene and modify the at least one cytidine in the target gene.

13. The method of claim 12, wherein if the nickase is a SpyCas9 nickase, then the gRNA is a SpyCas9 gRNA, and if the nickase is a NmeCas9 nickase, then the gRNA is a Nme gRNA.

14. The method of claim 12 or 13, wherein the ratio of the UGI polypeptide to the first polypeptide is from 10:1 to 50:1.

15. A cell, wherein the mRNA or composition of any one of claims 1-4 and 8-11 has been introduced to the cell, wherein the cell has been modified after the introduction.

16. An engineered cell altered by the method of claims 5-10 and 12-14.

17. An engineered cell comprising at least one base edit and/or indel, wherein the base edit and/or indel is made by contacting a cell with a composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI), wherein the second mRNA is different from the first mRNA.

18. The engineered cell of claim 17, wherein the first open reading frame does not comprise a sequence encoding a UGI.

19. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-18, wherein the cytidine deaminase is

(i) an enzyme of APOBEC family, optionally an enzyme of APOBEC3 subgroup;

(ii) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1023;

(iii) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, and 960-1013;

(iv) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009; or

(v) a cytidine deaminase comprising an amino acid sequence that is at least 80% identical to any one of SEQ ID NOs: 40, 976, 981, 984, 986, and 1014-1023.

20. The mRNA, composition, method, cell, or engineered cell of claim 19, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 41, and 960-1023.

21. The mRNA, composition, method, cell, or engineered cell of claim 19, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 41, and 960-1013.

22. The mRNA, composition, method, cell, or engineered cell of claim 19, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 41, 976, 977, 979, 980, 984-987, 993-1006, and 1009.

23. The mRNA, composition, method, cell, or engineered cell of claim 19, wherein the cytidine deaminase comprises an amino acid sequence with at least 80%, 85%, 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 40, 976, 981, 984, 986, 1014-1023.

24. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-23, wherein the cytidine deaminase is an APOBEC3A deaminase (A3A).

25. The mRNA, composition, method, cell, or engineered cell of claim 24, wherein the A3A comprises an amino acid sequence of SEQ ID NO: 40 or an amino acid sequence with at least 87%, at least 90%, at least 95%, at least 98%, at least 99% identity to SEQ ID NO: 40.

26. The mRNA, composition, method, cell, or engineered cell of any one of claims 24-25, wherein the A3A is a human A3A.

27. The mRNA, composition, method, cell, or engineered cell of any one of claims 24-26, wherein the A3A is a wild-type A3A.

28. The mRNA, composition, method, cell, or engineered cell of claim 24, wherein the A3A comprises an amino acid sequence with at least 87%, 90%, 95%, 98%, 99%, or 100% identity to SEQ ID NO: 976, 977, 993-1006, and 1009.

29. The composition, method, cell, or engineered cell of any one of claims 1-28, wherein the UGI comprises an amino acid sequence of SEQ ID NO: 27 or an amino acid sequence with at least 80%, at least 90%, at least 95%, at least 98%, at least 99% identity to SEQ ID NO: 27.

30. The composition, method, or cell of any one of claims 1-4, 8-10, and 15-29, further comprising at least one guide RNA (gRNA).

31. The composition, method, cell, or engineered cell of any one of claims 1-4, 8-10, and 15-30, comprising a gRNA, wherein the gRNA is an sgRNA.

32. The composition, method, cell, or engineered cell of any one of claims 1-4, 8-10, and 15-31, comprising a gRNA, wherein the gRNA is a short-single guide RNA (short-sgRNA) comprising a conserved portion of an sgRNA comprising a hairpin region, wherein the hairpin region lacks at least 5-10 nucleotides and wherein the short-sgRNA comprises a 5' end modification or a 3' end modification or both.

33. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-32, wherein the RNA-guided nickase is a Cas9 nickase.

34. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-33, wherein the RNA-guided nickase is an *S. pyogenes* (Spy) Cas9 nickase.

35. The mRNA, composition, method, cell, or engineered cell of claim 34, wherein the RNA-guided nickase is a D10A SpyCas9 nickase.

36. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-35, wherein the RNA-guided nickase comprises an amino acid sequence of any one of SEQ ID NOs: 70, 73, or 76 or an amino acid sequence having at least 80%, 90%, 95%, 98%, or 99% identity to any one of SEQ ID NOs: 70, 73, or 76.

37. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-36, wherein the sequence encoding the RNA-guided nickase comprises a nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78 or a nucleotide sequence having at least 80%, 90%, 95%, 98%, or 99% identity to the nucleotide sequence of any one of SEQ ID NOs: 72, 75, or 78.

38. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-37, wherein the sequence encoding the RNA-guided nickase comprises the nucleotide sequence of any one of SEQ ID NOs: 71, 72, 74, 75, or 77-90.

39. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-33, wherein the RNA-guided nickase is a *N. meningitidis* (Nme) Cas9 nickase.

40. The mRNA, composition, method, cell, or engineered cell of claim 39, wherein the RNA-guided nickase is a D16A NmeCas9 nickase, optionally a D16A Nme2Cas9.

41. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-33, 39, or 40, wherein the sequence encoding the RNA-guided nickase comprises the nucleotide sequence of any one of SEQ ID NOs: 380 and 387.

42. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-41, wherein the mRNA comprises a 5' UTR with at least 90% identity to any one of SEQ ID NOs: 91-98.

43. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-42, wherein the mRNA comprises a 3' UTR with at least 90% identity to any one of SEQ ID NOs: 99-106.

44. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-43, wherein the mRNA further comprises a 5' cap selected from Cap0, Cap1, Cap2, and a cap added co-transcriptionally or post-transcriptionally, optionally wherein the co-transcriptionally added cap is selected from anti-reverse cap analog (ARCA), AG (m7G(5')ppp(5')(2'OMeA)pG, or GG (m7G(5')ppp(5')(2'OMeG)pG, a cap added post-transcriptionally.

45. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-44, wherein the mRNA further comprises a poly-adenylated (poly-A) tail, optionally wherein the poly-A tail is added to the mRNA by PCR tailing or enzymatic tailing and optionally wherein the poly-A tail comprises a sequence of SEQ ID NO: 109.

46. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-45, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, and/or the open reading frame encoding a uracil glycosylase inhibitor (UGI) comprise (i) minimal adenine codons and/or minimal uridine codons; (ii) minimal adenine codons; (iii) codons that increase translation of the mRNA in a mammal; or (iv) codons that increase translation of the mRNA in a mammal, wherein the mammal is a human.

47. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-46, wherein the cytidine deaminase is located N-terminal to the RNA-guided nickase in the polypeptide.

48. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-47, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS).

49. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-48, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is at the C-terminus of the RNA-guided nickase.

50. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-49, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and

wherein the NLS is at the N-terminus of the RNA-guided nickase, or wherein an NLS is fused to both the N-terminus and C-terminus of the RNA-guided nickase.

51. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-50, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein a linker is present between the N-terminus of the RNA-guided nickase and the NLS, optionally wherein the linker is a peptide linker.

52. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-51, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS comprises a sequence having at least 80%, 85%, 90%, or 95% identity to any one of SEQ ID NOs: 63 and 110-122.

53. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-52, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS comprises the sequence of any one of SEQ ID NOs: 63 and 110-122.

54. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-53, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the NLS is encoded by a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of any one of SEQ ID NOs: 123-135.

55. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-54, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the cytidine deaminase is located N-terminal to the NLS in the polypeptide.

56. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-55, wherein the encoded RNA-guided nickase comprises a nuclear localization signal (NLS), and wherein the RNA-guided nickase is located N-terminal to the NLS in the polypeptide.

57. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-56, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of SEQ ID NO:1.

58. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-57, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98% or 100% identity to the sequence of SEQ ID NO: 4.

59. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-58, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and

an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98%, 99% or 100% identity to the sequence of SEQ ID NO: 321.

60. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-59, wherein the open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase comprises a sequence having at least 80%, 85%, 90%, 95%, 98%, 99% or 100% identity to the sequence of SEQ ID NO: 313.

61. The mRNA, composition, method, cell, or engineered cell of any one of claims 1-60, wherein at least 10% of the uridine in the mRNA is substituted with a modified uridine.

62. The mRNA, composition, method, cell or engineered cell of claim 61, wherein the modified uridine is one or more of N1-methyl-pseudouridine, pseudouridine, 5-methoxyuridine, or 5-iodouridine.

63. The mRNA, composition, method, cell, or engineered cell of any one of claims 61-62, wherein 15% to 45% of the uridine is substituted with the modified uridine.

64. The mRNA, composition, method, cell, or engineered cell of any one of claims 61-63, wherein at least 20% or at least 30%, at least 80% or at least 90%, or 100% of the uridine is substituted with the modified uridine.

65. The mRNA, composition, method, or cell of any one of the preceding claims 61-64, further encoding a peptide linker between the cytidine deaminase and RNA-guided nickase, optionally wherein the peptide linker is XTEN or the peptide linker comprises a sequence of GTKDSTKDIPETPSKD (SEQ ID NO: 268).

66. The mRNA, composition, method, or cell of any one of claims 61-65, further encoding a peptide linker between the cytidine deaminase and RNA-guided nickase, wherein the peptide linker comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, or more amino acids.

67. The mRNA, composition, method, or cell of any one of claims 61-66, further encoding a peptide linker between the cytidine deaminase and RNA-guided nickase, wherein the peptide linker comprises one or more sequences selected from SEQ ID NOs: 46-59, 61 and 211-272.

68. A polypeptide encoded by any one of the mRNAs of any one of claims 1-28 and 33-67.

69. A ribonucleoprotein complex (RNP) comprising (i) a polypeptide encoded by any one of the mRNAs of any one of claims 1-28 and 33-67; and (ii) a guide RNA.

70. A vector comprising any one of the mRNAs of any one of claims 1-28 and 33-67.

71. An expression construct comprising a promoter operably linked to a sequence encoding any one of the mRNAs of any one of claims 1-28, 33-60, and 65-67.

72. A plasmid comprising the expression construct of claim 71.

73. A host cell comprising the vector of claim 70, the expression construct of claim 71, or the plasmid of claim 72.

74. The mRNA or composition of any one of claims 1-4, 8-11, and 19-67, wherein the mRNA or composition is formulated as a lipid nucleic acid assembly composition, optionally a lipid nanoparticle.

75. Use of the mRNA or composition according to any one of claims 1-4, 8-11, and 19-67 for modifying a target gene in a cell.

76. Use of the mRNA or composition according to any one of claims 1-4, 8-11, and 19-67 for the manufacture of a medicament for modifying a target gene in a cell.

77. A method of modifying a target gene in a cell, comprising delivering to the cell one or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising:

(a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase;

(b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs.

78. The method of claim 77, wherein parts (a), (b), and (c) are each in separate lipid nucleic acid assembly compositions.

79. The method of claim 77, wherein parts (a), (b), and (c) are in the same lipid nucleic acid assembly composition.

80. The method of any one of claims 77-79, wherein the one or more guide RNAs are each in separate lipid nucleic acid assembly compositions.

81. The method of any one of claims 77-80, comprising delivering to the cell a lipid nucleic acid assembly composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase and a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI) in the same lipid nucleic acid assembly composition.

82. The method of any one of claims 77-81, comprising delivering to the cell a first lipid nucleic acid assembly composition comprising a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase,

and a second lipid nucleic acid assembly composition comprising a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI).

83. The method of any one of claims 77-82, further comprising delivering one or more guide RNAs in one or more lipid nucleic acid assembly compositions that are separate from the lipid nucleic acid assembly compositions comprising the cytidine deaminase and UGI.

84. The method of claims any one of claims 77-83, wherein at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 lipid nucleic acid assembly compositions are delivered to the cell.

85. The method of any one of claims 77-84, wherein at least one lipid nucleic acid assembly composition comprises lipid nanoparticle (LNPs), optionally wherein all lipid nucleic acid assembly compositions comprise LNPs.

86. The method of any one of claims 77-85, wherein at least one lipid nucleic acid assembly composition is a lipoplex composition.

87. The method of any one of claims 77-86, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid.

88. The method of any one of claims 77-87, wherein the lipid nucleic acid assembly composition comprises an ionizable lipid and wherein the ionizable lipid has a pKa in the range of from about 5.1 to about 7.4, such as from about 5.5 to about 6.6, from about 5.6 to about 6.4, from about 5.8 to about 6.2, or from about 5.8 to about 6.5.

89. The method of any one of claims 77-88, wherein the lipid nucleic acid assembly composition comprises (i) an amine lipid; (ii) a helper lipid; (iii) a stealth lipid; (iv) a neutral lipid; or combinations of one or more of (i)-(iv).

90. The method of claim 89, wherein (i) the amine lipid is Lipid A; (ii) the helper lipid is cholesterol; (iii) the stealth lipid is PEG2k-DMG; (iv) the neutral lipid is DSPC; or combinations of one or more of (i)-(iv).

91. The method of any one of claims 77-90, wherein the N/P ratio of the lipid nucleic acid assembly composition is about 6.

92. The method of any one of claims 77-91, wherein the lipid nucleic acid assembly composition comprises about 50 mol-% amine lipid such as Lipid A; about 9 mol-% neutral lipid such as DSPC; about 3 mol-% of stealth lipid such as a PEG lipid, such as PEG2k-DMG, and the remainder of the lipid component is helper lipid such as cholesterol wherein the N/P ratio of the is about 6.

93. The method of any one of claims 77-92, wherein the lipid nucleic acid assembly composition comprises about 35 mol-% amine lipid such as Lipid A; about 15 mol-% neutral lipid such as DSPC; about 2.5 mol-% of stealth lipid such as a PEG lipid, such as PEG2k-

DMG, and the remainder of the lipid component is helper lipid such as cholesterol wherein the N/P ratio of the is about 6.

94. The method of any one of claims 77-93, comprising one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, and /or one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and/or one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

95. The method of any one of claims 77-94, comprising at least two gRNAs selected from: one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

96. The method of any one of claims 77-95, comprising one gRNA that targets a gene that reduces or eliminates MHC class I expression on the surface of a cell, one gRNA that targets a gene that reduces or eliminates MHC class II expression on the surface of a cell, and one gRNA that targets a gene that reduces or eliminates endogenous TCR expression.

97. The method of any one of claims 77-96, comprising one gRNA selected from a gRNA that targets TRAC, TRBC, B2M, HLA-A, or CIITA.

98. The method of any one of claims 77-97, wherein the gRNA targets TRBC, wherein the gRNA comprises a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

99. The method of any one of claims 77-97, wherein the gRNA targets TRBC, wherein the gRNA comprises a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v).

100. The method of any one of claims 77-97, comprising at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or B2M, wherein the two guide RNAs do not target the same gene.

101. The method of any one of claims 77-97, comprising at least two gRNAs selected from a gRNA that targets TRAC, TRBC, or HLA-A, wherein the two guide RNAs do not target the same gene.

102. The method of any one of claims 77-97, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC.

103. The method of any one of claims 77-97, comprising one guide RNA that targets B2M, and one gRNA that targets CIITA.

104. The method of any one of claims 77-97, comprising one guide RNA that targets HLA-A, and one gRNA that targets CIITA, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

105. The method of any one of claims 77-97, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA that targets B2M.

106. The method of any one of claims 77-97, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, and one gRNA that targets HLA-A, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

107. The method of any one of claims 77-97, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets B2M, and one gRNA that targets CIITA.

108. The method of any one of claims 77-97, comprising one guide RNA that targets TRAC, and one gRNA that targets TRBC, one gRNA that targets HLA-A, and one gRNA that targets CIITA, optionally wherein the cell is homozygous for HLA-B and homozygous for HLA-C.

109. The method of any one of claims 5-10, 12-14, 19-67, and 77-108, wherein the method generates a cytosine (C) to thymine (T) conversion when present within a target sequence, optionally wherein if the nickase is a SpyCas9 nickase, the C to T conversion comprises 1-12 C to T conversions, and if the nickase is a NmeCas9 nickase, the C to T conversion comprises 1-20 C to T conversions.

110. The method of any one of claims 5-10, 12-14, 19-67, and 77-109, wherein the method causes at least 60% C-to-T conversion relative to the total edits in the target sequence.

111. The method of any one of claims 5-10, 12-14, 19-67, and 77-110, wherein the method causes at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% C-to-T conversion relative to the total edits in the target sequence.

112. The method of any one of claims 5-10, 12-14, 19-67, and 77-111, wherein the ratio of C-to-T conversion to unintended edits is larger than 1:1.

113. The method of any one of claims 5-10, 12-14, 19-67, and 77-112, wherein the ratio of C-to-T conversion to unintended edits is from 2:1 to 99:1.

114. The method of any one of claims 5-10, 12-14, 19-67, and 77-113, wherein the ratio of C-to-T conversion to unintended edits is 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, or 8:1.

115. The method of any one of claims 5-10, 12-14, 19-67, and 77-114, wherein the method causes the cytidine deaminase to make a base edit corresponding to any one of positions -1 to 10 relative to the 5' end of the guide sequence.

116. The method of any one of claims 5-10, 12-14, 19-67, and 77-115, wherein the method causes the cytidine deaminase to make a base edit at a cytidine present at position 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the guide sequence.

117. The method of any one of claims 5-10, 12-14, 19-67, and 77-116, wherein the nickase is a SpyCas9 nickase, and the method causes the cytidine deaminase to make a base edit at a cytidine present at position 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 nucleotides from the 5' end of the guide sequence.

118. The method of any one of claims 5-10, 12-14, 19-67, and 77-116, wherein the nickase is a NmeCas9 nickase, and the method causes the cytidine deaminase to make a base edit at a cytidine present at position 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides from the 5' end of the guide sequence.

119. The method of any one of claims 5-10, 12-14, 19-67, and 77-118, wherein the first mRNA, the second mRNA, and the guide RNA if present, delivered at a ratio of about 6:2:3 (w:w:w).

120. The method, cell, or engineered cell of any one of claims 5-10, 12-67, and 77-119, wherein the cell is a lymphocyte.

121. The method or use of any one of claims 5-10, 12-14, 19-67, and 75-120, wherein the modification of the target gene is *in vivo*.

122. The method or use of any one of claims 5-10, 12-14, 19-67, and 75-120, wherein the modification of the target gene is *ex vivo*.

123. The method or use of any one of claims 5-10, 12-14, 19-67, and 75-122, wherein the modification of the target gene reduces or eliminates expression of the target gene.

124. The method or use of any one of claims 5-10, 12-14, 19-67, and 75-123, wherein the genome editing or modification of the target gene reduces expression of the target gene by at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%.

125. The method or use of any one of claims 5-10, 12-14, 19-67, and 75-124, wherein the genome editing or modification of the target gene produces a missense mutation in the gene.

126. A polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

127. A ribonucleoprotein complex (RNP) comprising the polypeptide of claim 126 and a guide RNA, wherein if the RNP comprises a SpyCas9 nickase, then the guide RNA is a Spy guide RNA, and wherein if the RNP comprises a NmeCas nickase, then the guide RNA is a Nme guide RNA.

128. A composition comprising a first polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the first polypeptide does not comprise a uracil glycosylase inhibitor (UGI), and a second polypeptide comprising a UGI, wherein the second polypeptide is different from the first polypeptide.

129. The polypeptide, RNP, or composition of any one of claims 126-128, wherein the cytidine deaminase is fused to the RNA-guided nickase via a peptide linker, optionally XTEN or a peptide linker comprising a sequence of GTKDSTKDIPETPSKD (SEQ ID NO: 268).

130. The polypeptide, RNP, or composition of any one of claims 126-128, wherein the cytidine deaminase is attached to a linker comprising an organic molecule, polymer, or chemical moiety.

131. A pharmaceutical composition comprising the mRNA, RNP, composition, or polypeptide of claims 1-4, 8-11, 19-69, 74, and 126-130 and a pharmaceutically acceptable carrier.

132. A kit comprising the mRNA, RNP, composition, or polypeptide of any of 1-4, 8-11, 19-69, 74, and 126-130.

133. The mRNA, RNP, composition, method, use, cell, or engineered cell of any one of 1-132, wherein the polypeptide comprising a cytidine deaminase and an RNA-guided nickase includes: the cytidine deaminase, a linker, and the RNA-guided nickase in amino to carboxy terminal order.

134. A method of altering a DNA sequence within a TRAC gene, comprising delivering to a cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ

ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

135. A method of reducing the expression of a TRAC gene, comprising delivering to a cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

136. A method of immunotherapy comprising administering a composition comprising an engineered cell to a subject,

wherein the cell comprises a genomic modification of at least one nucleotide within the genomic coordinates selected from:

chr14: 22547596-22547616; chr14: 22550570-22550590; chr14: 22547763-22547783; chr14: 22550596-22550616; chr14: 22550566-22550586; chr14: 22547753-22547773; chr14: 22550601-22550621; chr14: 22550599-22550619; chr14: 22547583-22547603; chr14: 22547671-22547691; chr14: 22547770-22547790; chr14: 22547676-22547696; chr14: 22547772-22547792; chr14: 22547771-22547791; chr14: 22547733-22547753; chr14: 22547776-22547796; or

wherein the cell is engineered by delivering to the cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v)

at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a gRNA of (a.).

137. A method of altering a DNA sequence within a TRBC1 and/or TRBC2 gene, comprising delivering a composition to a cell, wherein the composition comprises:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a guide RNA of (a.).

138. A method of reducing the expression of a TRBC1 and/or TRBC2 gene, comprising delivering a composition to a cell, wherein the composition comprises:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a guide RNA of (a.).

139. A method of immunotherapy comprising administering a composition comprising an engineered cell to a subject,

wherein the cell comprises a modification of at least one nucleotide within the genomic coordinates selected from: chr7: 142791757-142791777; chr7: 142801104-142801124; chr7: 142791811-142791831; chr7: 142801158-142801178; chr7: 142792728-142792748; chr7: 142791719-142791739; chr7: 142791766-142791786; chr7: 142801113-142801133; chr7: 142791928-142791948; chr7: 142801275-142801295; chr7: 142792062-142792082; chr7: 142801409-142801429; chr7: 142792713-142792733; chr7: 142802126-142802146; chr7:

142791808-142791828; chr7: 142801155-142801175; chr7: 142792003-142792023; chr7: 142801350-142801370; chr7: 142791760-142791780; chr7: 142791715-142791735; chr7: 142792781-142792801; chr7: 142792040-142792060; chr7: 142801387-142801407; chr7: 142791862-142791882; chr7: 142791716-142791736; chr7: 142791787-142791807; chr7: 142791759-142791779; chr7: 142801106-142801126; chr7: 142791807-142791827; chr7: 142801154-142801174; chr7: 142791879-142791899; chr7: 142801226-142801246; chr7: 142791805-142791825; chr7: 142791700-142791720; chr7: 142791765-142791785; chr7: 142801112-142801132; chr7: 142791820-142791840; chr7: 142791872-142791892; chr7: 142801219-142801239; chr7: 142791700-142791720; chr7: 142791806-142791826; chr7: 142801153-142801173; chr7: 142792035-142792055; chr7: 142792724-142792744; chr7: 142792754-142792774; chr7: 142791804-142791824; chr7: 142792684-142792704; chr7: 142791823-142791843; chr7: 142792728-142792748; chr7: 142792721-142792741; chr7: 142792749-142792769; chr7: 142792685-142792705; chr7: 142791816-142791836; chr7: 142801163-142801183; chr7: 142792686-142792706; chr7: 142791793-142791813; chr7: 142793110-142793130; chr7: 142791815-142791835; chr7: 142801162-142801182; chr7: 142792770-142792790; chr7: 142792047-142792067; chr7: 142801394-142801414; chr7: 142791871-142791891; chr7: 142801218-142801238; chr7: 142791894-142791914; chr7: 142792723-142792743; chr7: 142792724-142792744; chr7: 142791897-142791917; chr7: 142801244-142801264; chr7: 142792757-142792777; chr7: 142792740-142792760; chr7: 142792758-142792778; or

wherein the cell is engineered by delivering to a cell:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); or

b. a nucleic acid encoding a guide RNA of (a.).

140. A composition comprising:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 706-721; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 706-721; iii) a guide sequence at least 95%, 90%, or 85% identical to a

sequence selected from SEQ ID NOs: 706-721; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5A; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); and optionally

b. the mRNA or composition of any one of the preceding claims relating to mRNA or compositions.

141. A composition comprising:

a. a gRNA comprising a guide sequence chosen from: i) SEQ ID NOs: 618-669; ii) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence selected from SEQ ID NOs: 618-669; iii) a guide sequence at least 95%, 90%, or 85% identical to a sequence selected from SEQ ID NOs: 618-669; iv) a sequence that comprises 10 contiguous nucleotides ±10 nucleotides of a genomic coordinate listed in Table 5B; v) at least 17, 18, 19, or 20 contiguous nucleotides of a sequence from (iv); or vi) a guide sequence that is at least 95%, 90%, or 85% identical to a sequence selected from (v); and optionally

b. the mRNA or composition of any one of the preceding claims relating to mRNA or compositions.

142. An engineered cell which has reduced or eliminated surface expression of TRAC, comprising a genetic modification in a human TRAC gene, wherein the genetic modification comprises a modification of at least one nucleotide within the genomic coordinates selected from:

chr14: 22547596-22547616; chr14: 22550570-22550590; chr14: 22547763-22547783; chr14: 22550596-22550616; chr14: 22550566-22550586; chr14: 22547753-22547773; chr14: 22550601-22550621; chr14: 22550599-22550619; chr14: 22547583-22547603; chr14: 22547671-22547691; chr14: 22547770-22547790; chr14: 22547676-22547696; chr14: 22547772-22547792; chr14: 22547771-22547791; chr14: 22547733-22547753; chr14: 22547776-22547796.

143. An engineered cell which has reduced or eliminated surface expression of TRBC1/2, comprising a genetic modification in a human TRBC1/2 gene, wherein the genetic modification comprises a modification of at least one nucleotide within the genomic coordinates selected from:

chr7: 142791757-142791777; chr7: 142801104-142801124; chr7: 142791811-142791831; chr7: 142801158-142801178; chr7: 142792728-142792748; chr7: 142791719-142791739;

chr7: 142791766-142791786; chr7: 142801113-142801133; chr7: 142791928-142791948; chr7: 142801275-142801295; chr7: 142792062-142792082; chr7: 142801409-142801429; chr7: 142792713-142792733; chr7: 142802126-142802146; chr7: 142791808-142791828; chr7: 142801155-142801175; chr7: 142792003-142792023; chr7: 142801350-142801370; chr7: 142791760-142791780; chr7: 142791715-142791735; chr7: 142792781-142792801; chr7: 142792040-142792060; chr7: 142801387-142801407; chr7: 142791862-142791882; chr7: 142791716-142791736; chr7: 142791787-142791807; chr7: 142791759-142791779; chr7: 142801106-142801126; chr7: 142791807-142791827; chr7: 142801154-142801174; chr7: 142791879-142791899; chr7: 142801226-142801246; chr7: 142791805-142791825; chr7: 142791700-142791720; chr7: 142791765-142791785; chr7: 142801112-142801132; chr7: 142791820-142791840; chr7: 142791872-142791892; chr7: 142801219-142801239; chr7: 142791700-142791720; chr7: 142791806-142791826; chr7: 142801153-142801173; chr7: 142792035-142792055; chr7: 142792724-142792744; chr7: 142792754-142792774; chr7: 142791804-142791824; chr7: 142792684-142792704; chr7: 142791823-142791843; chr7: 142792728-142792748; chr7: 142792721-142792741; chr7: 142792749-142792769; chr7: 142792685-142792705; chr7: 142791816-142791836; chr7: 142801163-142801183; chr7: 142792686-142792706; chr7: 142791793-142791813; chr7: 142793110-142793130; chr7: 142791815-142791835; chr7: 142801162-142801182; chr7: 142792770-142792790; chr7: 142792047-142792067; chr7: 142801394-142801414; chr7: 142791871-142791891; chr7: 142801218-142801238; chr7: 142791894-142791914; chr7: 142792723-142792743; chr7: 142792724-142792744; chr7: 142791897-142791917; chr7: 142801244-142801264; chr7: 142792757-142792777; chr7: 142792740-142792760; chr7: 142792758-142792778.

144. A lipid nucleic acid assembly composition comprising an mRNA comprising an open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase, wherein the polypeptide does not comprise a uracil glycosylase inhibitor (UGI).

145. One or more lipid nucleic acid assembly compositions, optionally lipid nanoparticles, comprising:

(a) a first mRNA comprising a first open reading frame encoding a polypeptide comprising a cytidine deaminase and an RNA-guided nickase;

(b) a second mRNA comprising a second open reading frame encoding a uracil glycosylase inhibitor (UGI); and

(c) one or more guide RNAs.

146. The method, cell, or engineered cell of any one of claims 134-143, wherein the cell is an immune cell, a lymphocyte, or a T cell.

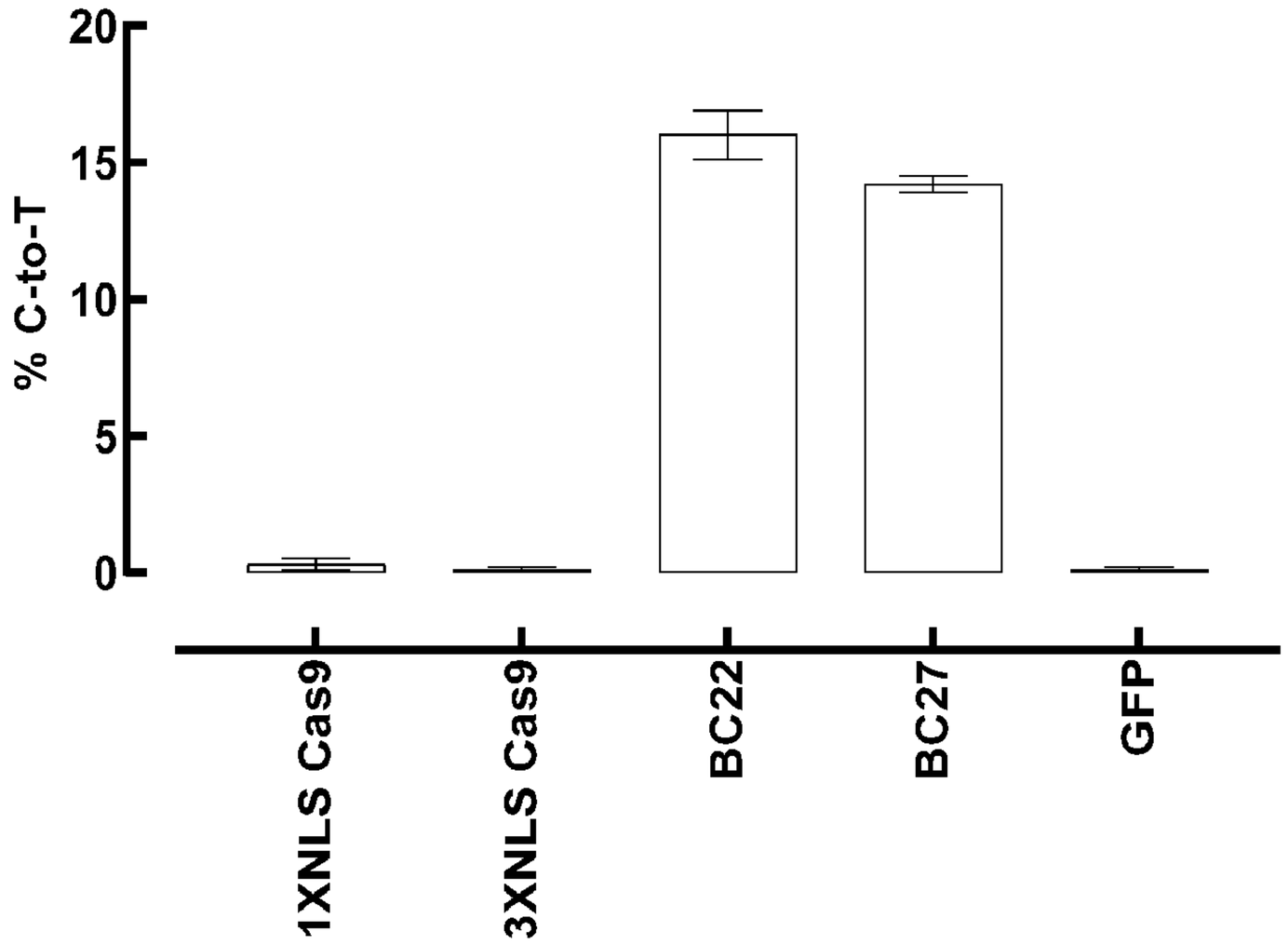

sg000296
% C-to-T
0
5
10
15
20
1XNLS Cas9
3XNLS Cas9
BC22
BC27
GFP

FIG. 1A

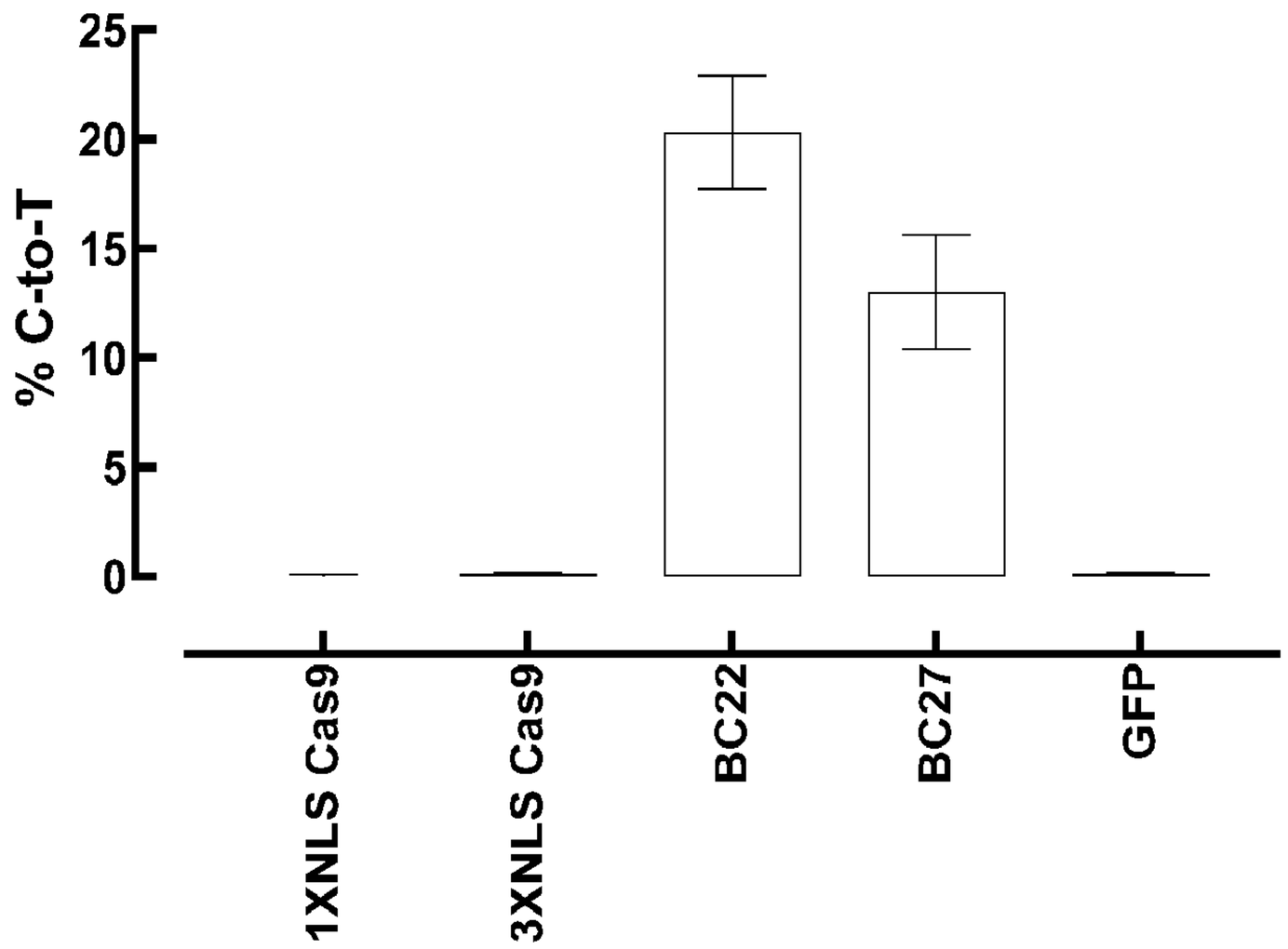

sg001373
% C-to-T
0
5
10
15
20
25
1XNLS Cas9
3XNLS Cas9
BC22
BC27
GFP

FIG. 1B

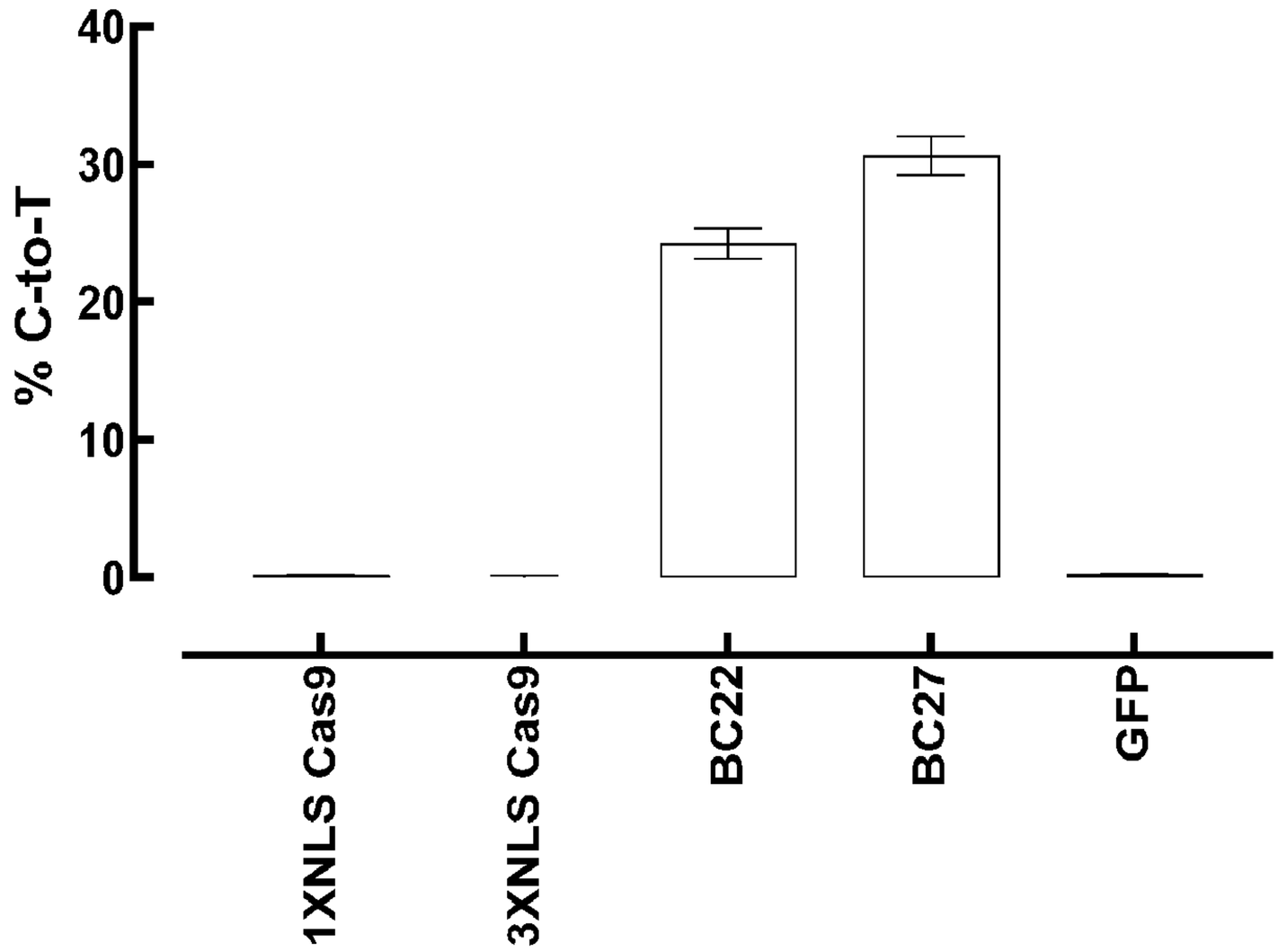

sg001400
% C-to-T
40
30
20
10
0
1XNLS Cas9
3XNLS Cas9
BC22
BC27
GFP

FIG. 1C

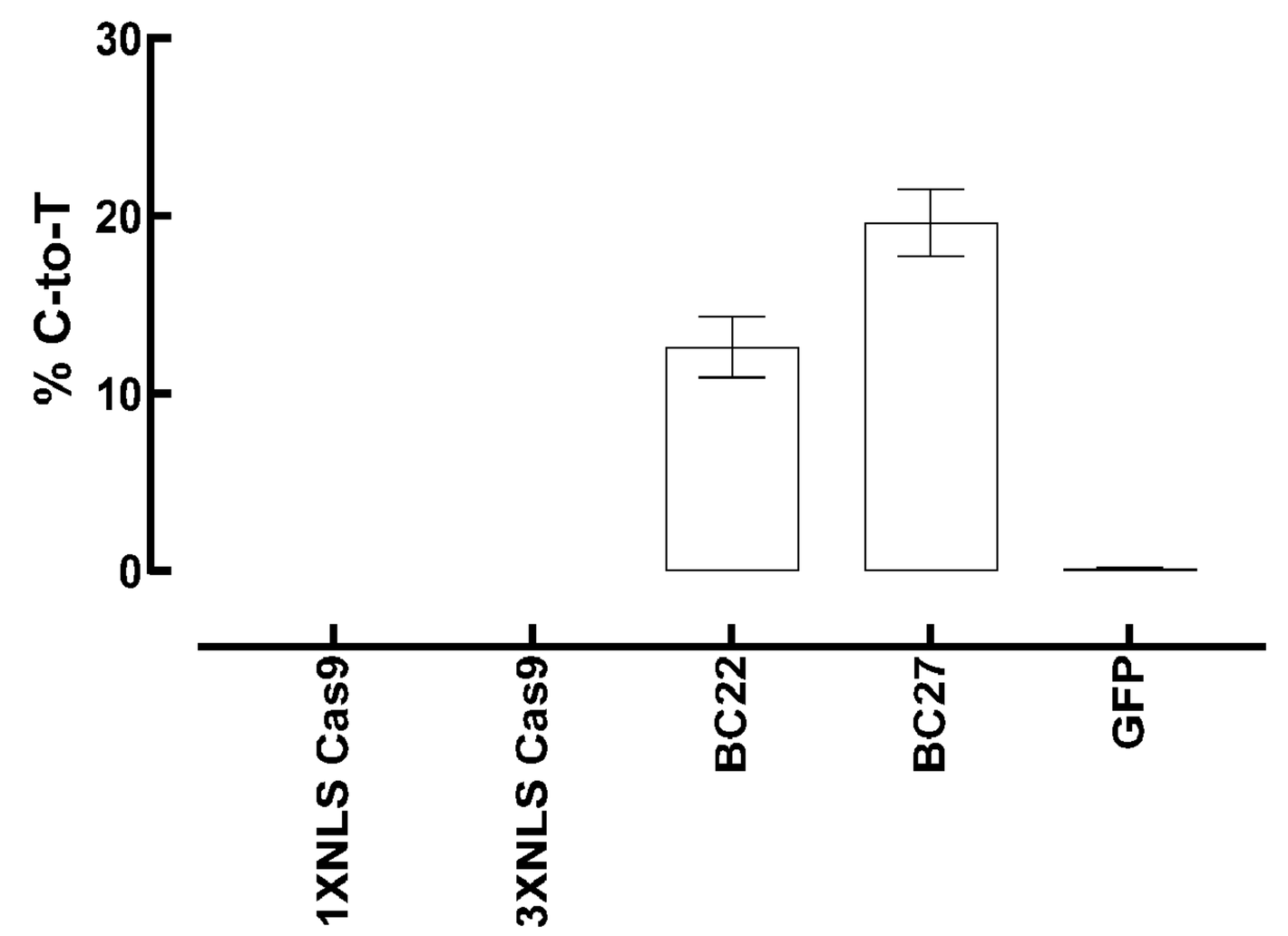

sg003018
% C-to-T
30
20
10
0
1XNLS Cas9
3XNLS Cas9
BC22
BC27
GFP

FIG. 1D

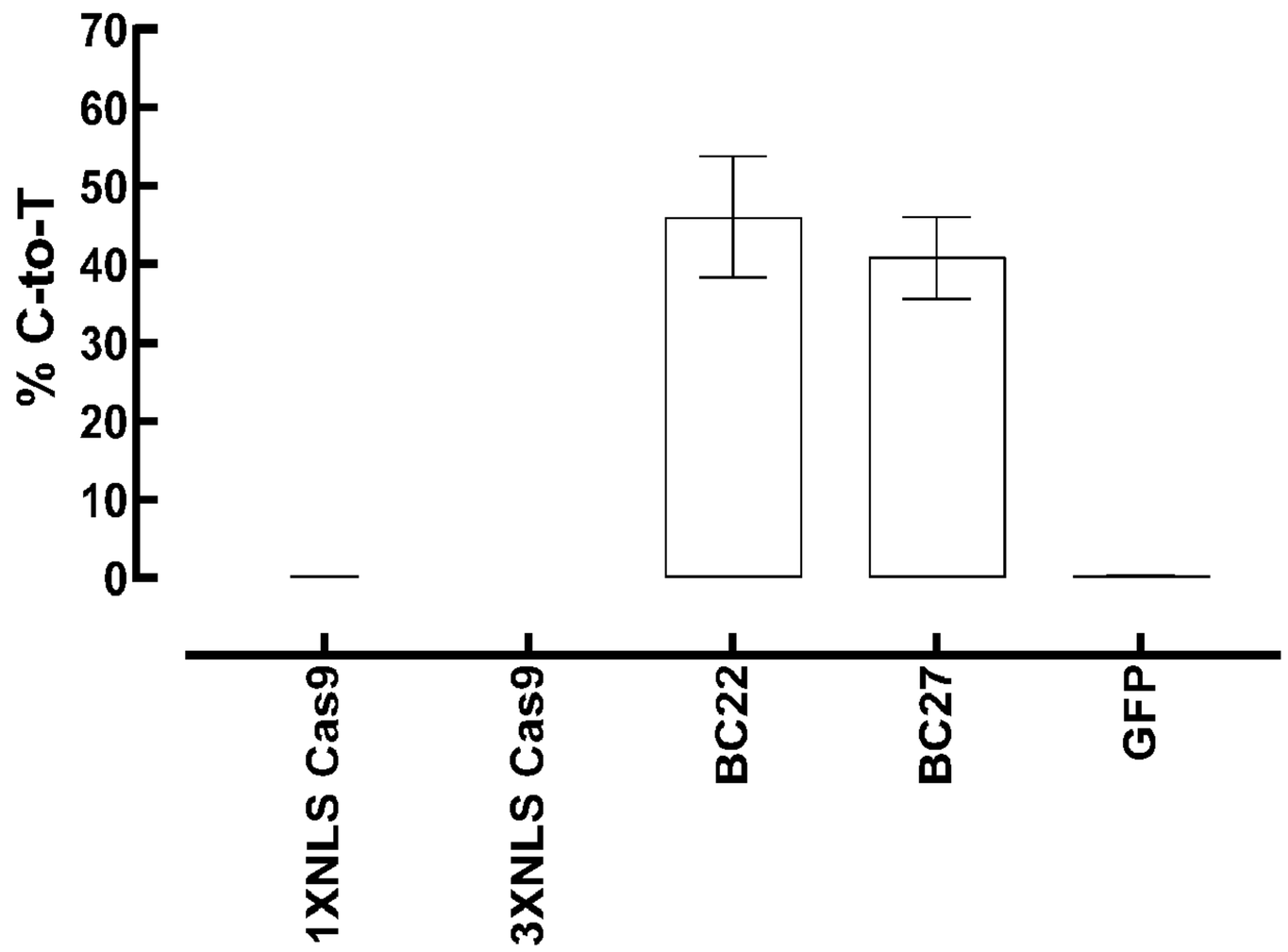

sg005883
% C-to-T
70
60
50
40
30
20
10
0
1XNLS Cas9
3XNLS Cas9
BC22
BC27
GFP

FIG. 1E

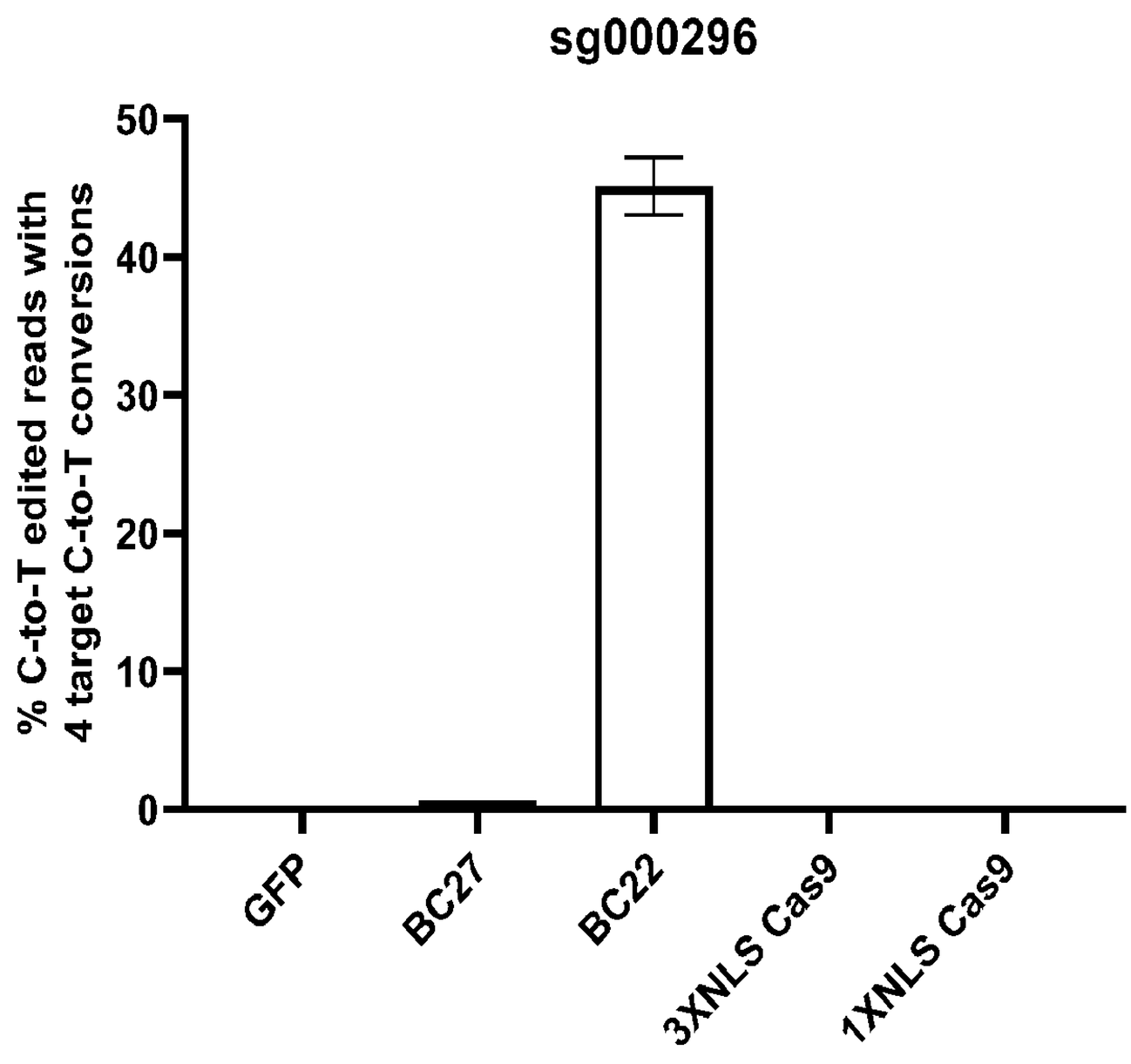

sg000296
% C-to-T edited reads with
4 target C-to-T conversions
50
40
30
20
10
0
GFP
BC27
BC22
3XNLS Cas9
1XNLS Cas9

FIG. 2A

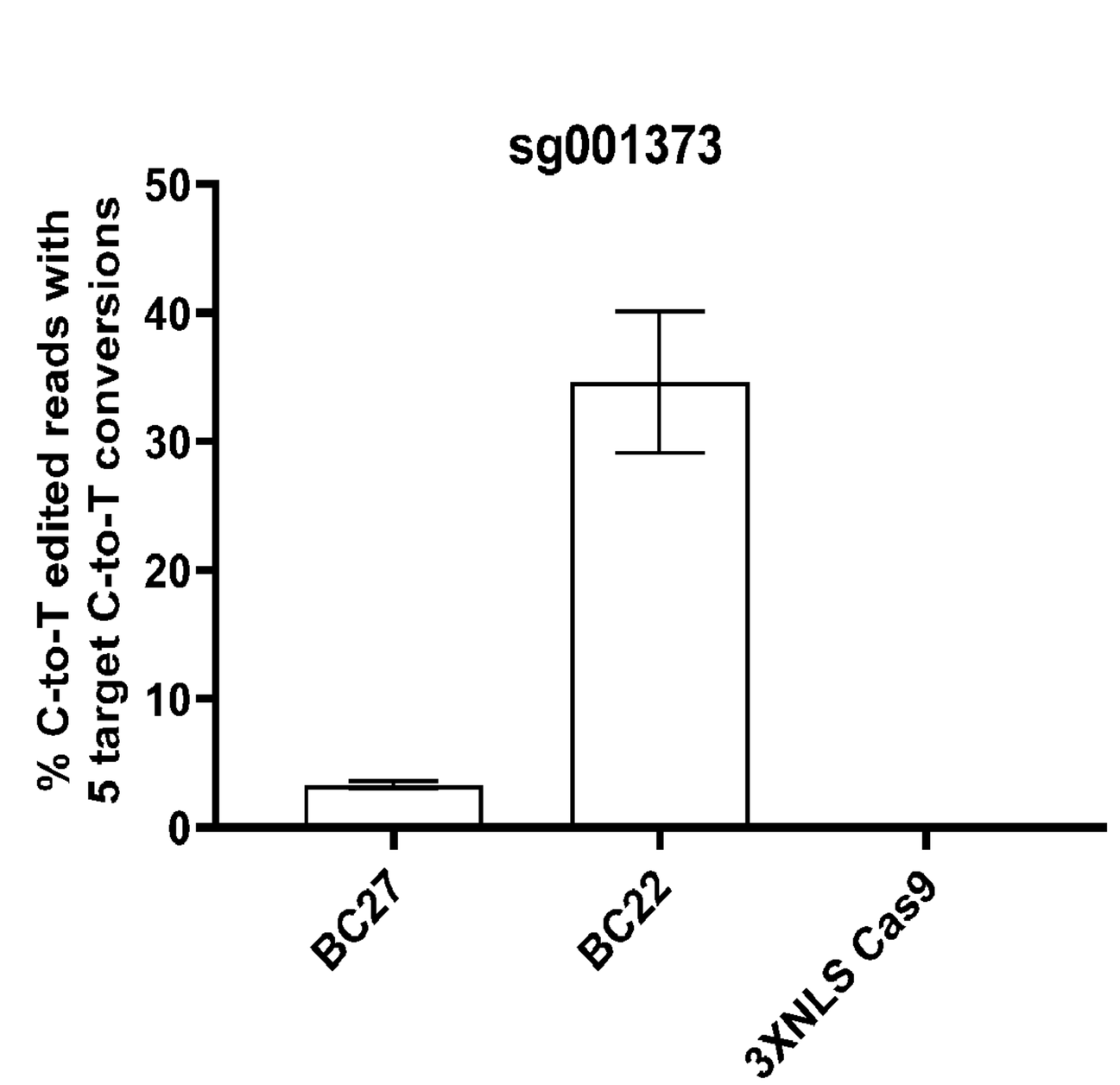

sg001373
% C-to-T edited reads with
5 target C-to-T conversions
50
40
30
20
10
0
BC27
BC22
3XNLS Cas9

FIG. 2B

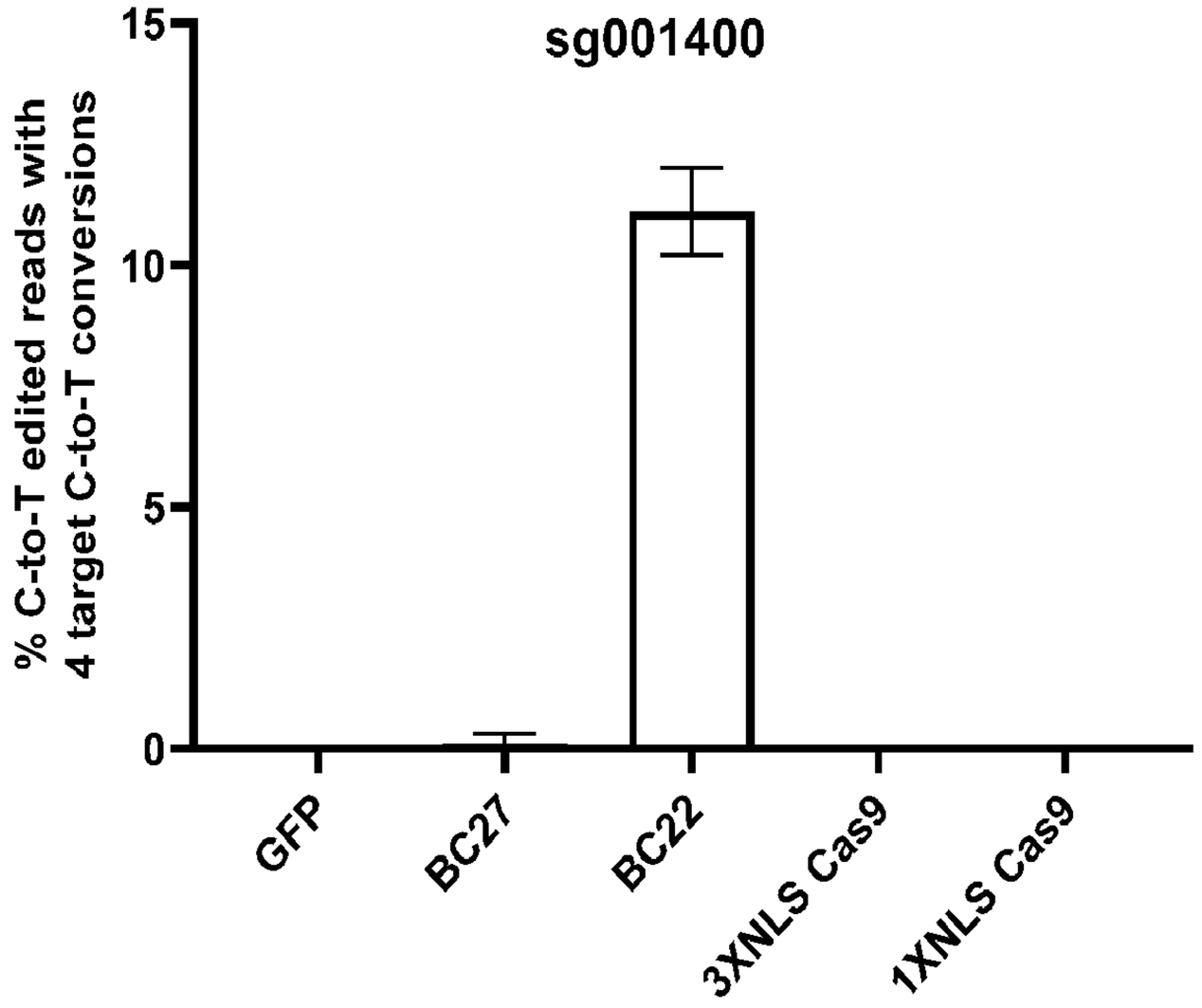

sg001400
% C-to-T edited reads with
4 target C-to-T conversions
15
10
5
0
GFP
BC27
BC22
3XNLS Cas9
1XNLS Cas9

FIG. 2C

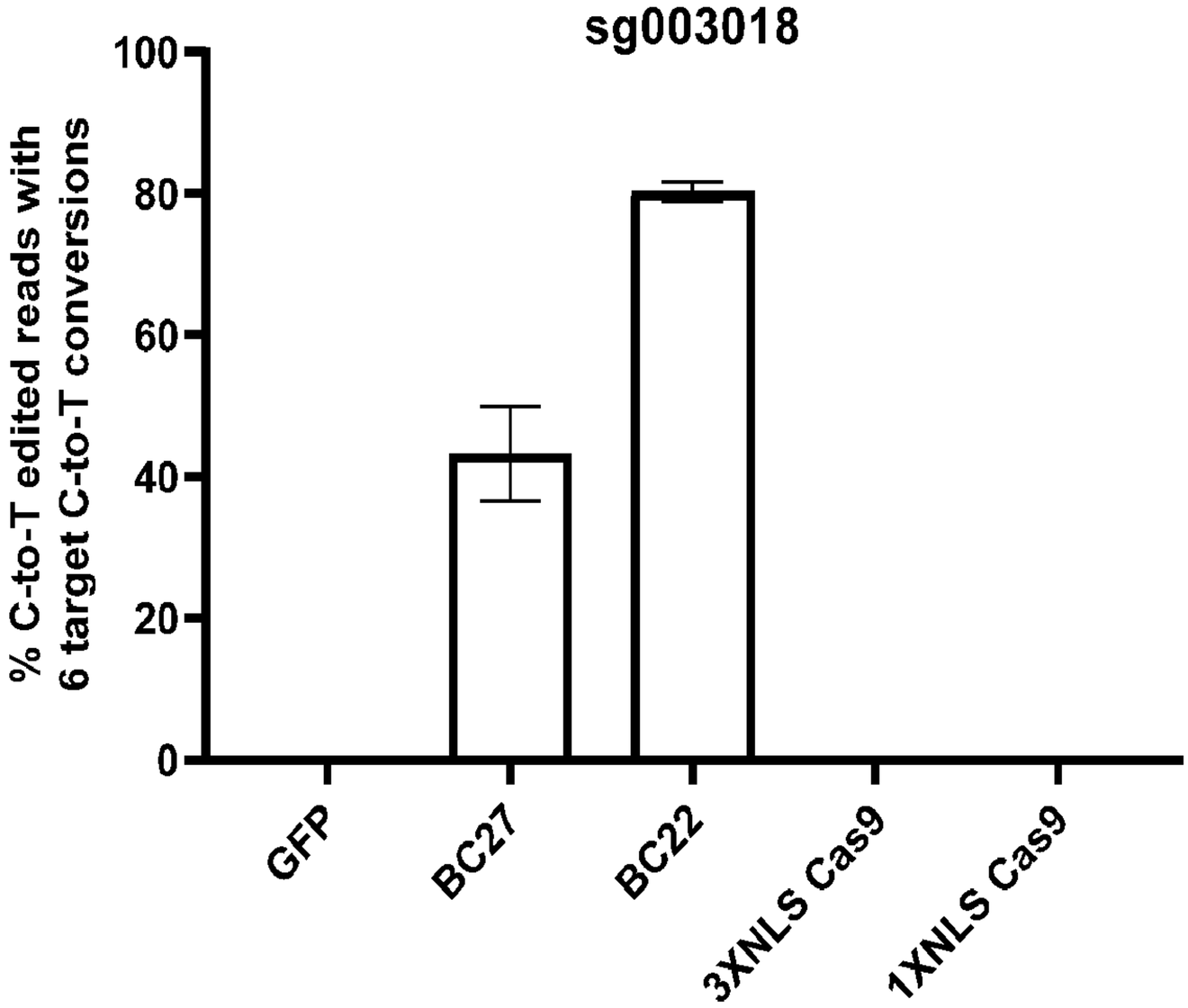

sg003018
% C-to-T edited reads with
6 target C-to-T conversions
100
80
60
40
20
0
GFP
BC27
BC22
3XNLS Cas9
1XNLS Cas9

FIG. 2D

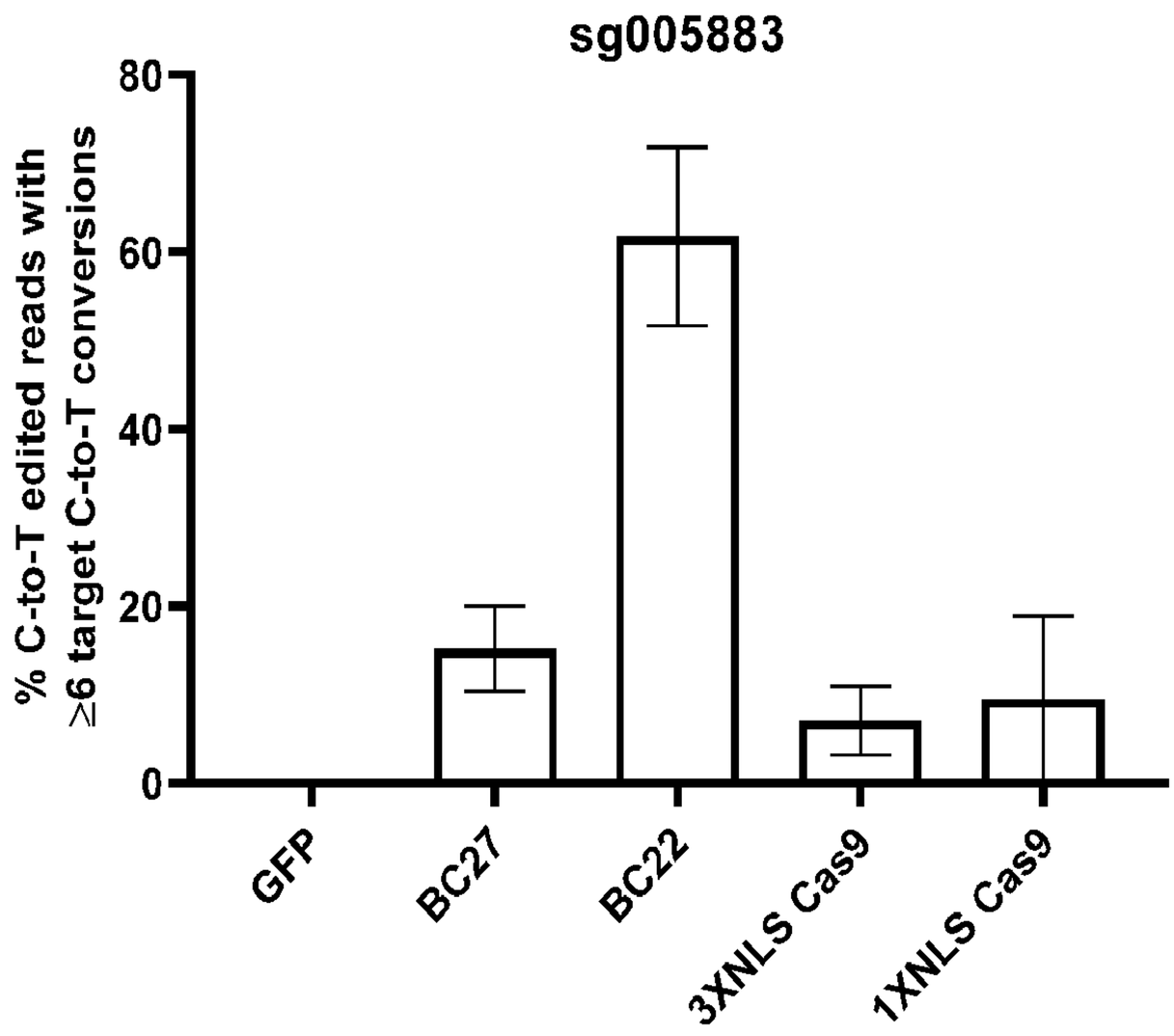

sg005883
% C-to-T edited reads with ≥6 target C-to-T conversions
80
60
40
20
0
GFP
BC27
BC22
3XNLS Cas9
1XNLS Cas9

FIG. 2E

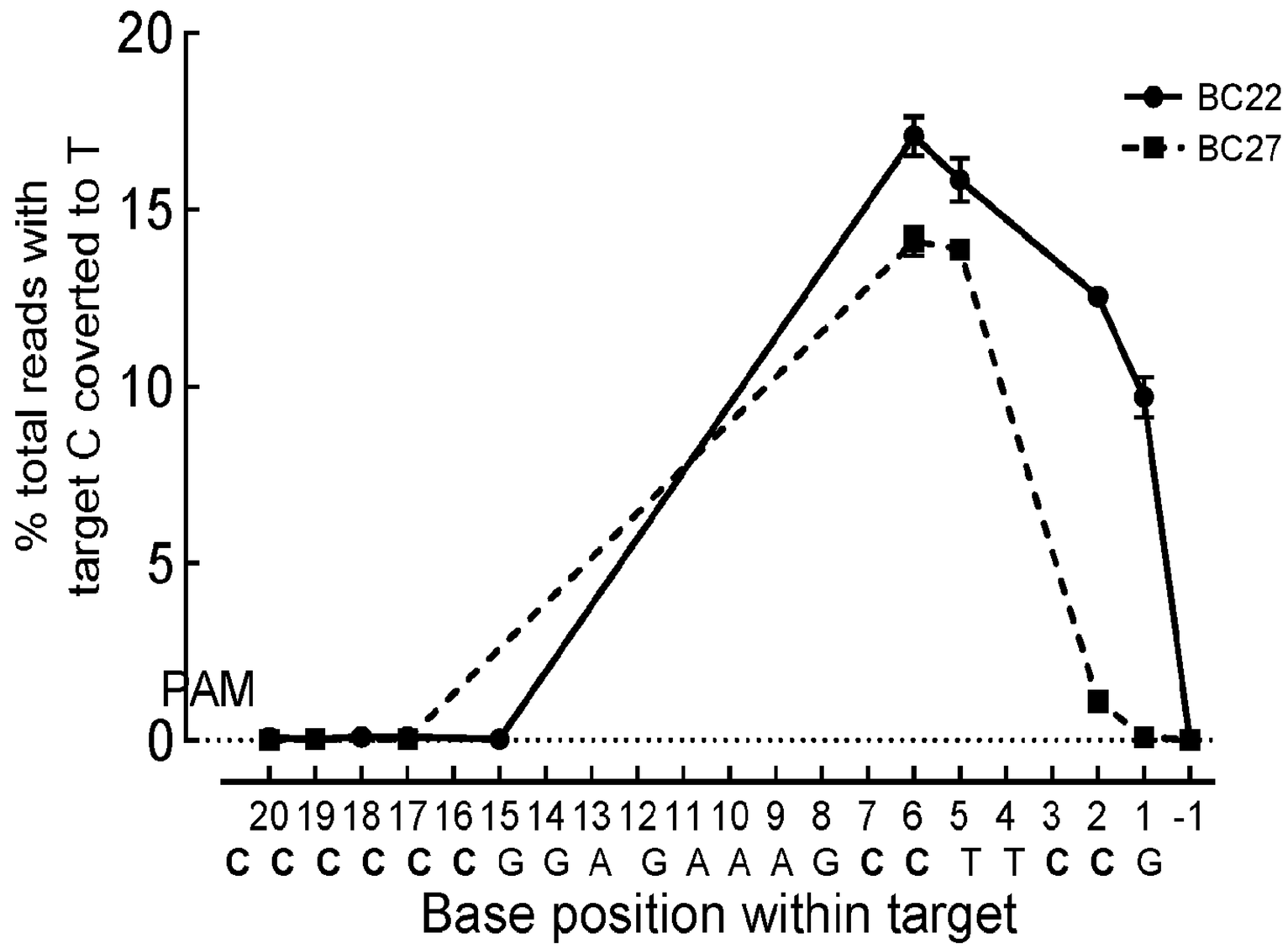

% total reads with target C coverted to T
20
15
10
5
PAM
0
BC22
BC27
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 -1
C C C C C C G G A G A A A G C C T T C C G
Base position within target

FIG. 3A

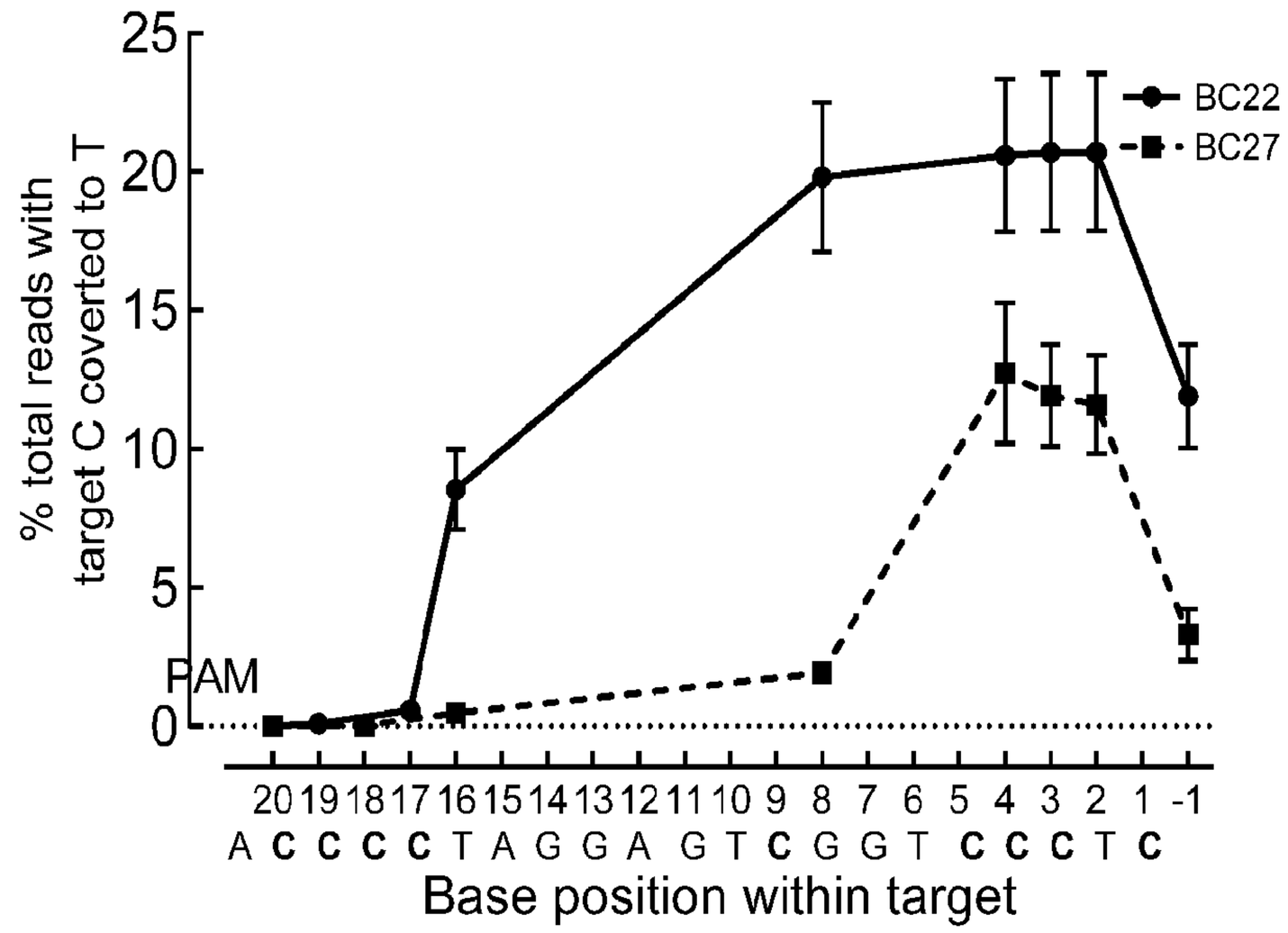

% total reads with target C coverted to T
25
20
15
10
5
PAM
0
BC22
BC27
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 -1
A C C C C T A G G A G T C G G T C C C T C
Base position within target

FIG. 3B

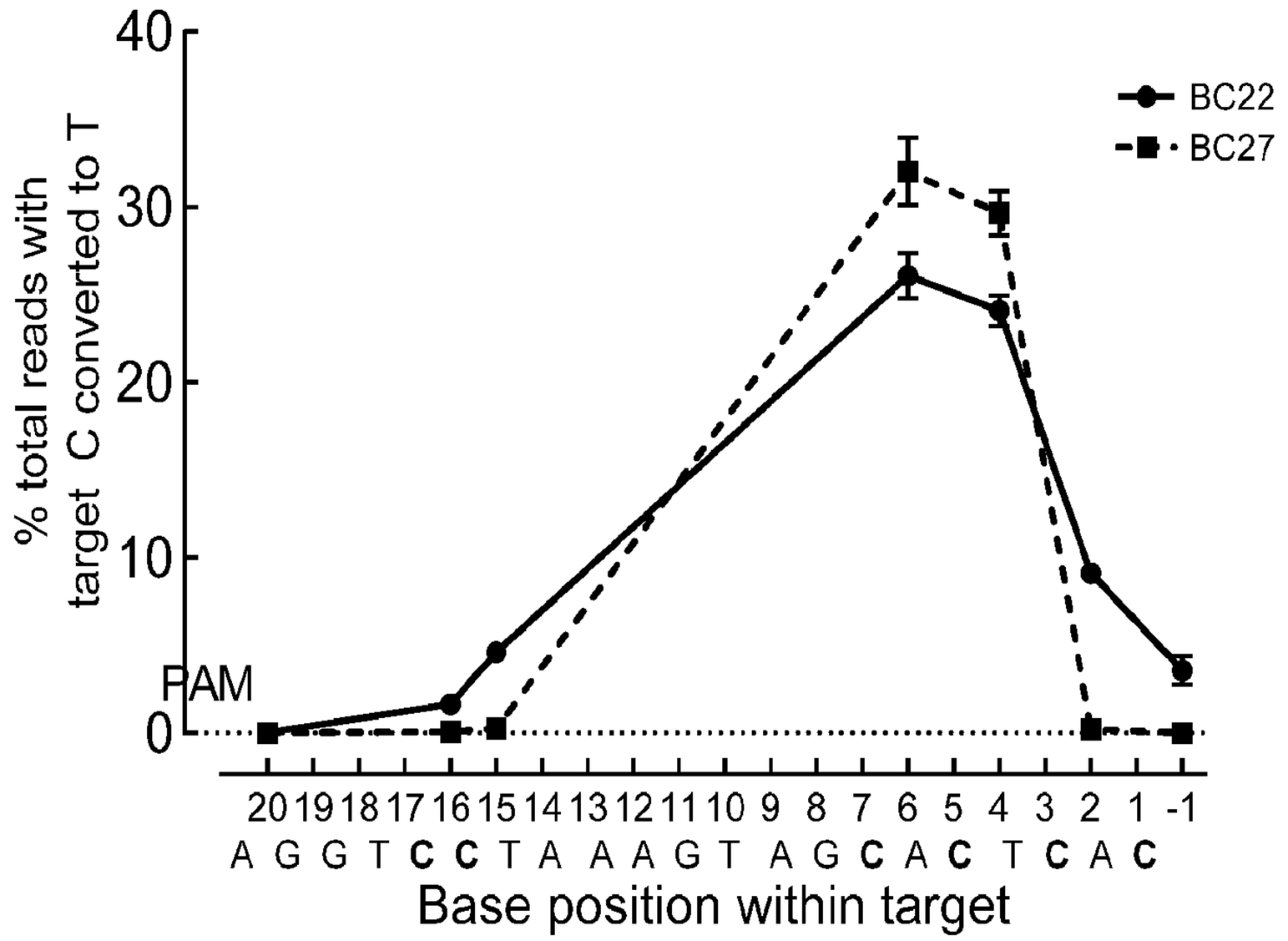

% total reads with target C converted to T
40
30
20
10
0
PAM
BC22
BC27
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 -1
A G G T C C T A A A G T A G C A C T C A C
Base position within target

FIG. 3C

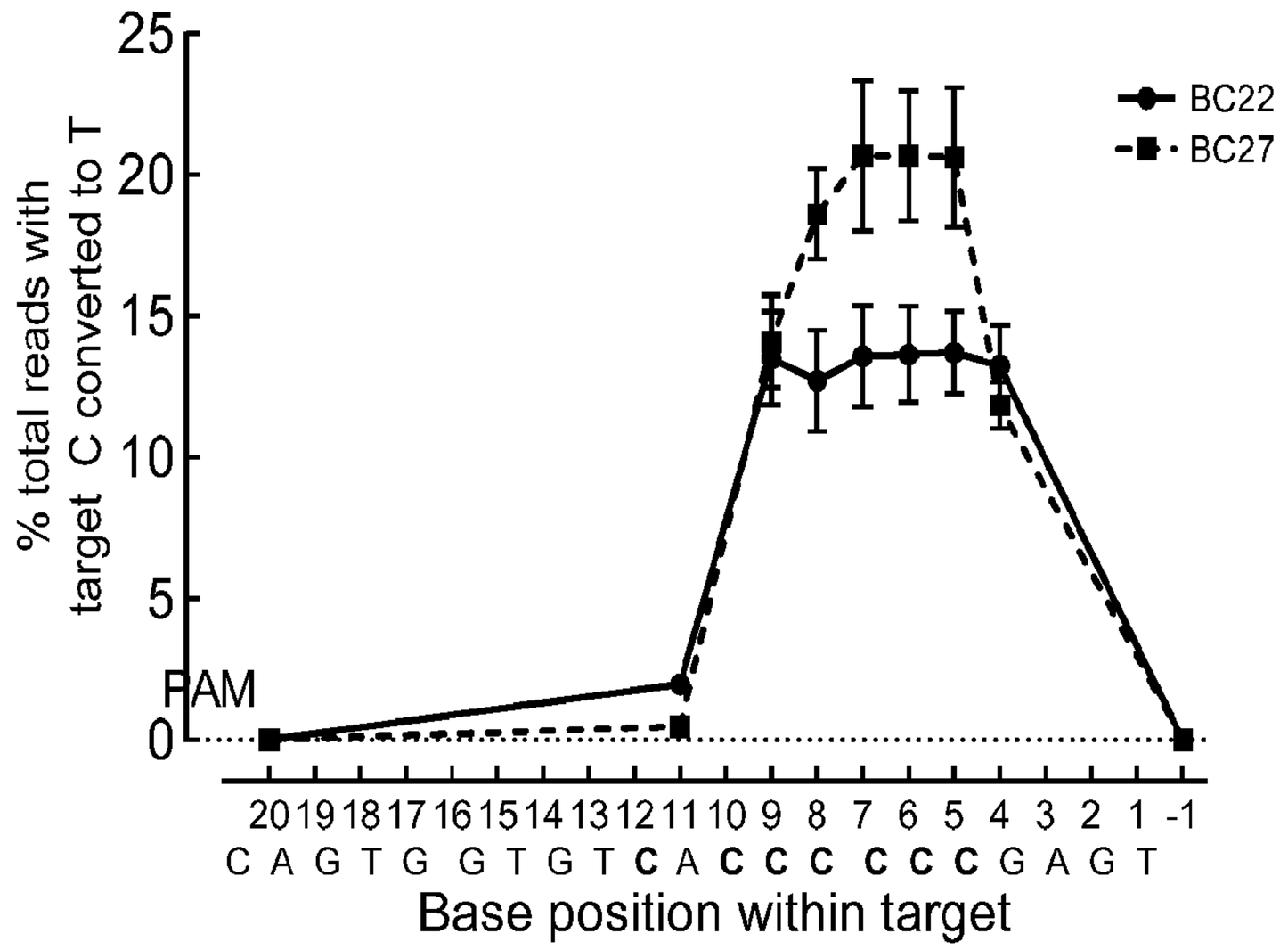

% total reads with target C converted to T
25
20
15
10
5
0
PAM
BC22
BC27
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 -1
C A G T G G T G T C A C C C C C C G A G T
Base position within target

FIG. 3D

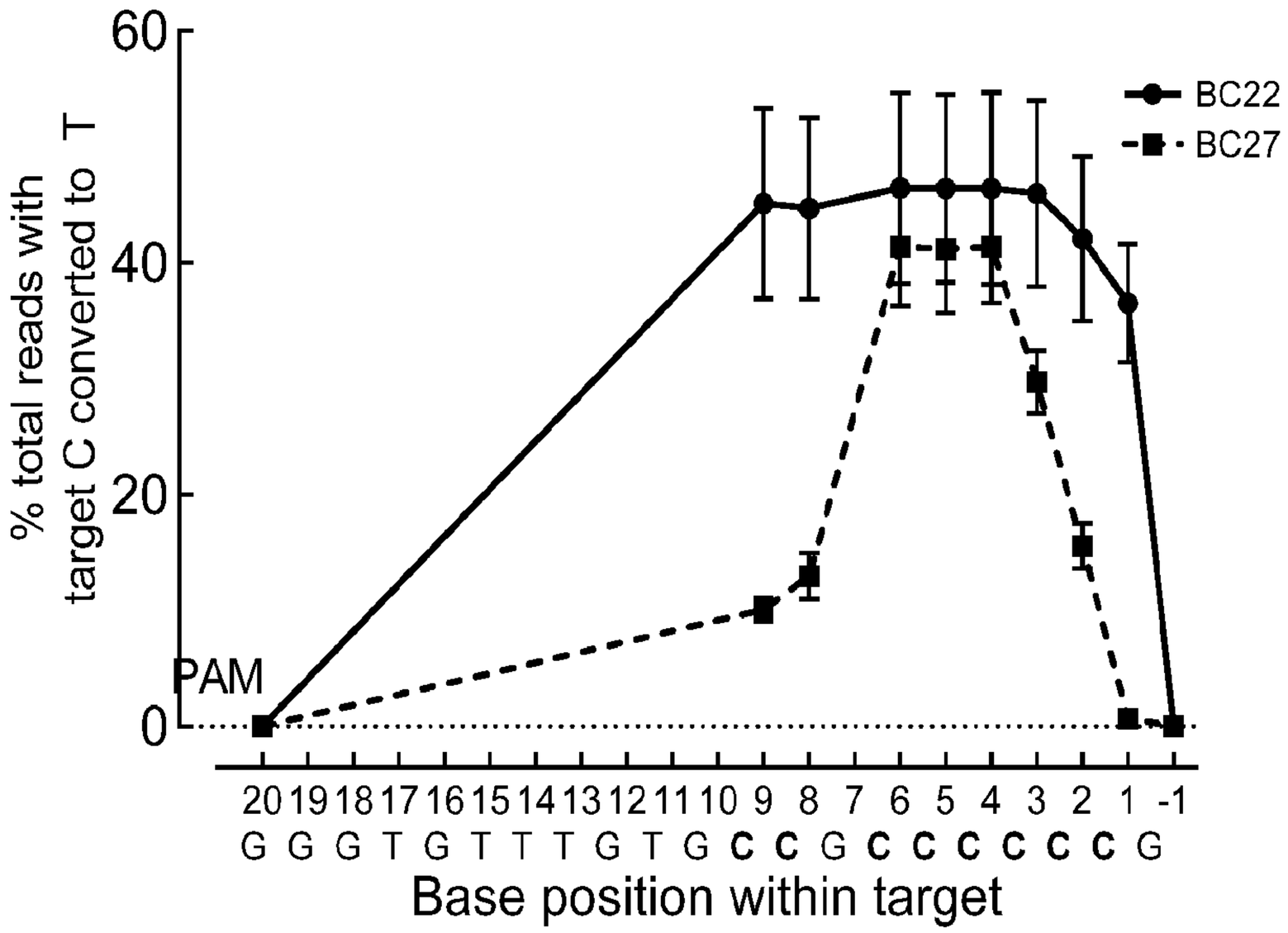

60
40
20
0
% total reads with target C converted to T
PAM
BC22
BC27
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 -1
G G G T G T T T G T G C C G C C C C C C G
Base position within target

FIG. 3E

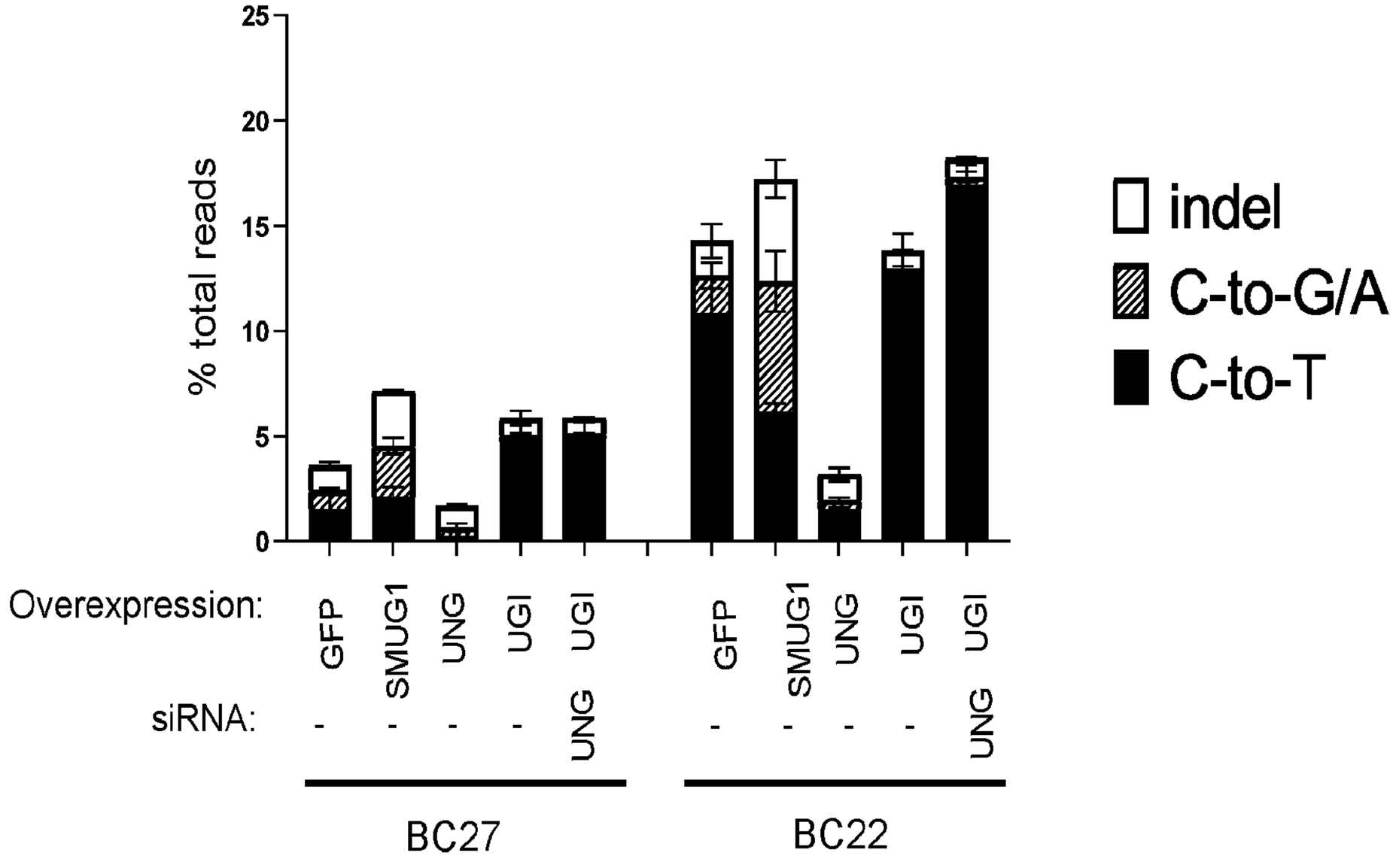

% total reads
25
20
15
10
5
0
indel
C-to-G/A
C-to-T
Overexpression:
GFP
SMUG1
UNG
UGI
UGI
siRNA:
-
-
-
-
UNG
BC27
BC22

FIG. 4A

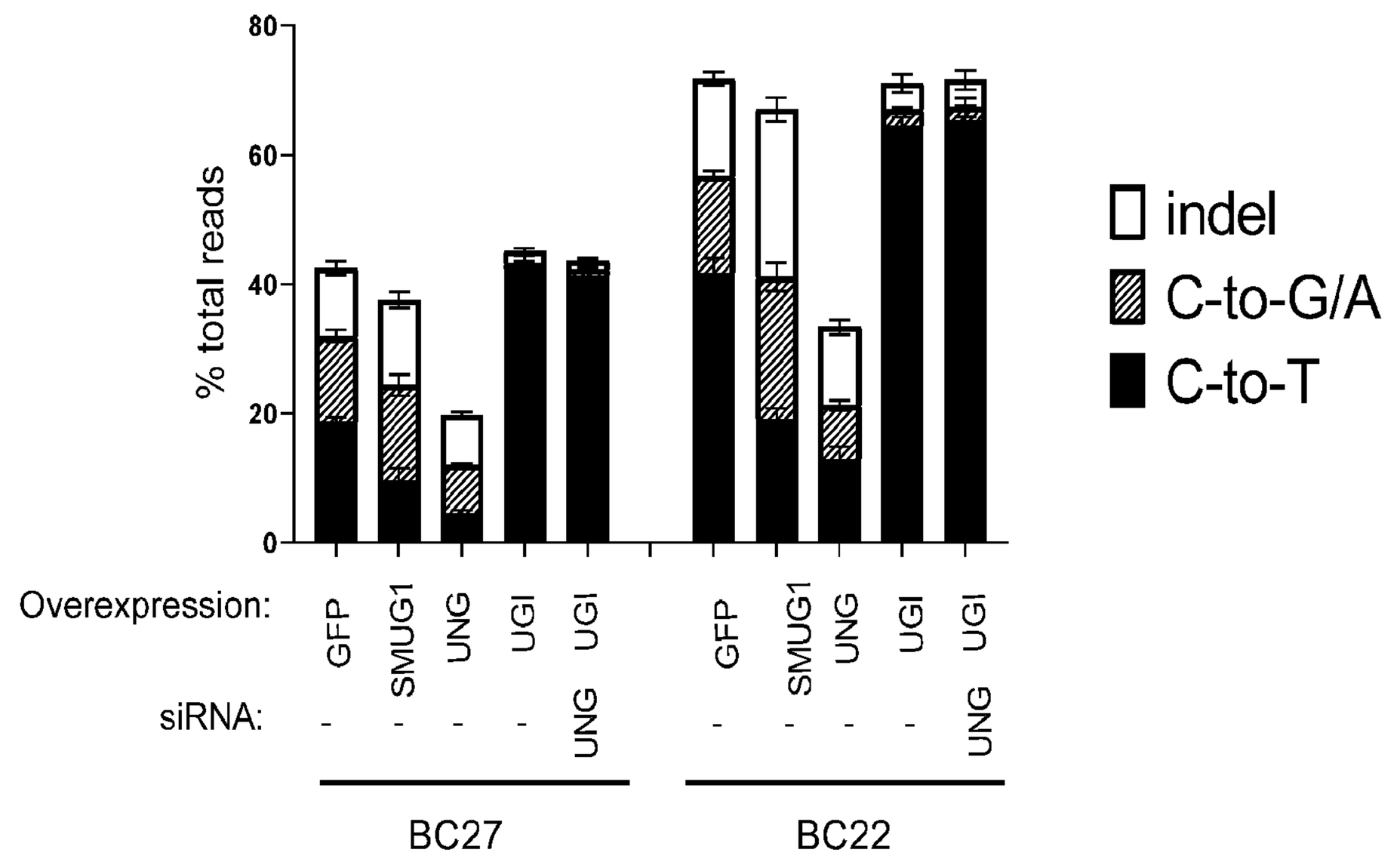

% total reads
80
60
40
20
0
indel
C-to-G/A
C-to-T
Overexpression:
GFP
SMUG1
UNG
UGI
UGI
siRNA:
-
-
-
-
UNG
BC27
BC22

FIG. 4B

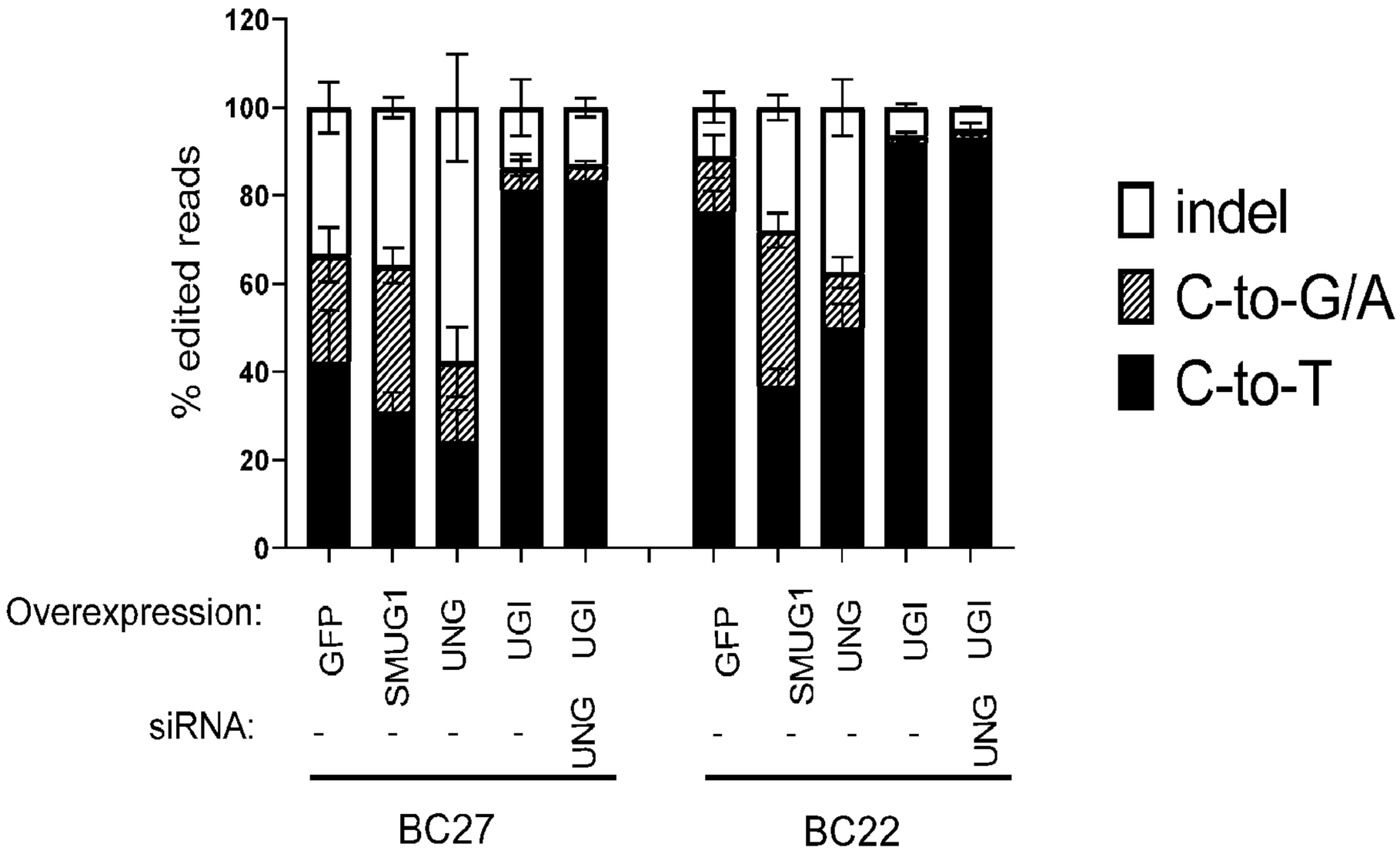

120
100
80
60
40
20
0
% edited reads
indel
C-to-G/A
C-to-T
Overexpression:
GFP
SMUG1
UNG
UGI
UGI
GFP
SMUG1
UNG
UGI
UGI
siRNA:
-
-
-
-
UNG
-
-
-
-
UNG
BC27
BC22

FIG. 5A

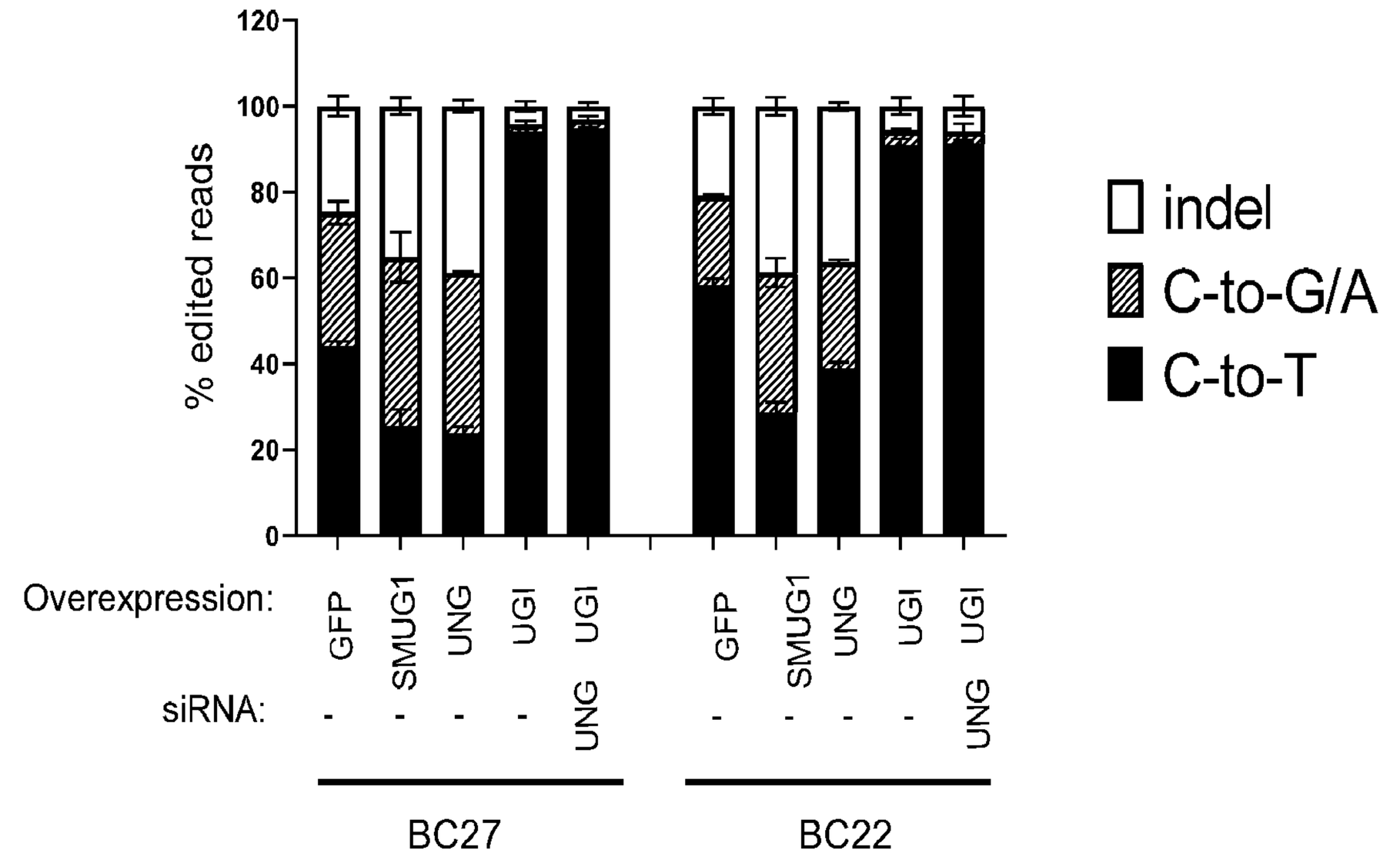

120
100
80
60
40
20
0
% edited reads
indel
C-to-G/A
C-to-T
Overexpression:
GFP
SMUG1
UNG
UGI
UGI
GFP
SMUG1
UNG
UGI
UGI
siRNA:
-
-
-
-
UNG
-
-
-
-
UNG
BC27
BC22

FIG. 5B

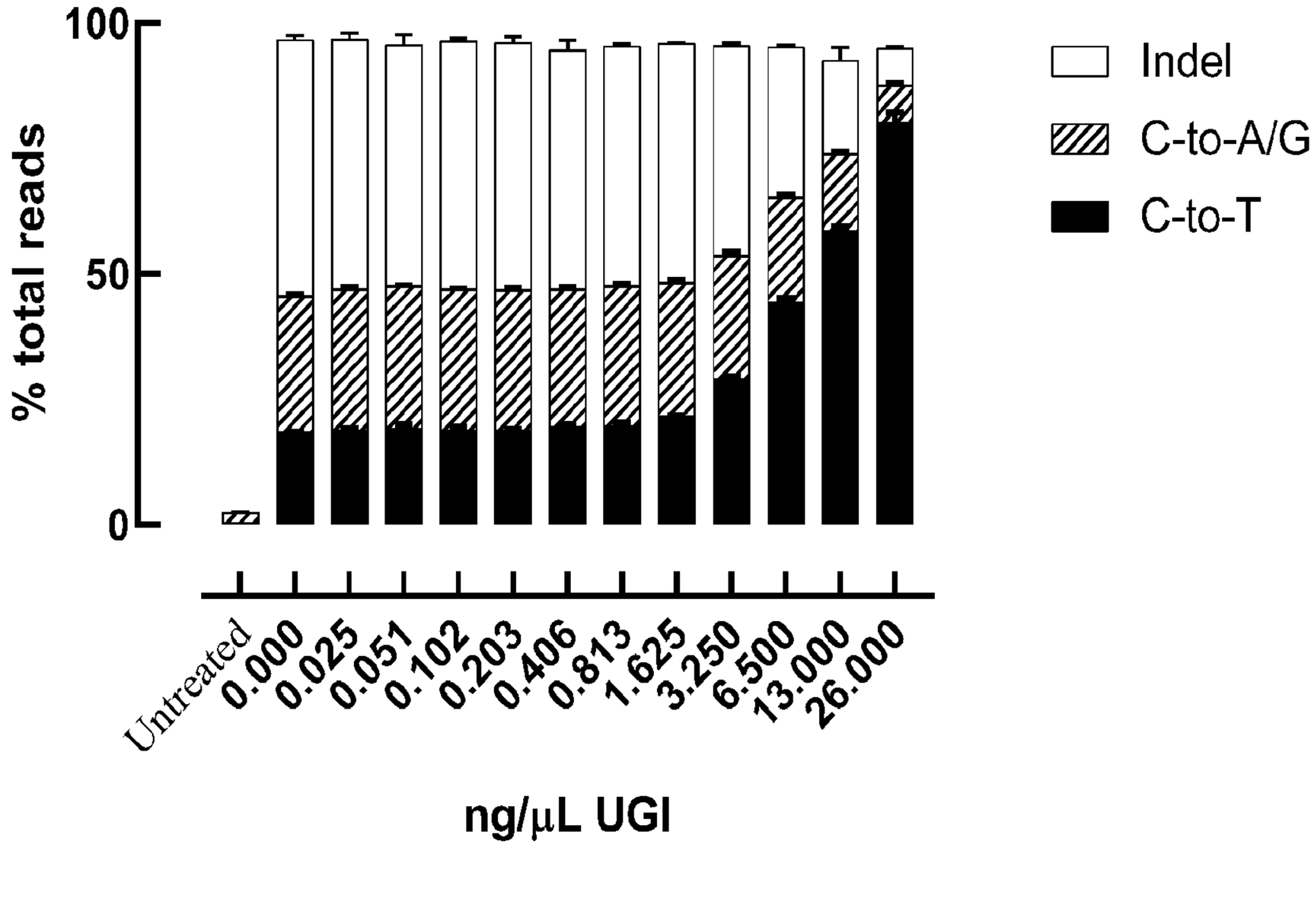

100
50
0
% total reads
Indel
C-to-A/G
C-to-T
Untreated
0.000
0.025
0.051
0.102
0.203
0.406
0.813
1.625
3.250
6.500
13.000
26.000
ng/μL UGI

FIG. 6A

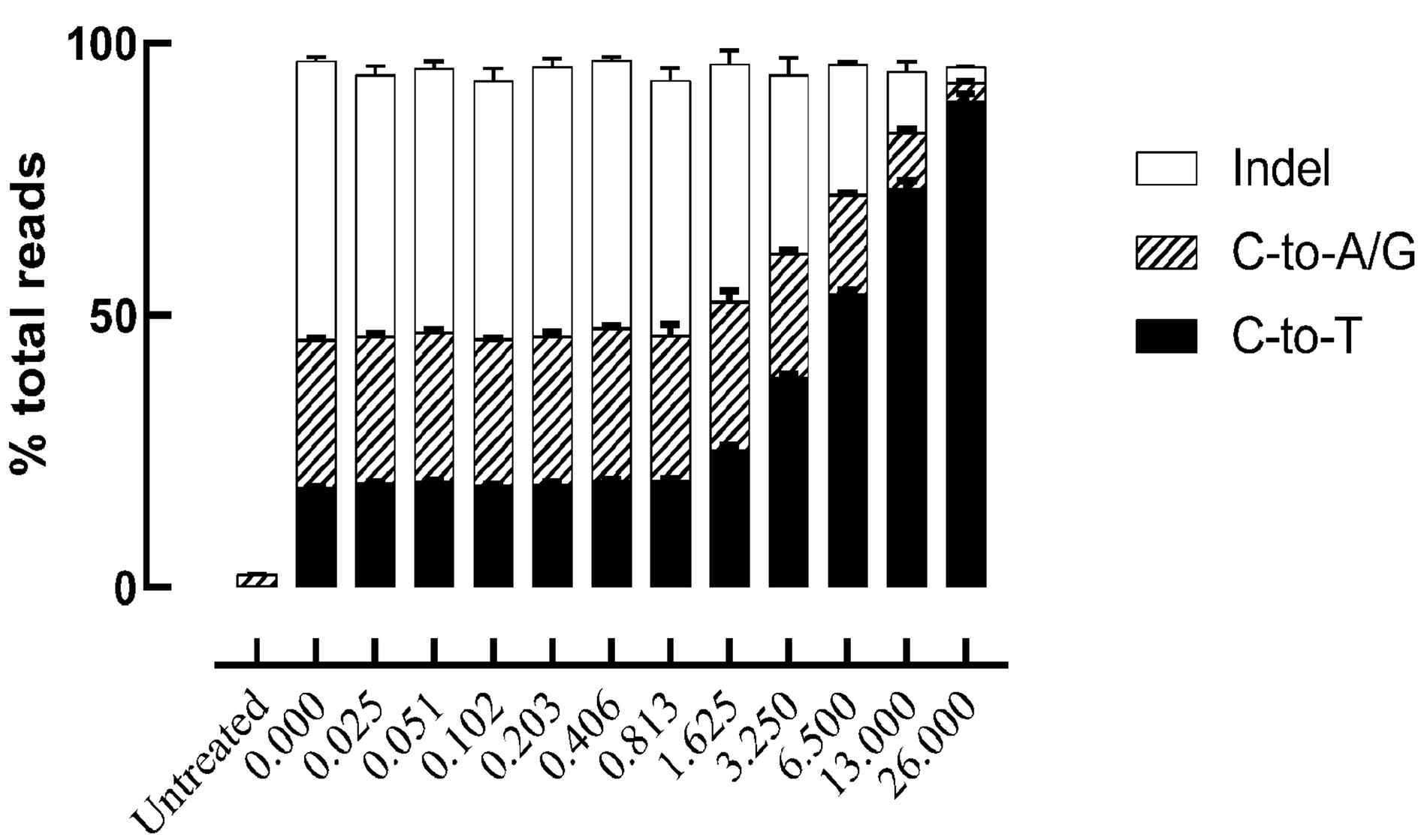

100
50
0
% total reads
Indel
C-to-A/G
C-to-T
Untreated
0.000
0.025
0.051
0.102
0.203
0.406
0.813
1.625
3.250
6.500
13.000
26.000
ng/μL UGI

FIG. 6B

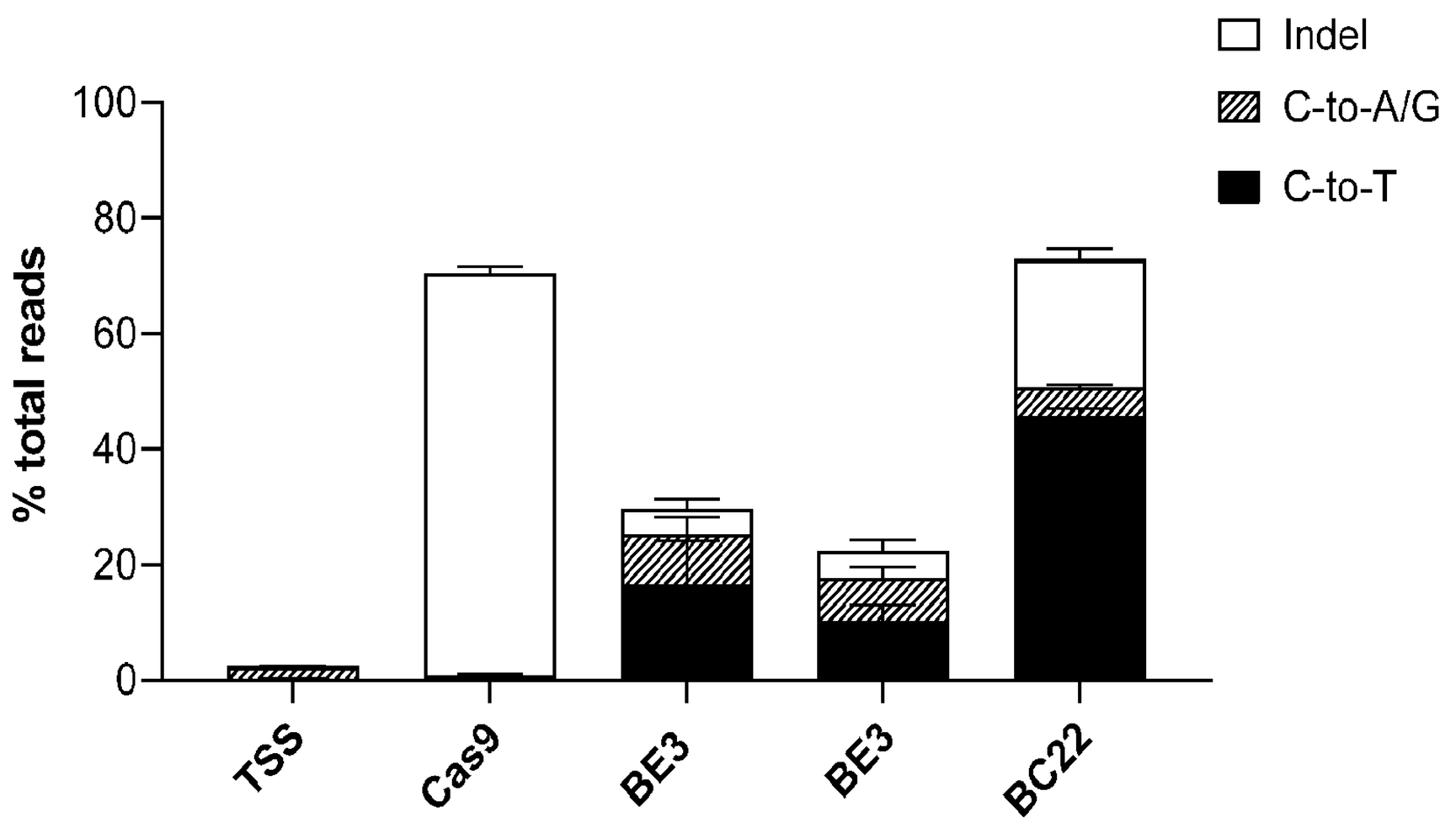

Indel
C-to-A/G
C-to-T
% total reads
100
80
60
40
20
0
TSS
Cas9
BE3
BE3
BC22

FIG. 7

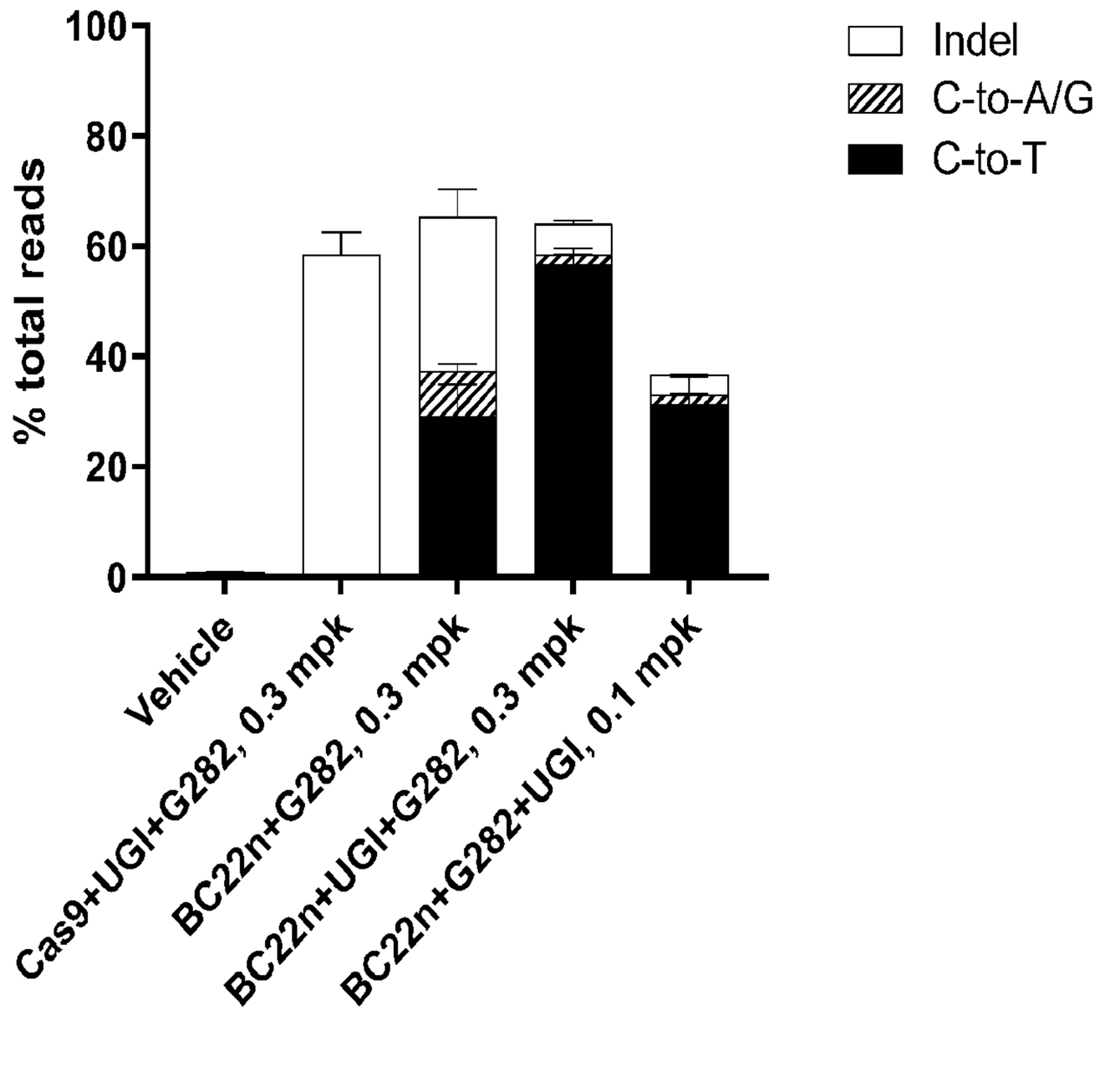

Indel
C-to-A/G
C-to-T
% total reads
100
80
60
40
20
0
Vehicle
Cas9+UGI+G282, 0.3 mpk
BC22n+G282, 0.3 mpk
BC22n+UGI+G282, 0.3 mpk
BC22n+G282+UGI, 0.1 mpk

FIG. 8

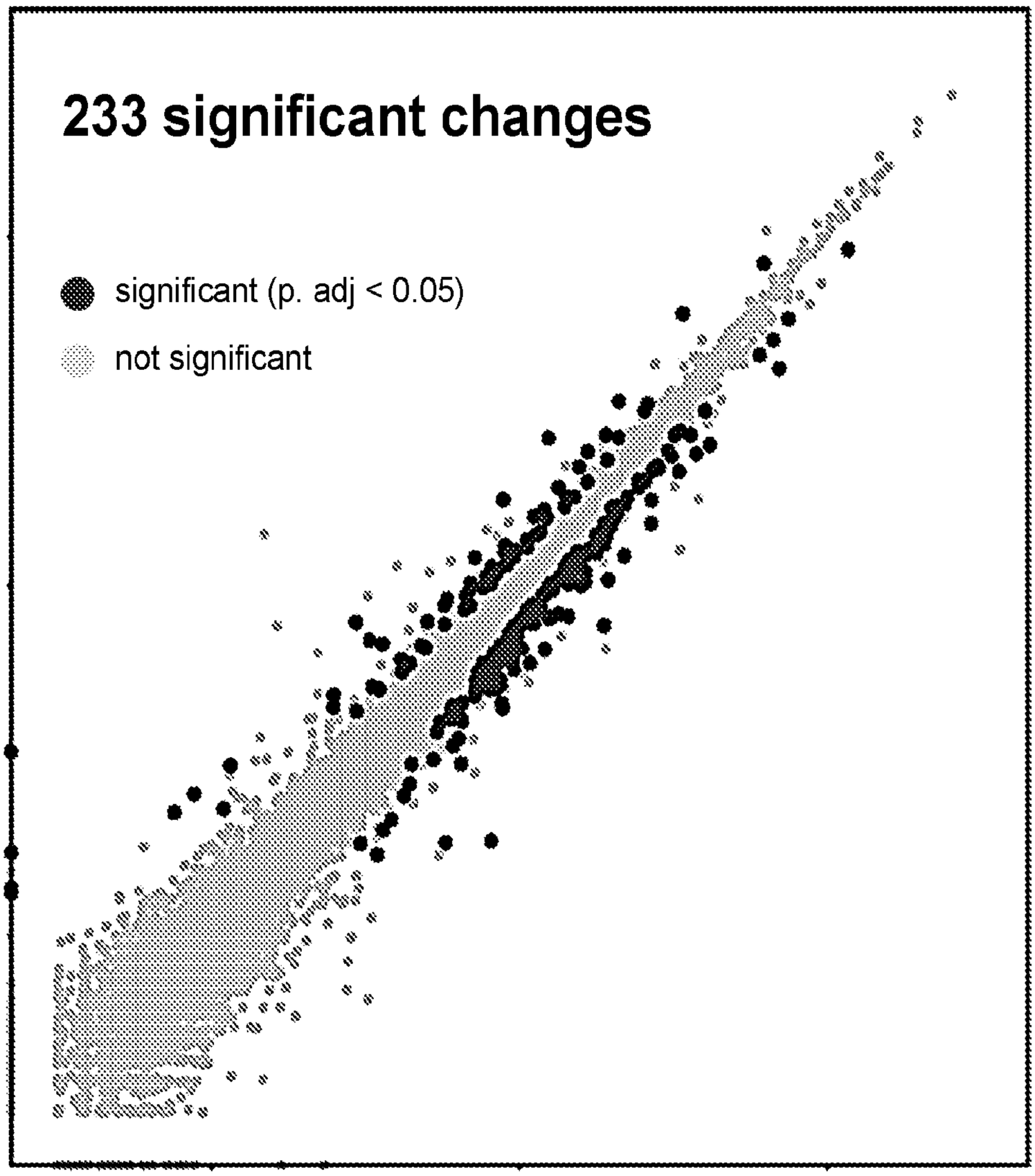

233 significant changes
significant (p. adj < 0.05)
not significant
BC22n + G000282
Vehicle (TSS)

FIG. 9A

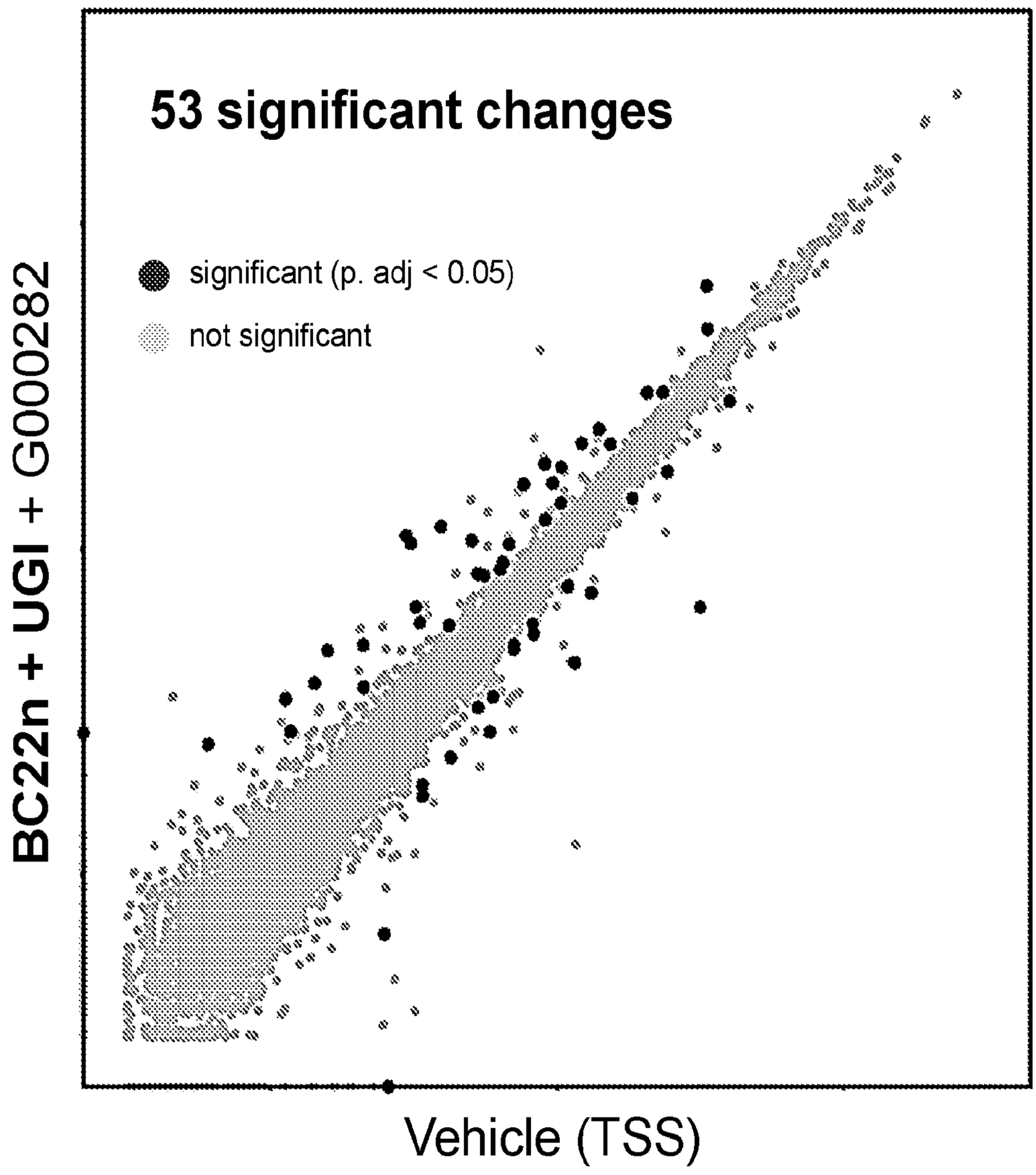

53 significant changes
significant (p. adj < 0.05)
not significant
BC22n + UGI + G000282
Vehicle (TSS)

FIG. 9B

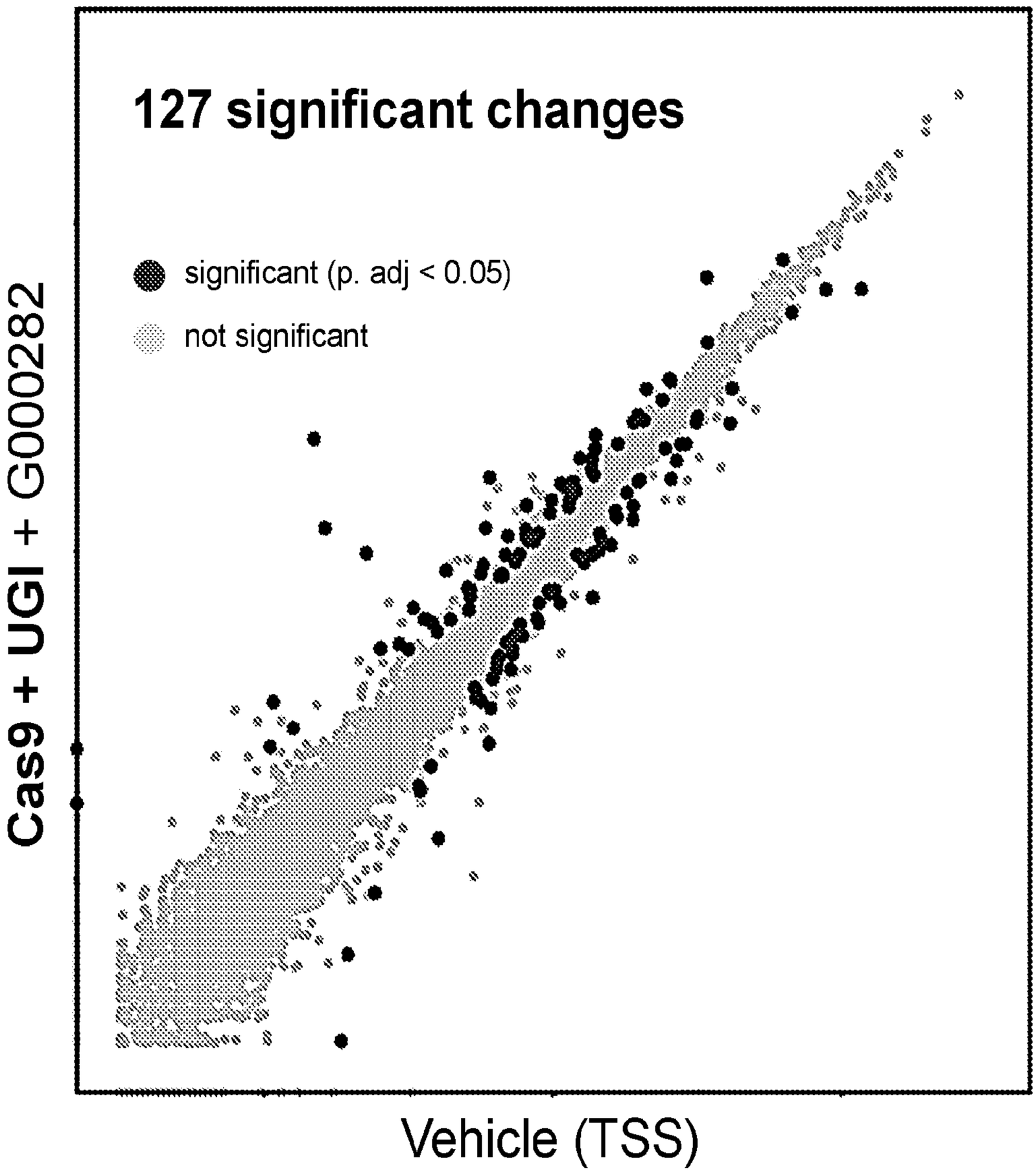

127 significant changes
significant (p. adj < 0.05)
not significant
Cas9 + UGI + G000282
Vehicle (TSS)

FIG. 9C

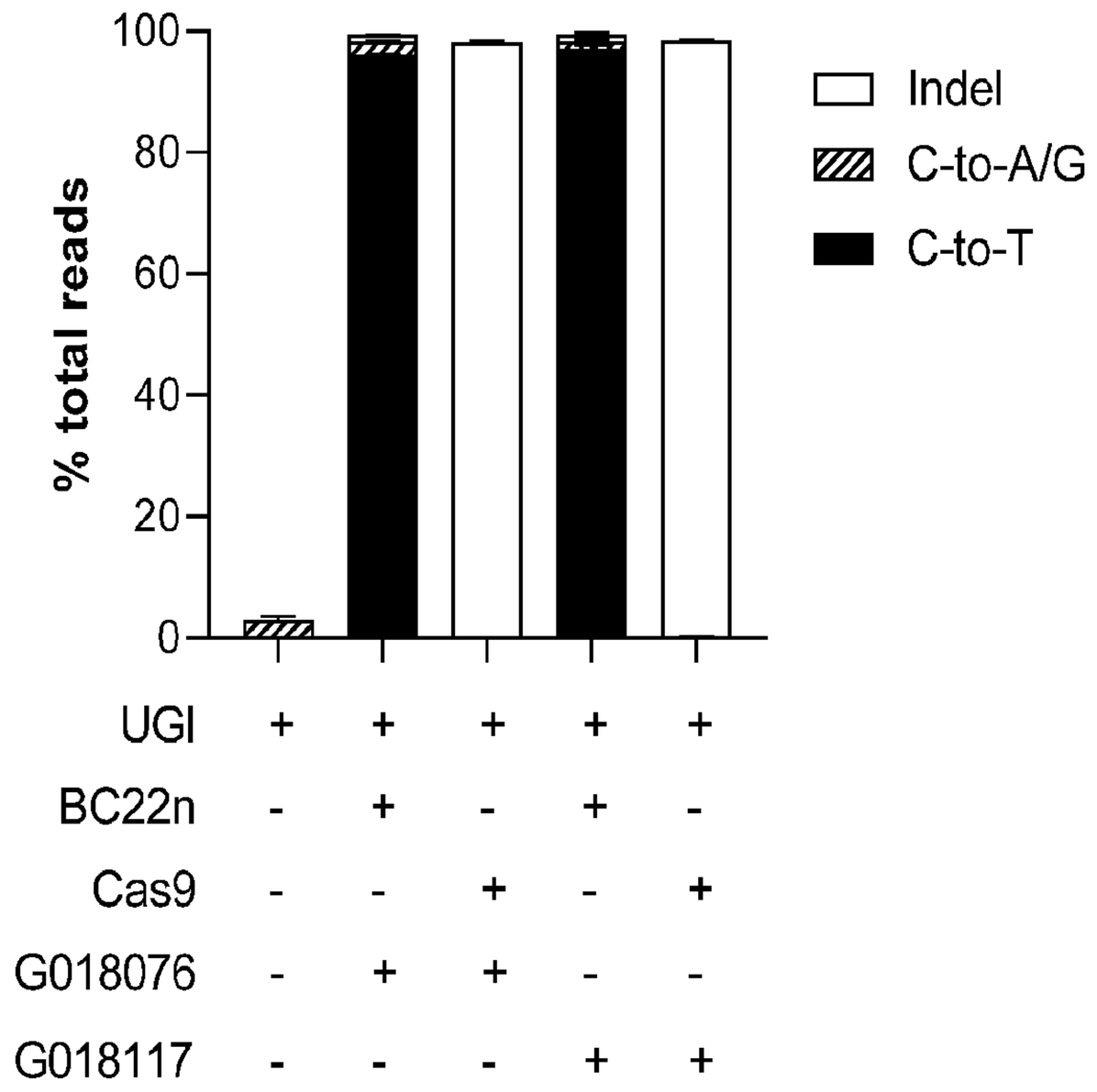

100
80
60
40
20
0
% total reads
Indel
C-to-A/G
C-to-T
UGI + + + + +
BC22n - + - + -
Cas9 - - + - +
G018076 - + + - -
G018117 - - - + +

FIG. 10

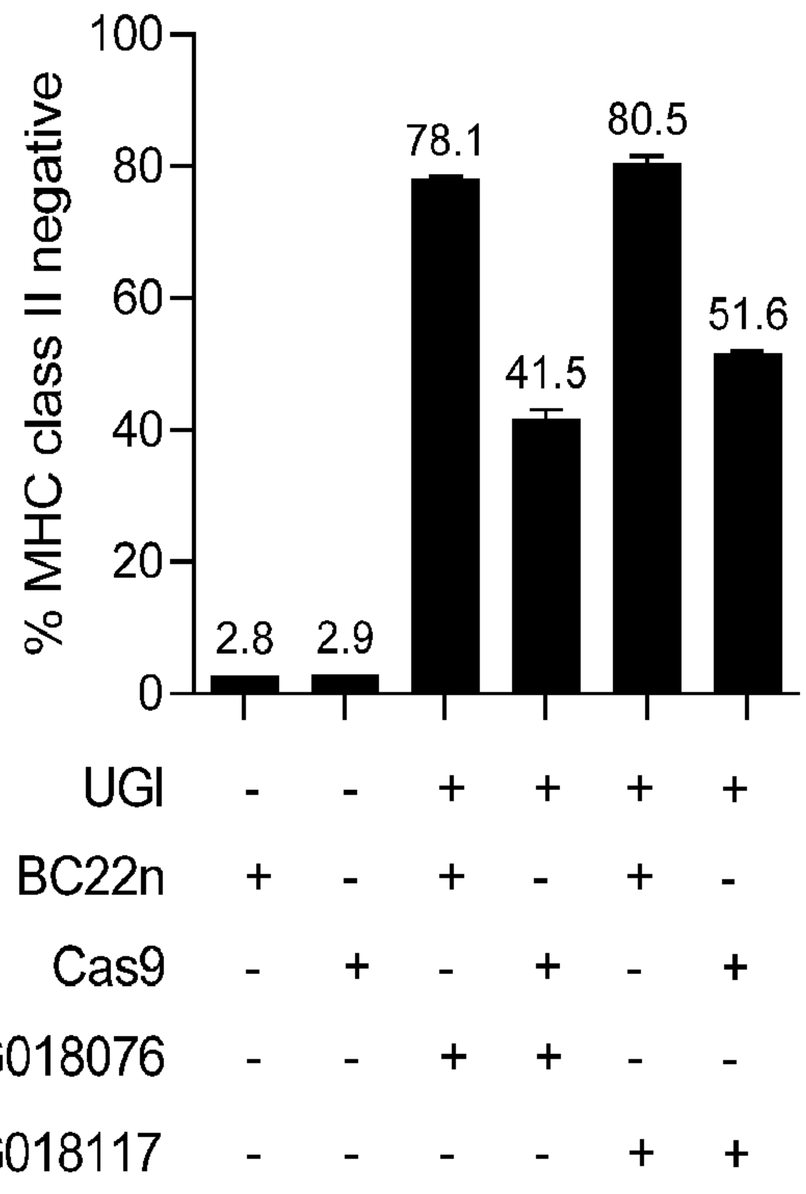

100
80
60
40
20
0
% MHC class II negative
2.8
2.9
78.1
41.5
80.5
51.6
UGI - - + + + +
BC22n + - + - + -
Cas9 - + - + - +
G018076 - - + + - -
G018117 - - - - + +

FIG. 11

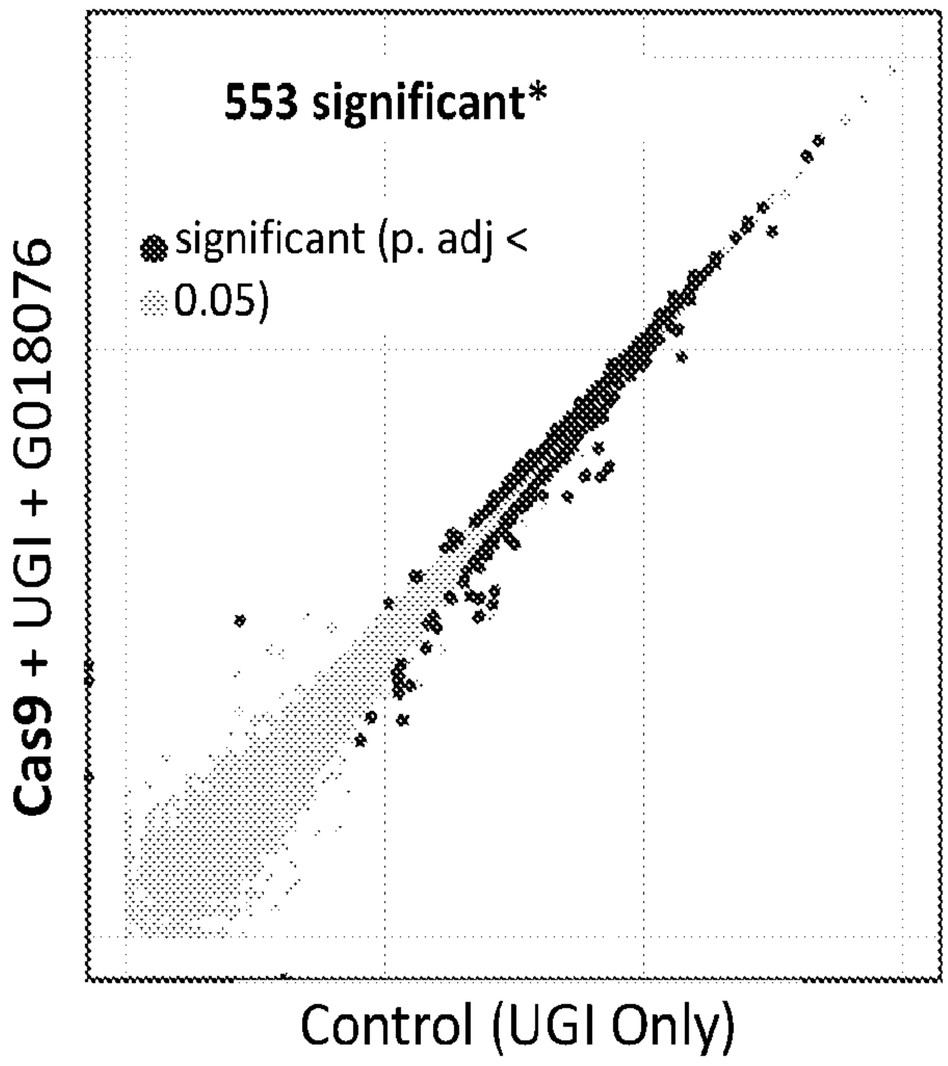


FIG. 12A

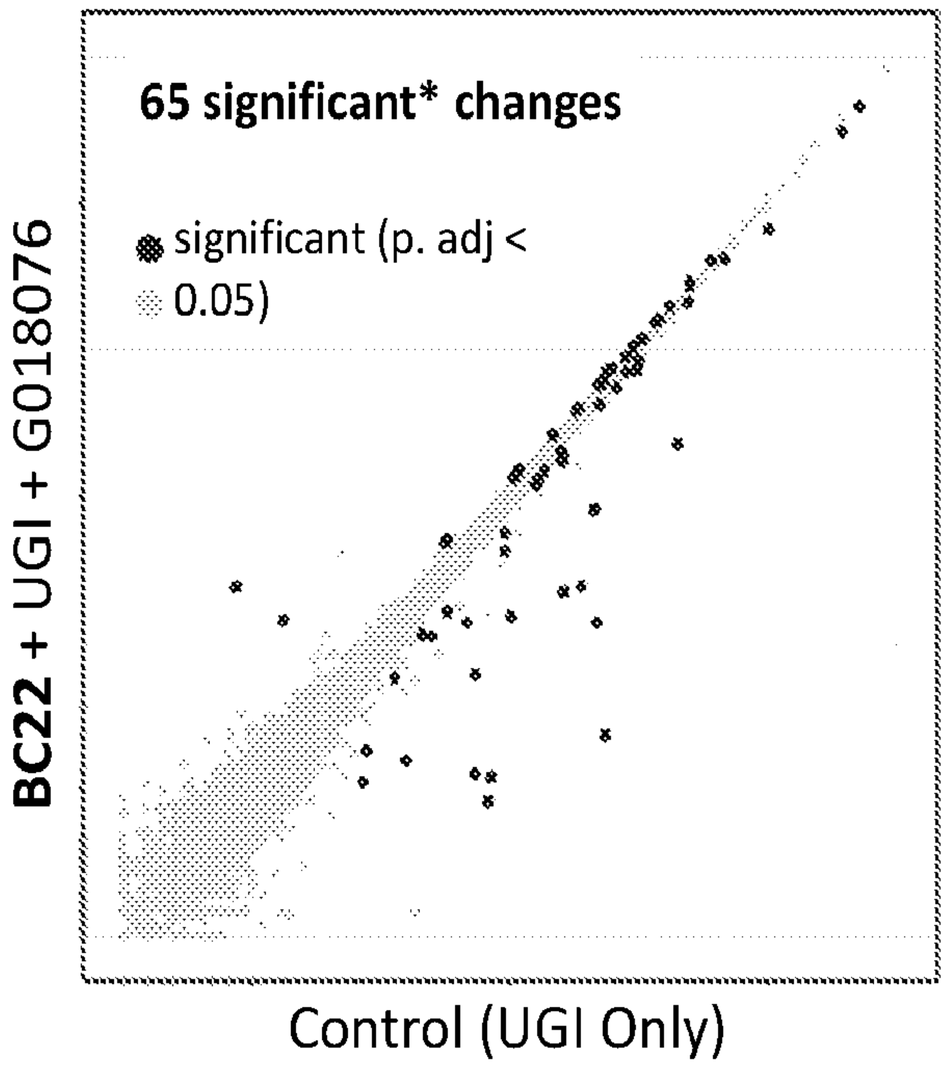


FIG. 12B

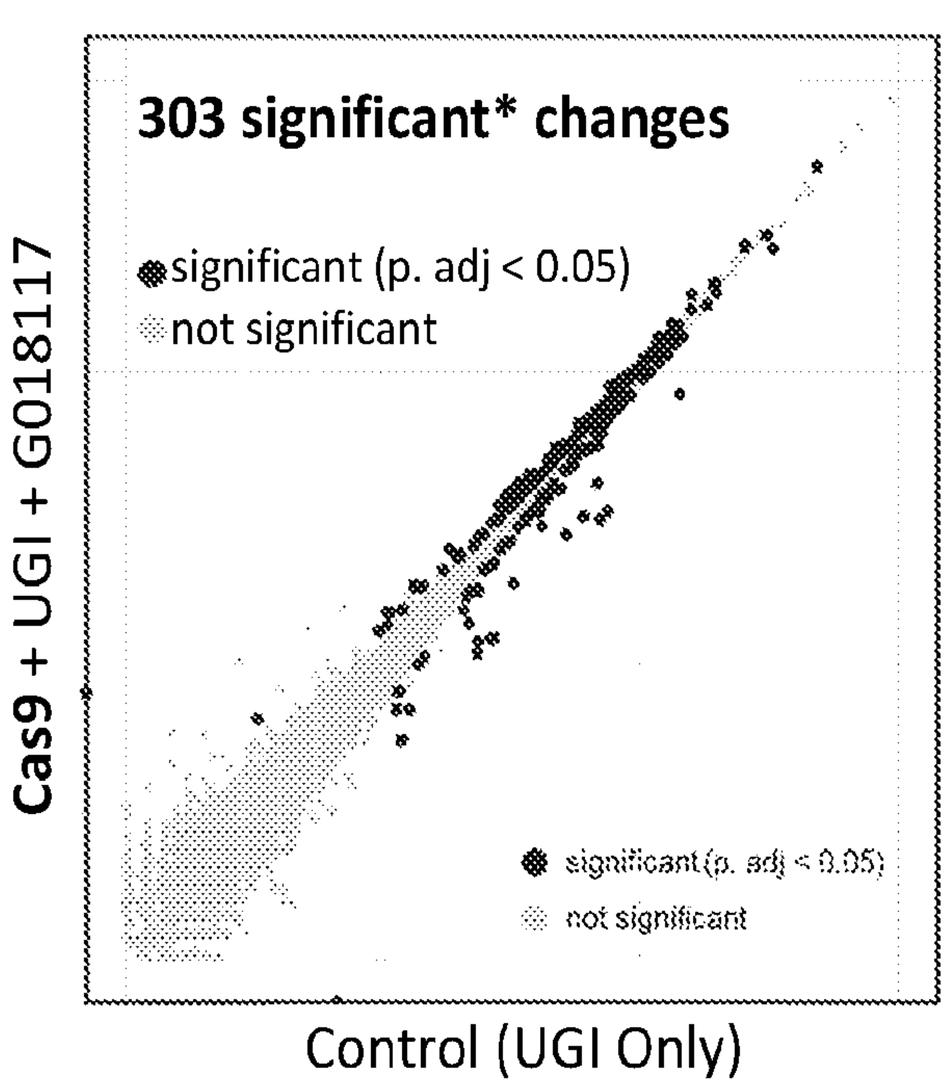


FIG. 13A

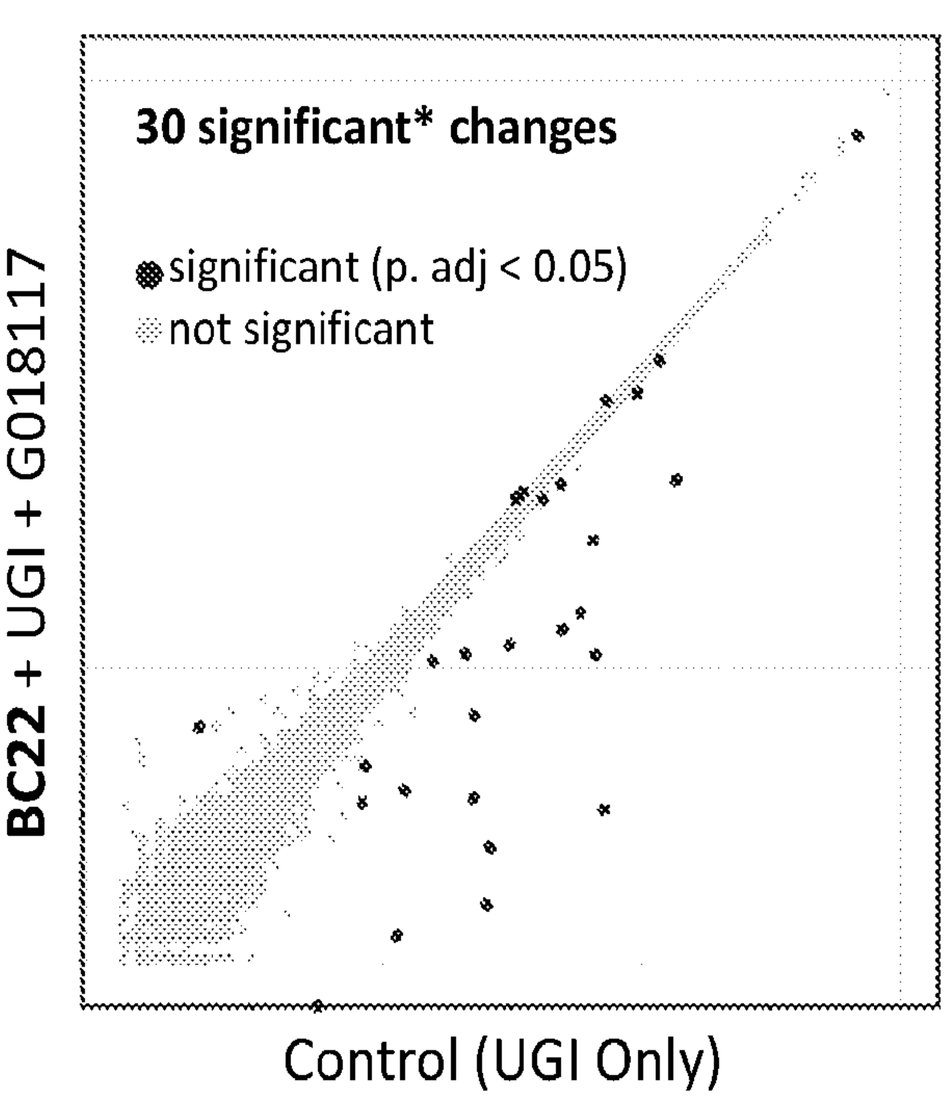


FIG. 13B

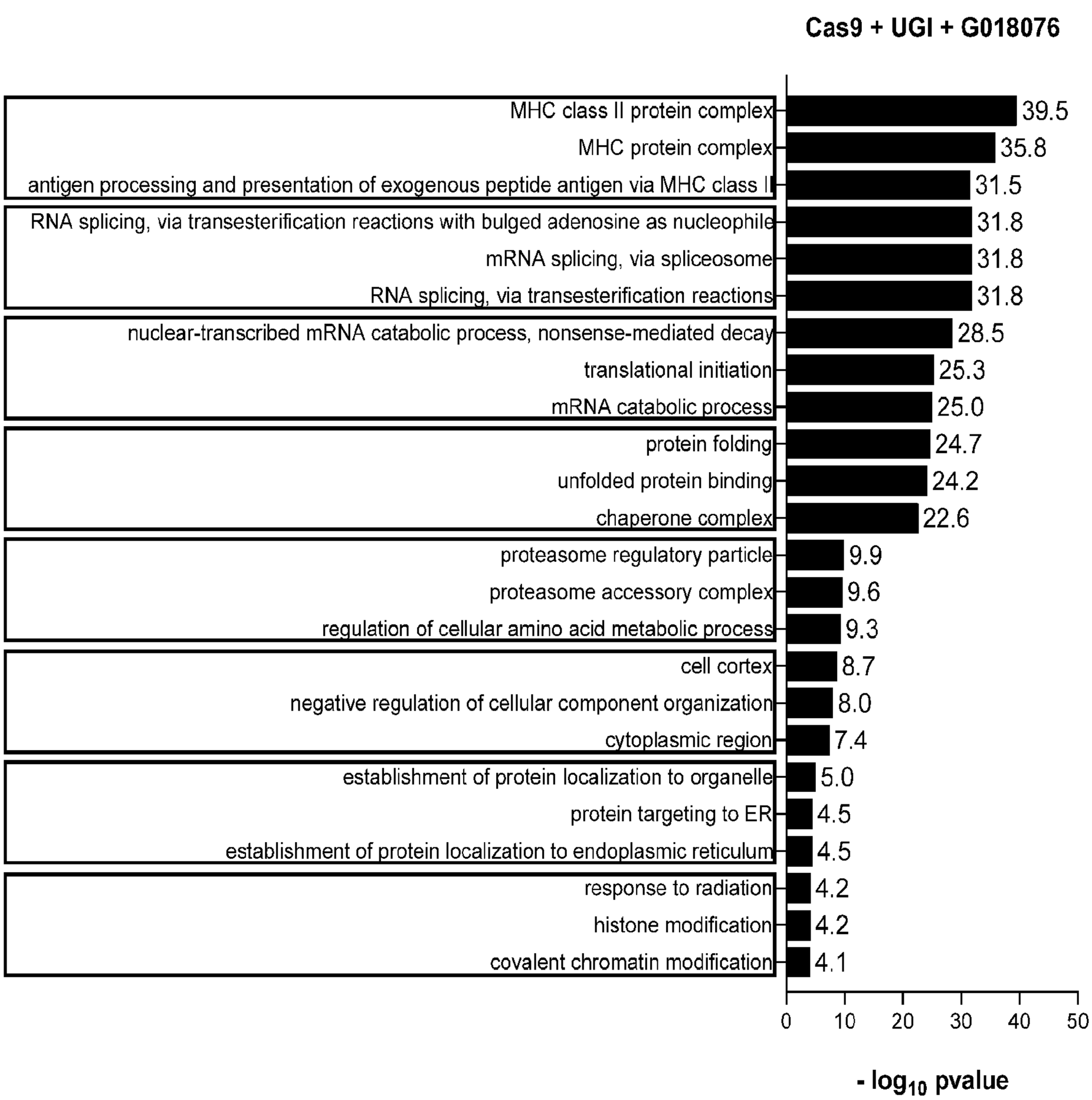


FIG. 14A

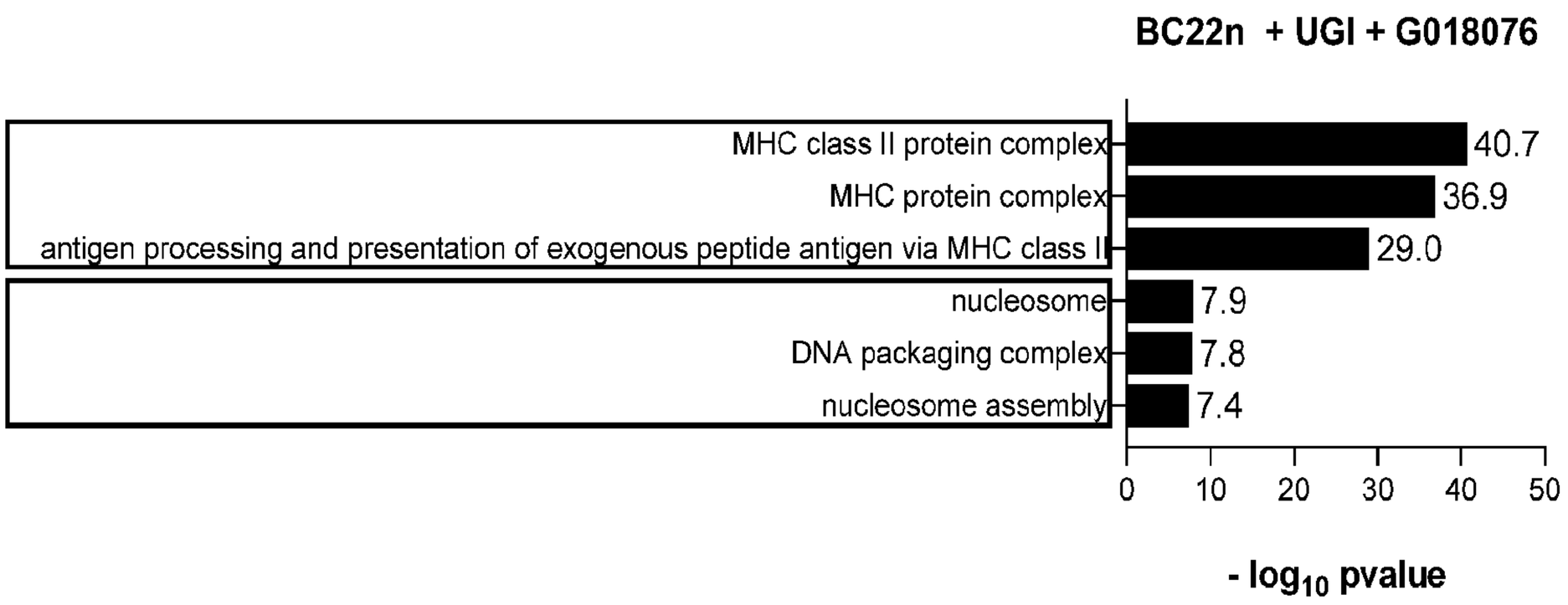


FIG. 14B

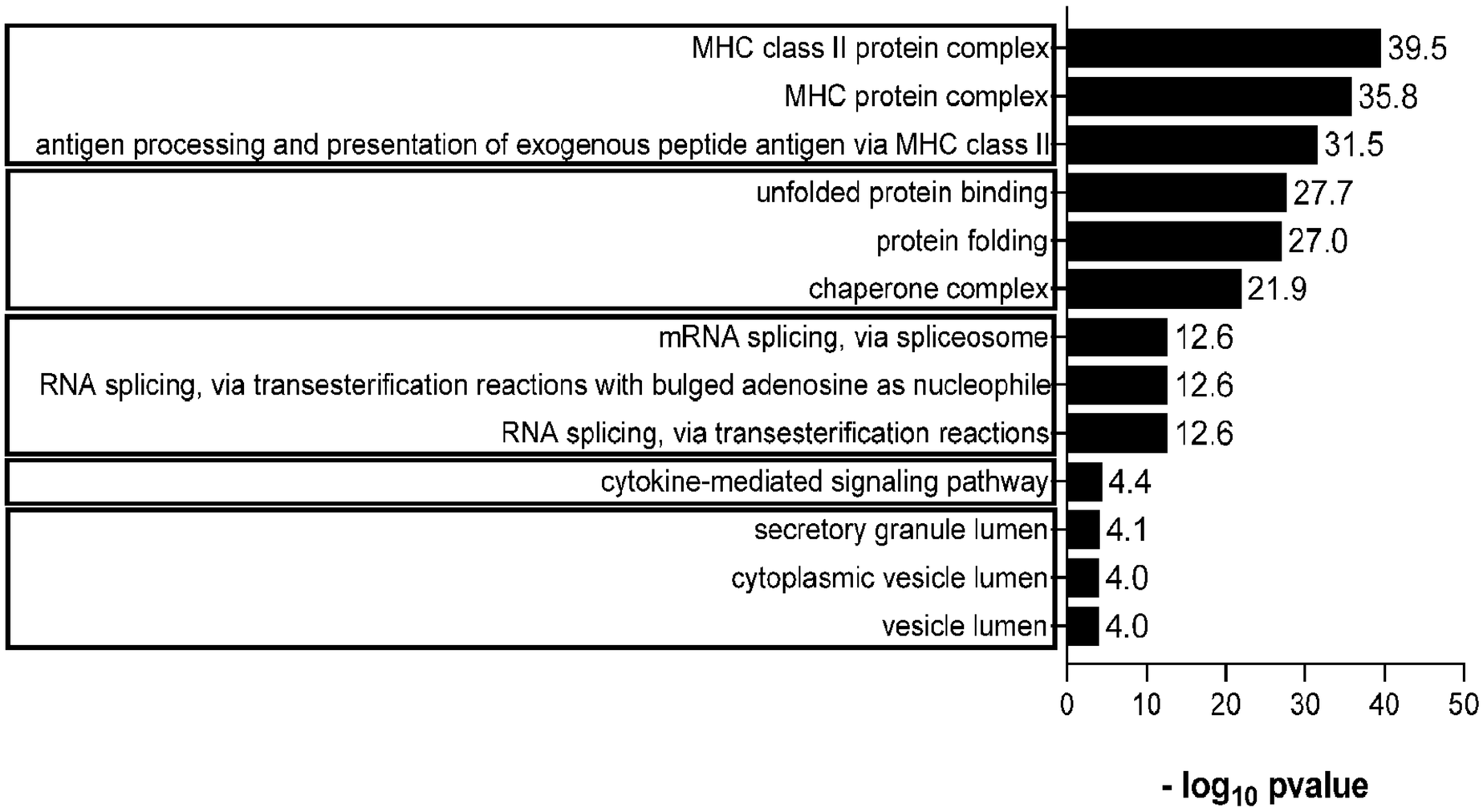


FIG. 15A

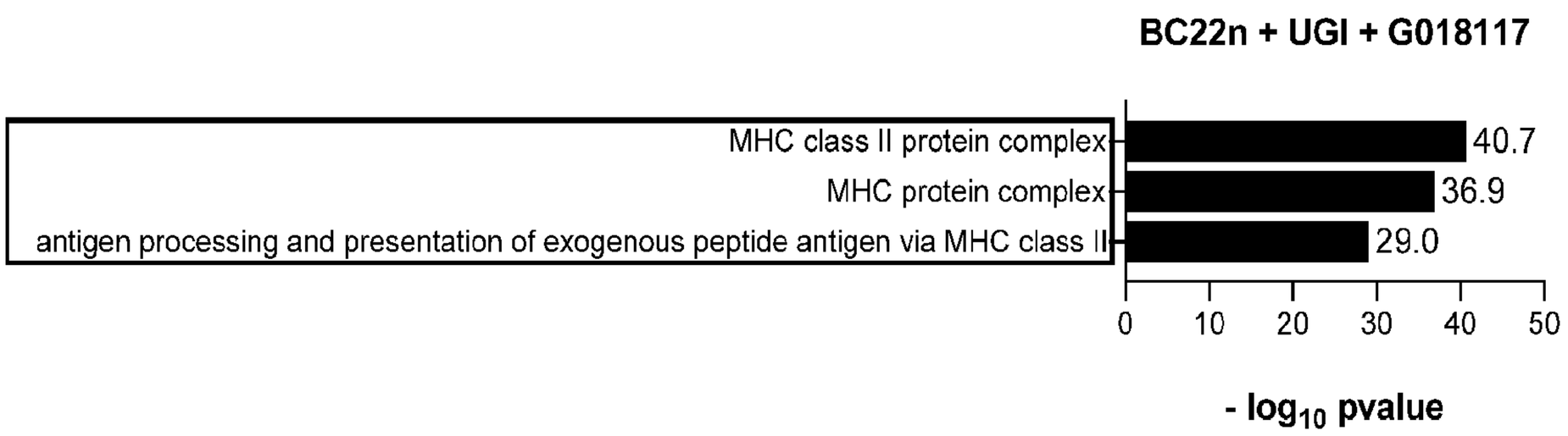


FIG. 15B

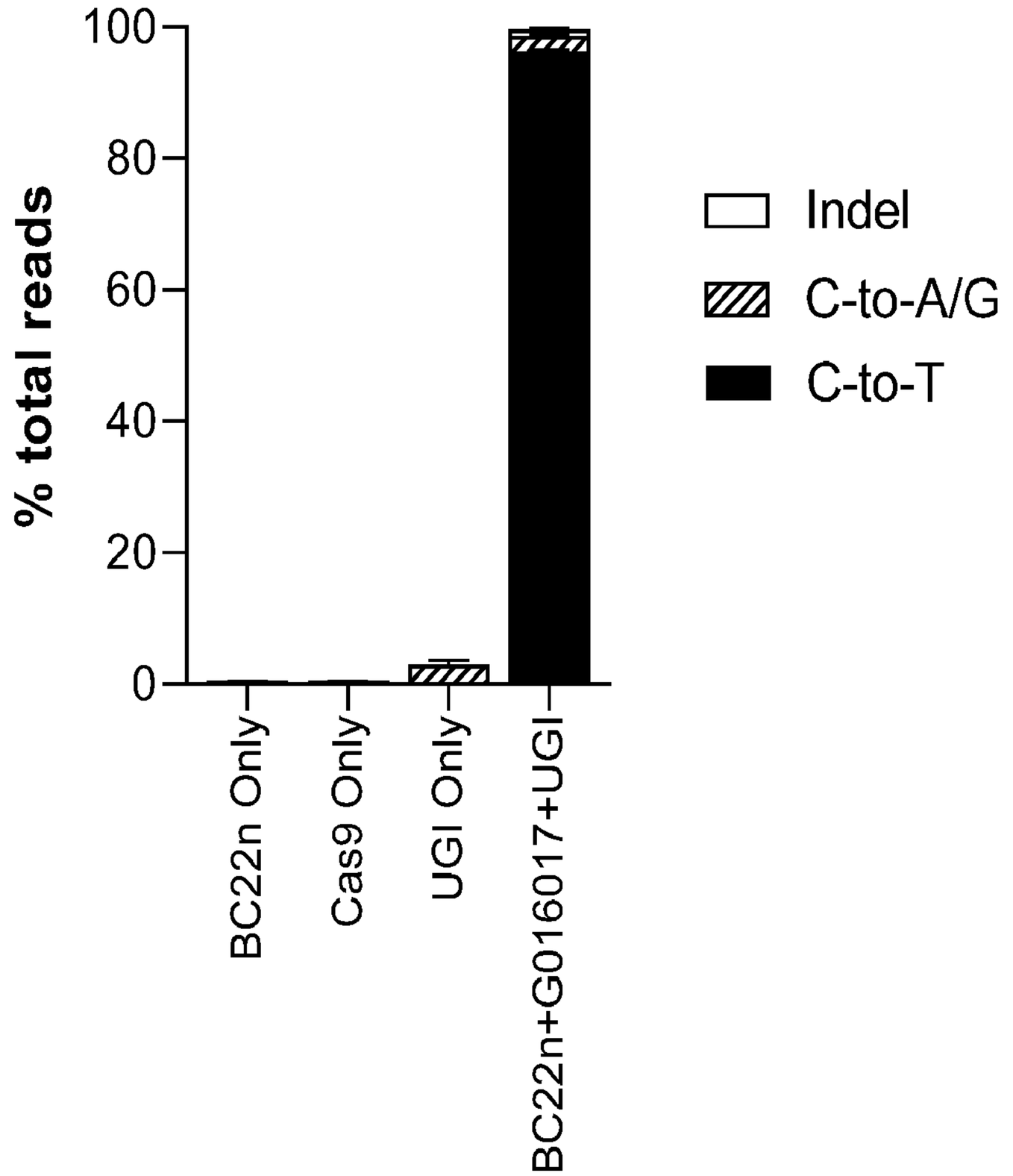

100
80
60
40
20
0
% total reads
Indel
C-to-A/G
C-to-T
BC22n Only
Cas9 Only
UGI Only
BC22n+G016017+UGI

FIG. 16A

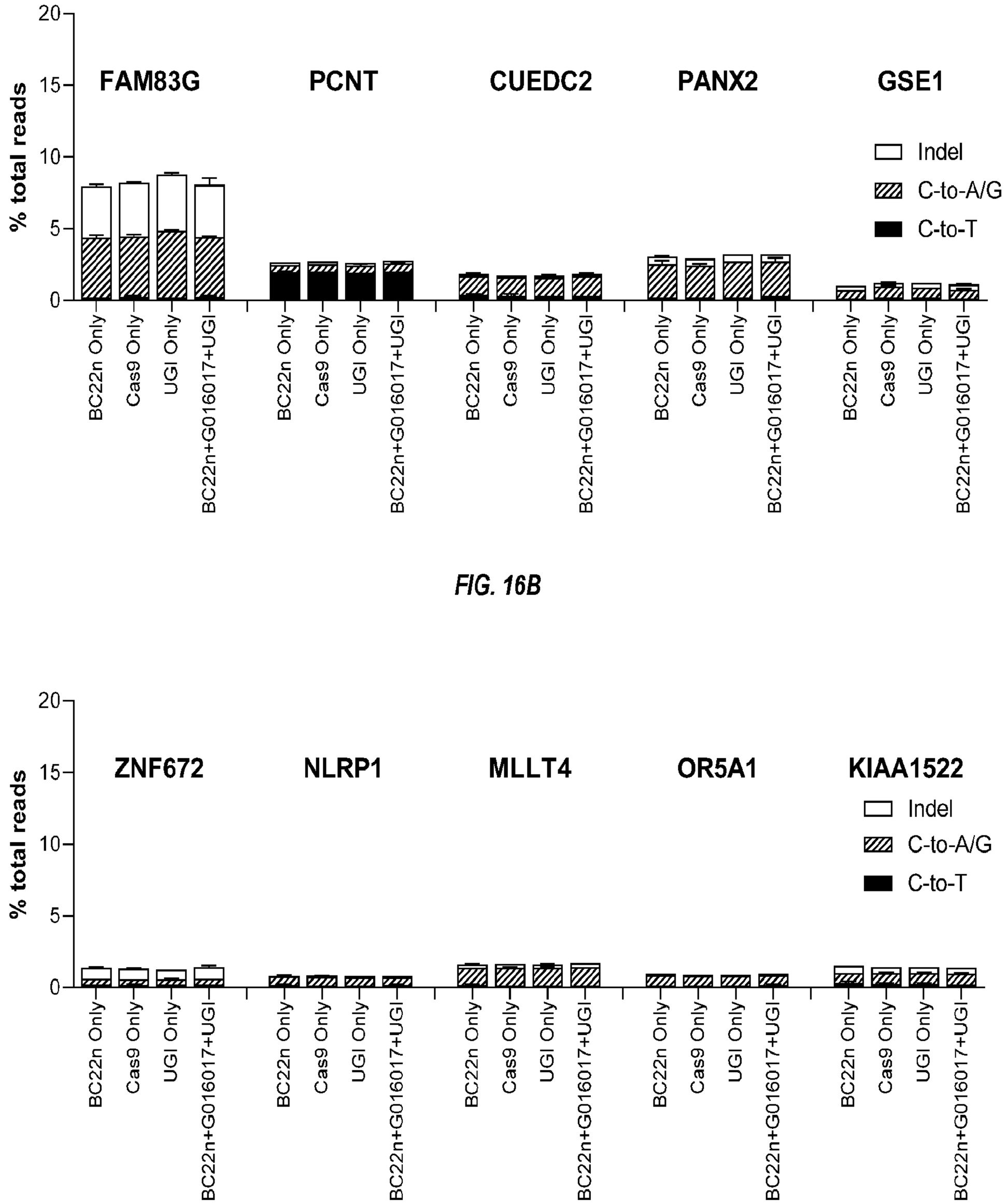


FIG. 16B

FIG. 16C

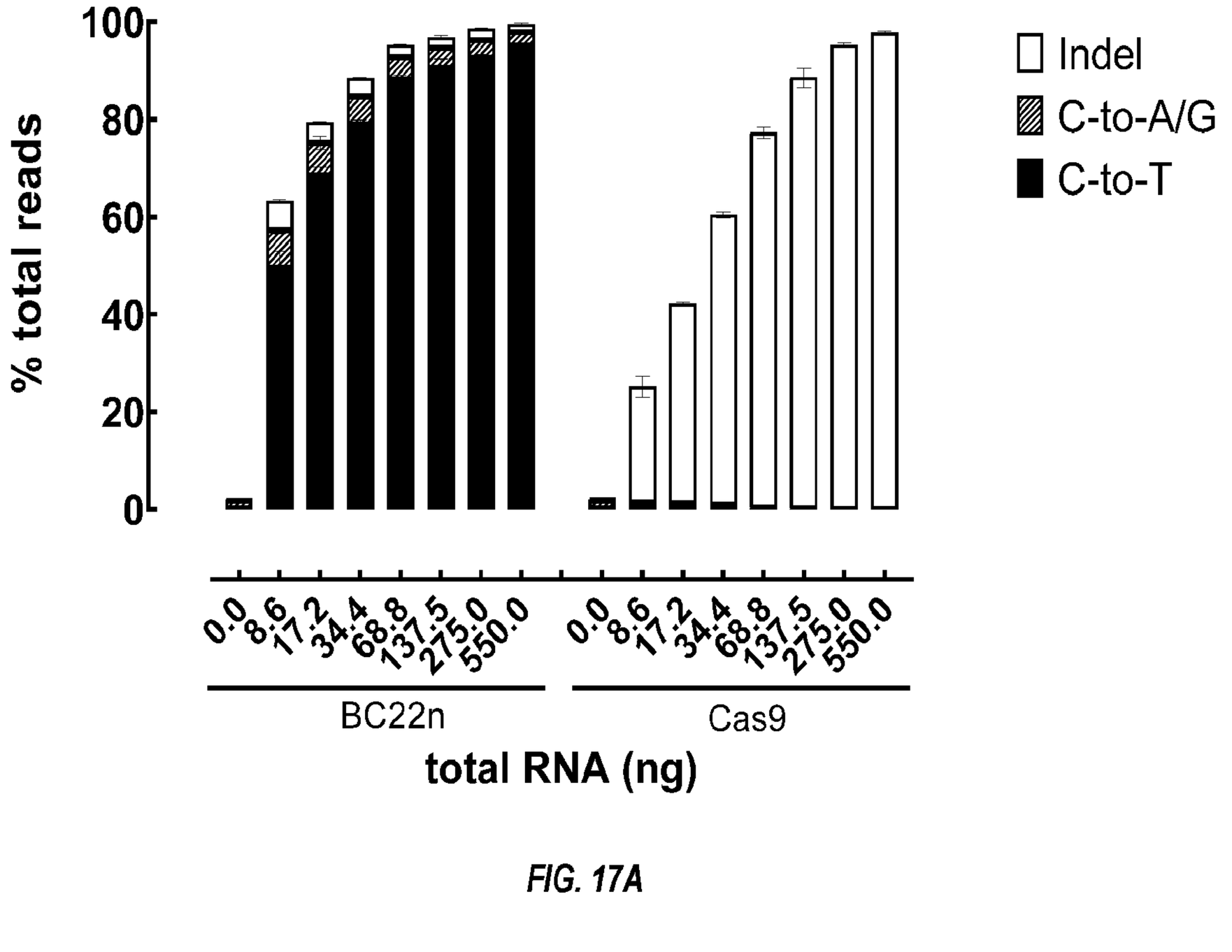


FIG. 17A

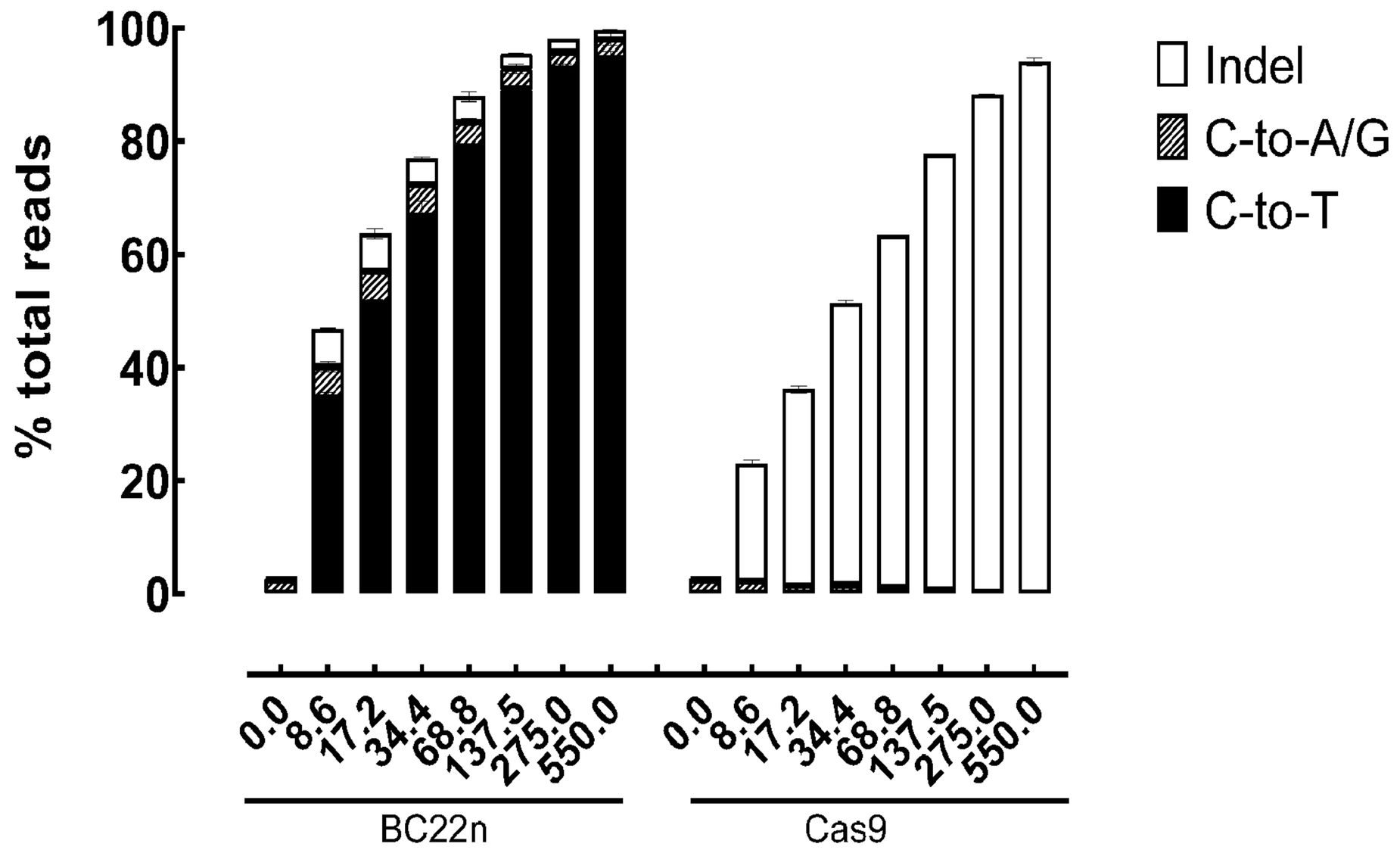


FIG. 17B

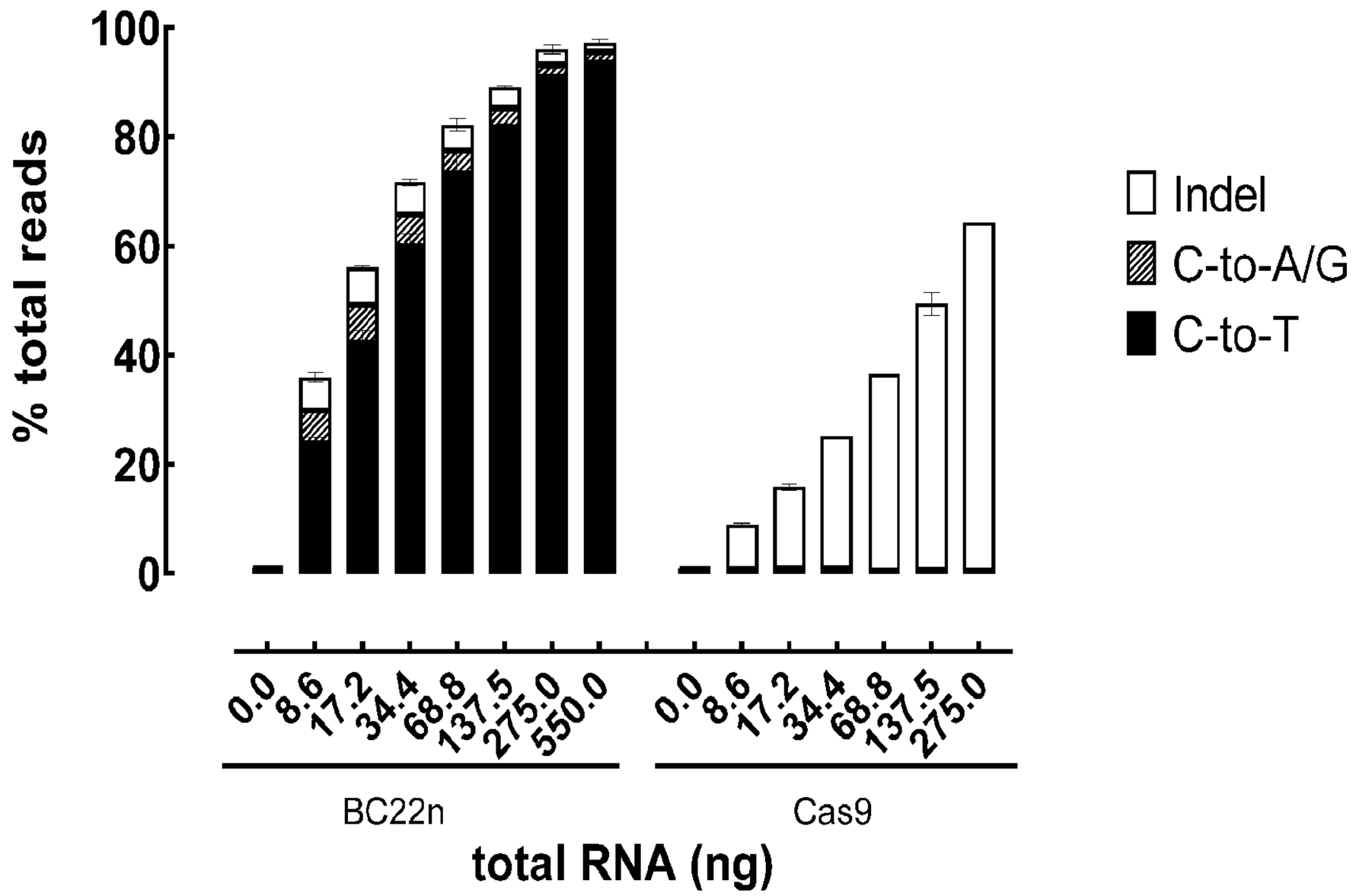


FIG. 17C

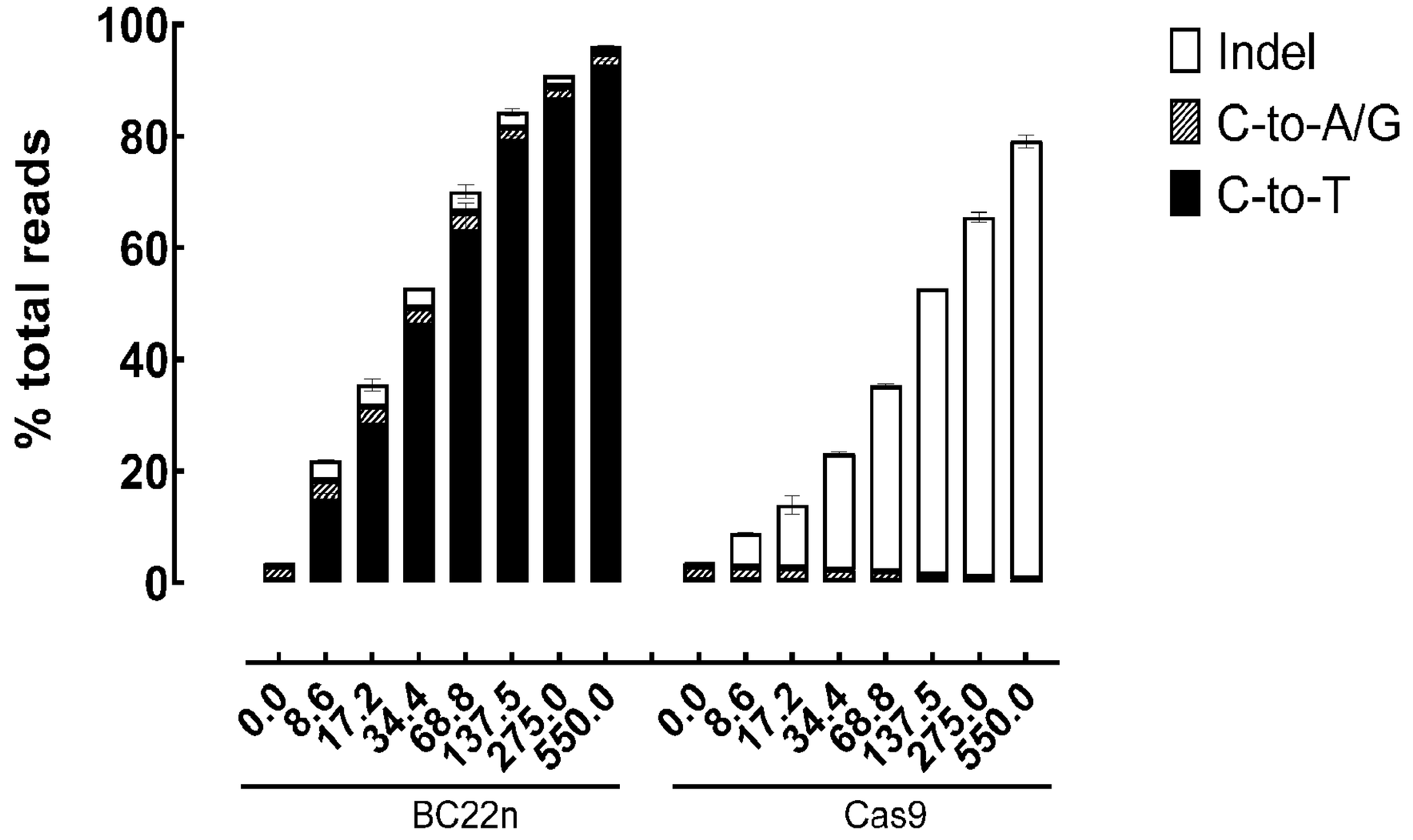


FIG. 17D

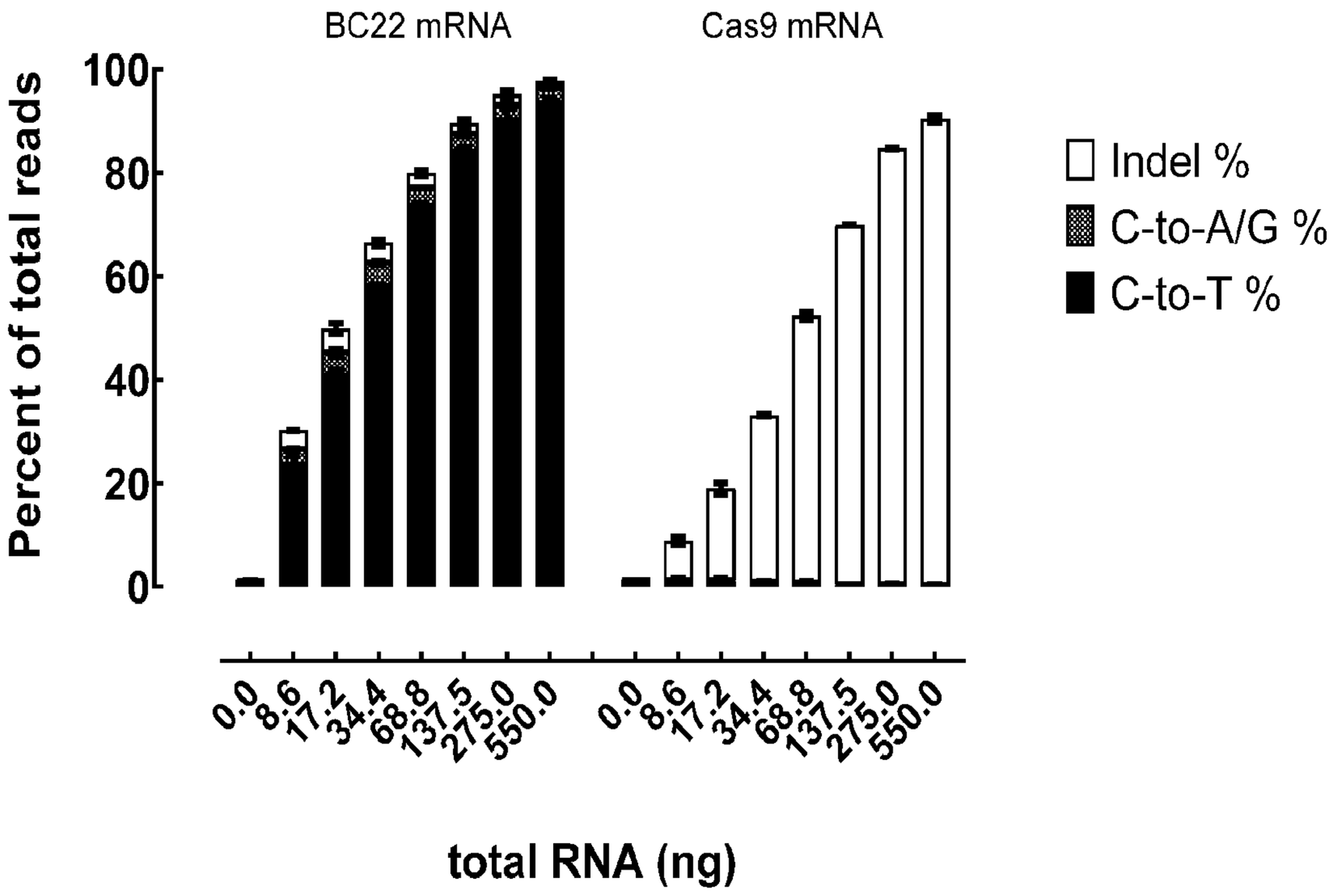

BC22 mRNA
Cas9 mRNA
Percent of total reads
100
80
60
40
20
0
Indel %
C-to-A/G %
C-to-T %
0.0
8.6
17.2
34.4
68.8
137.5
275.0
550.0
total RNA (ng)

Fig. 17E

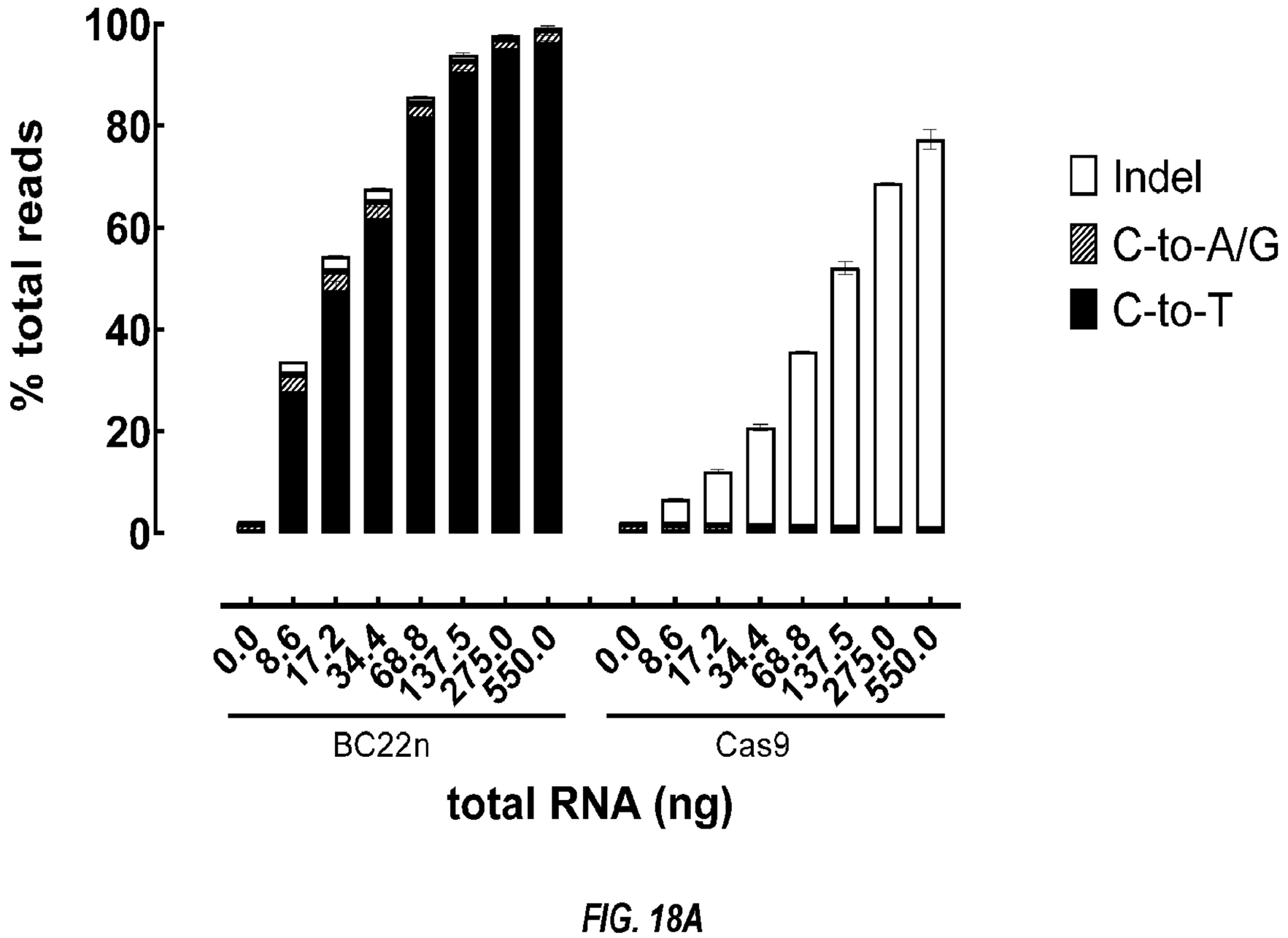


FIG. 18A

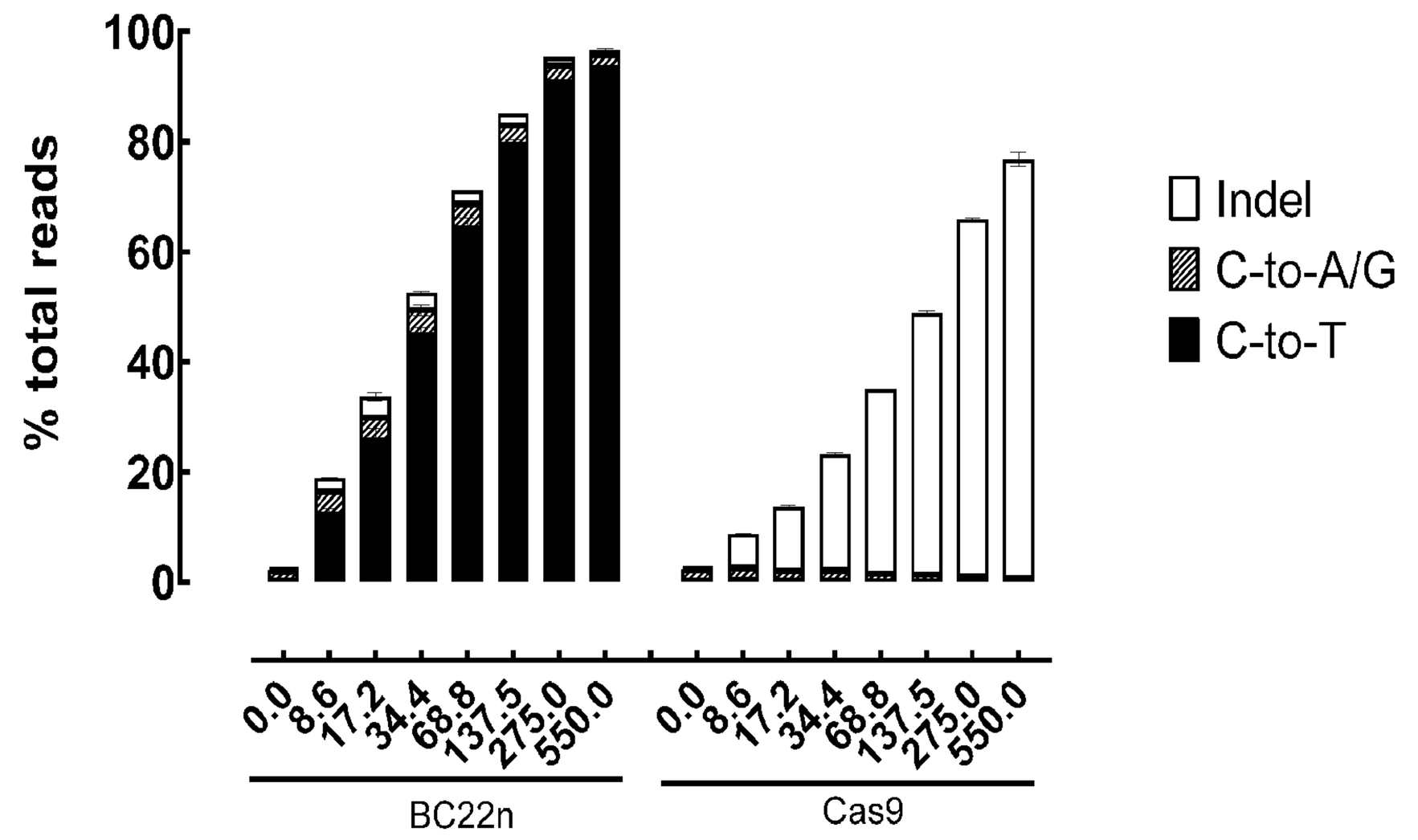


FIG. 18B

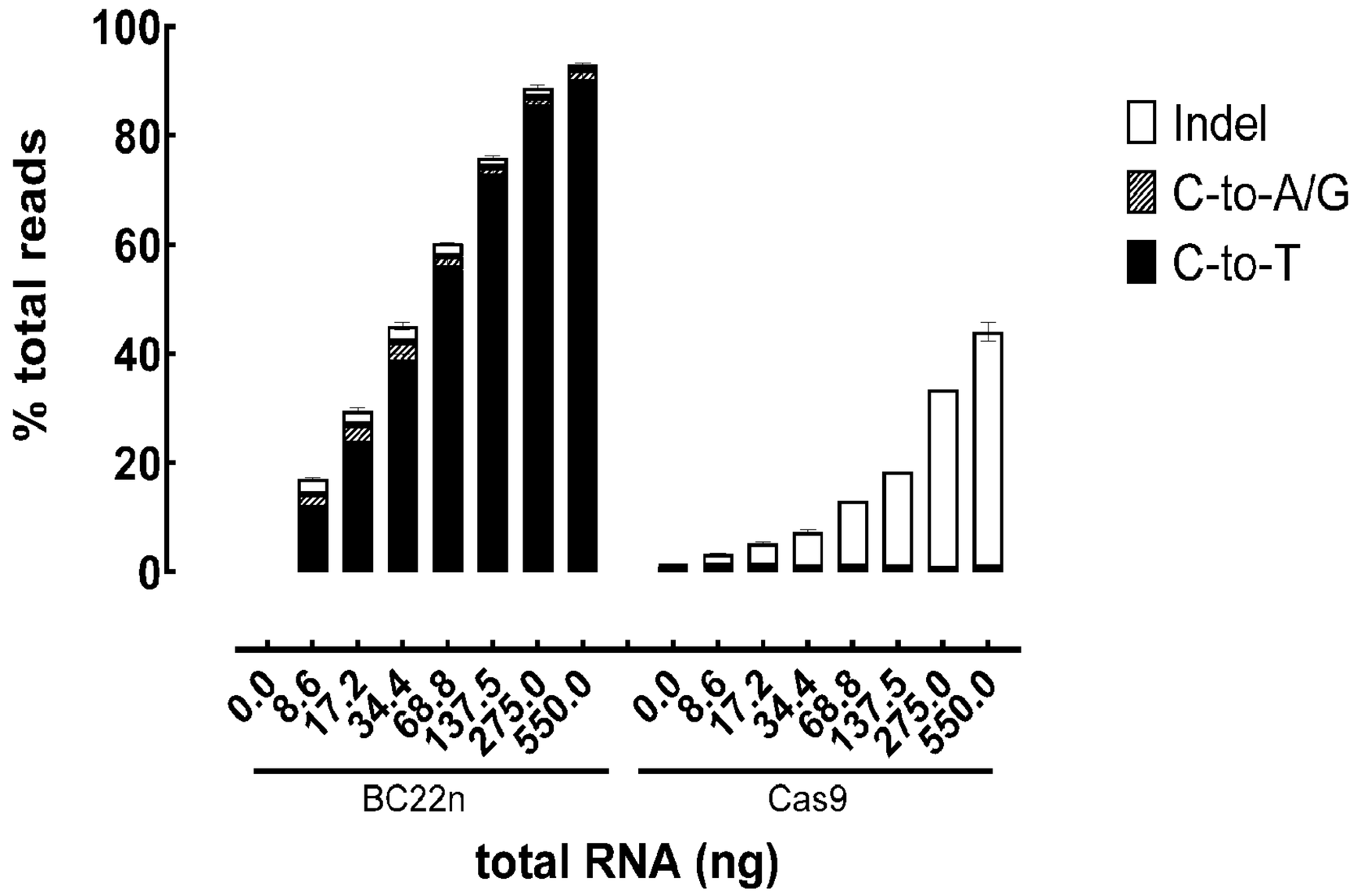


FIG. 18C

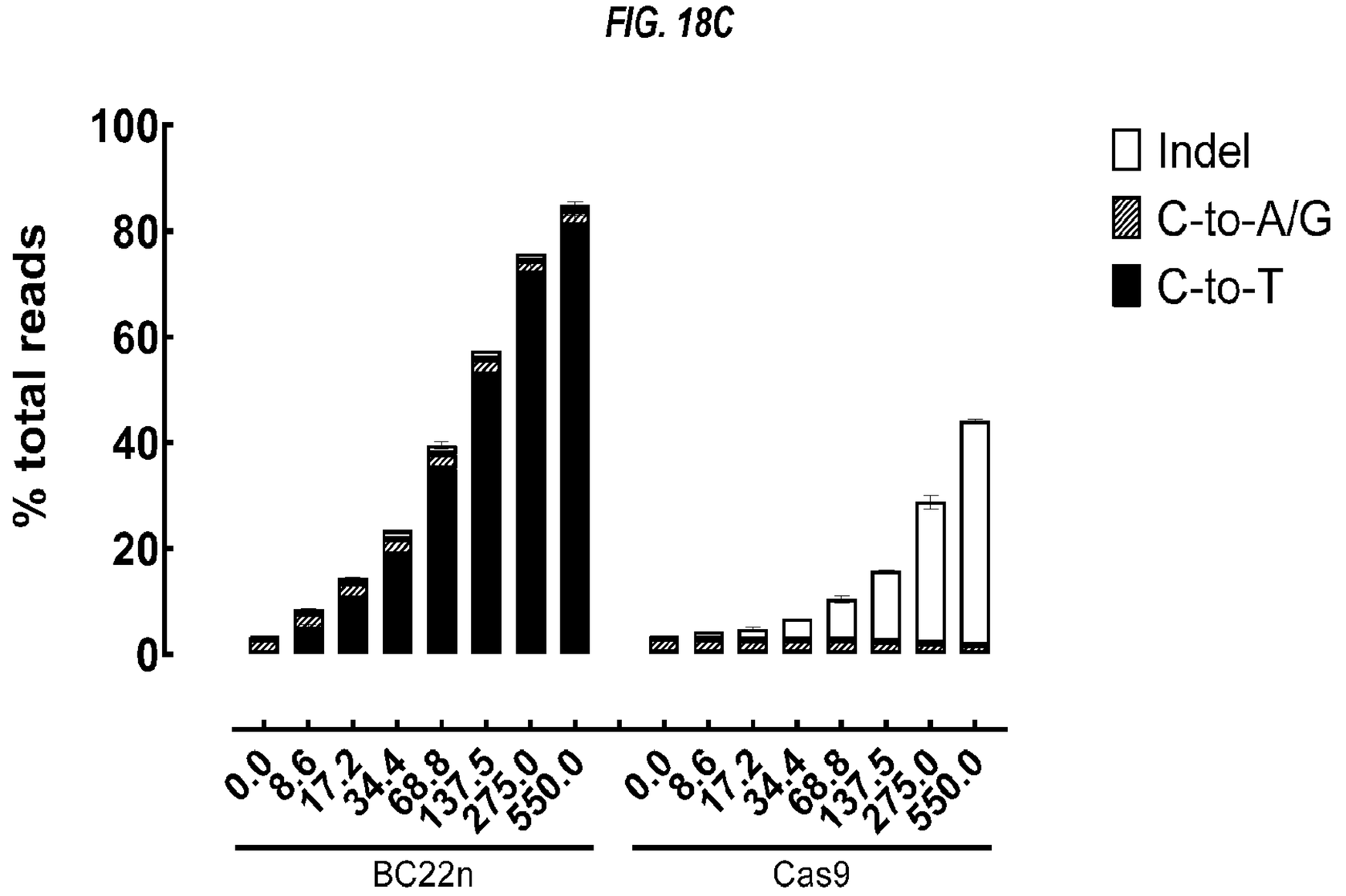


FIG. 18D

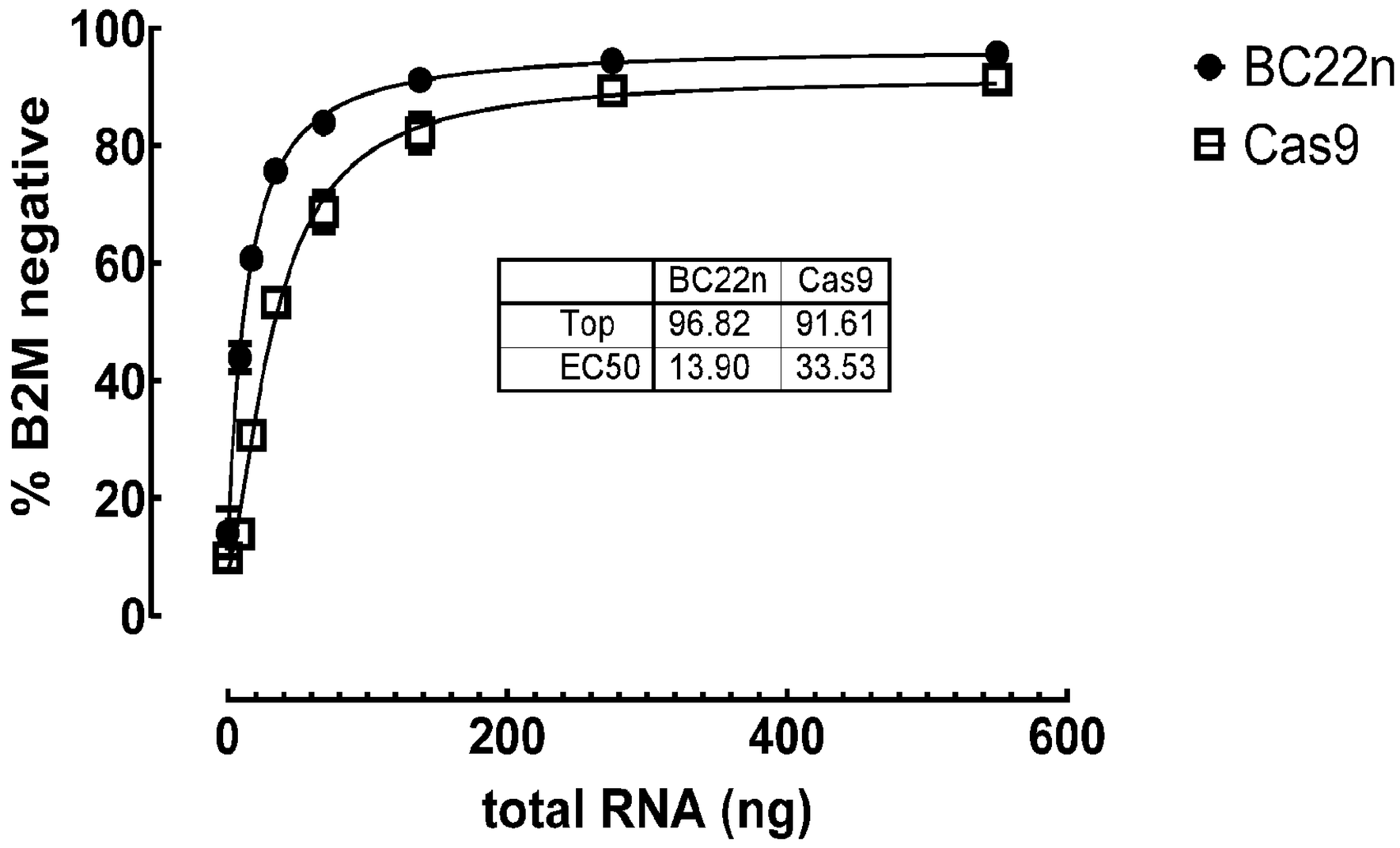


FIG. 19A

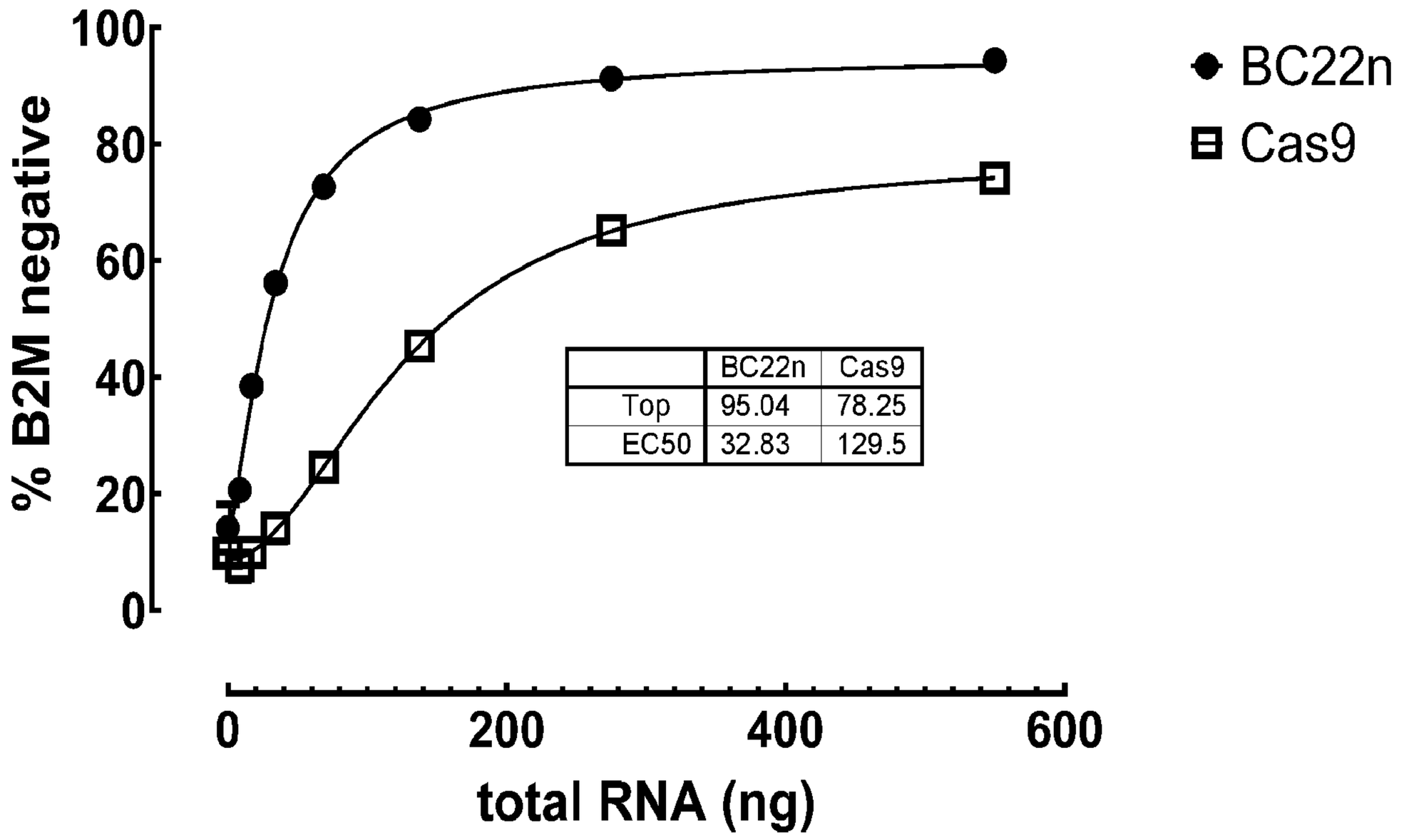


FIG. 19B

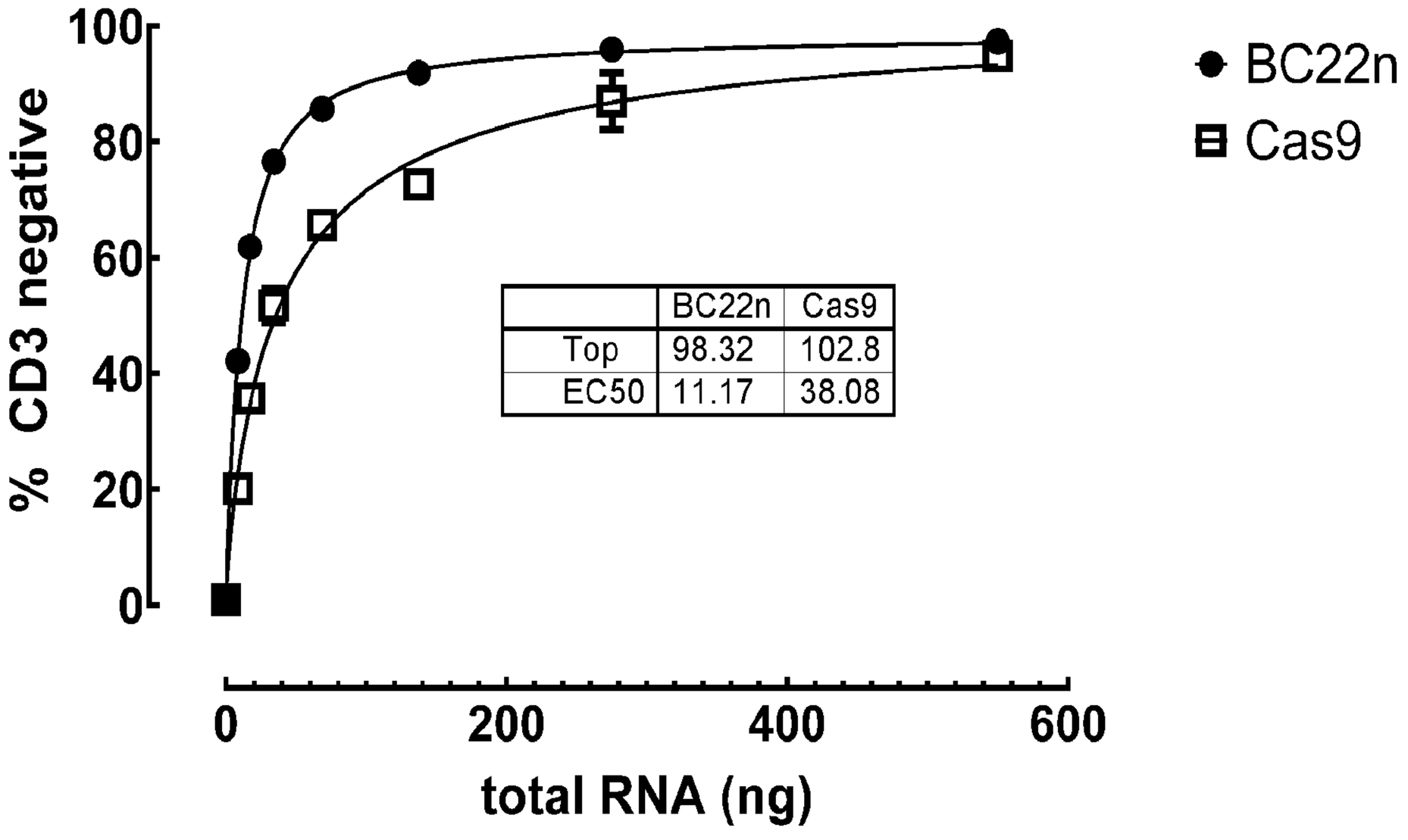


FIG. 19C

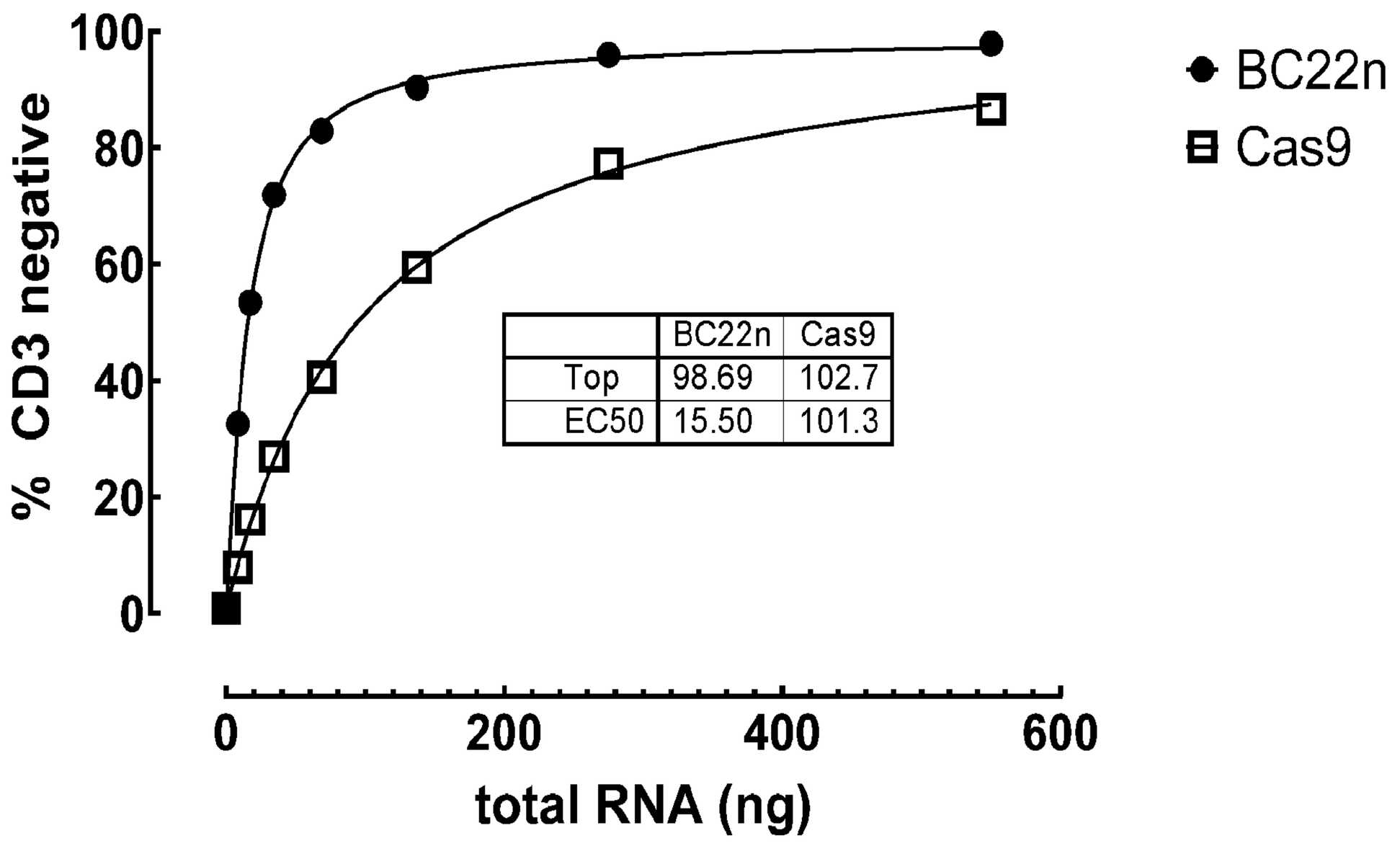
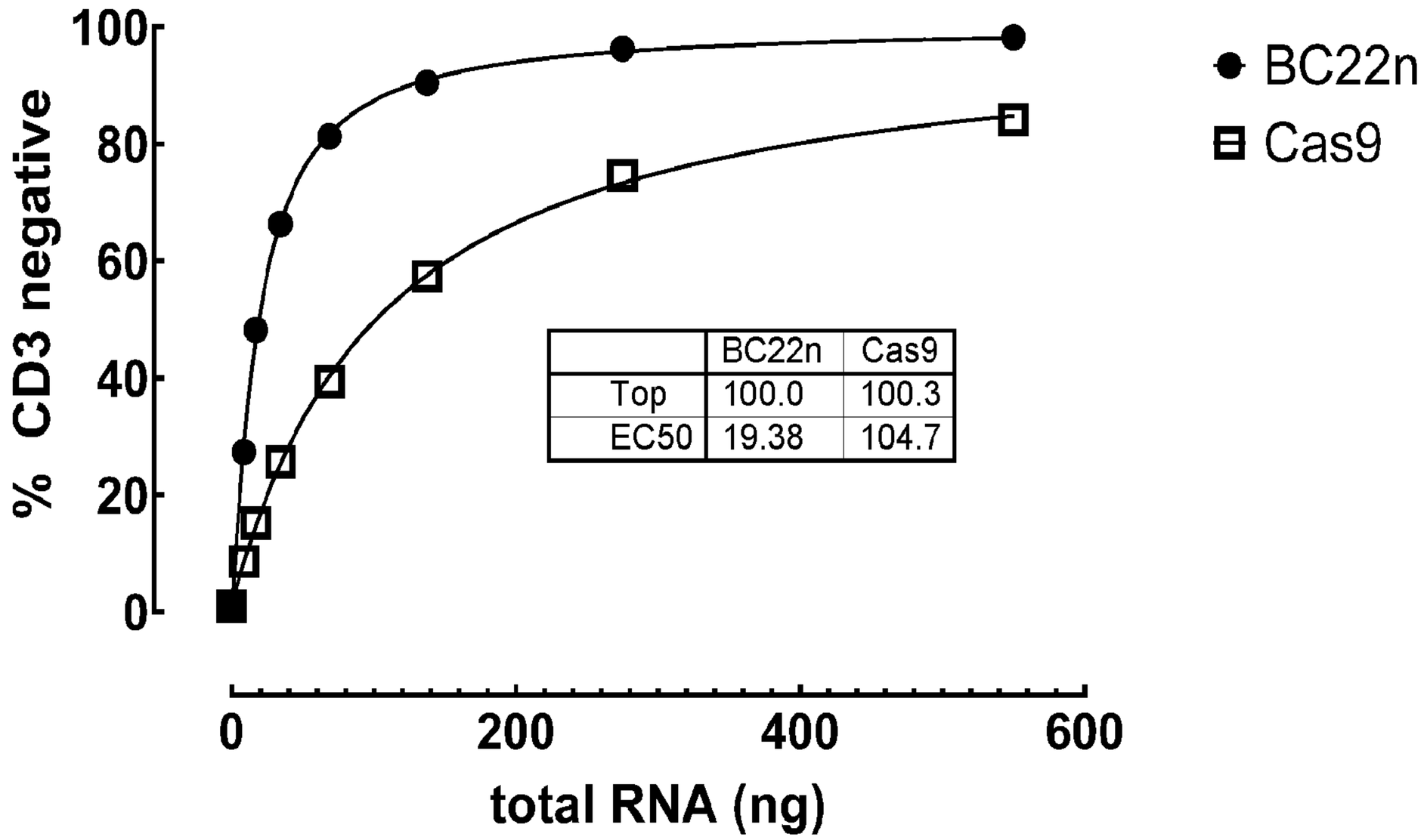
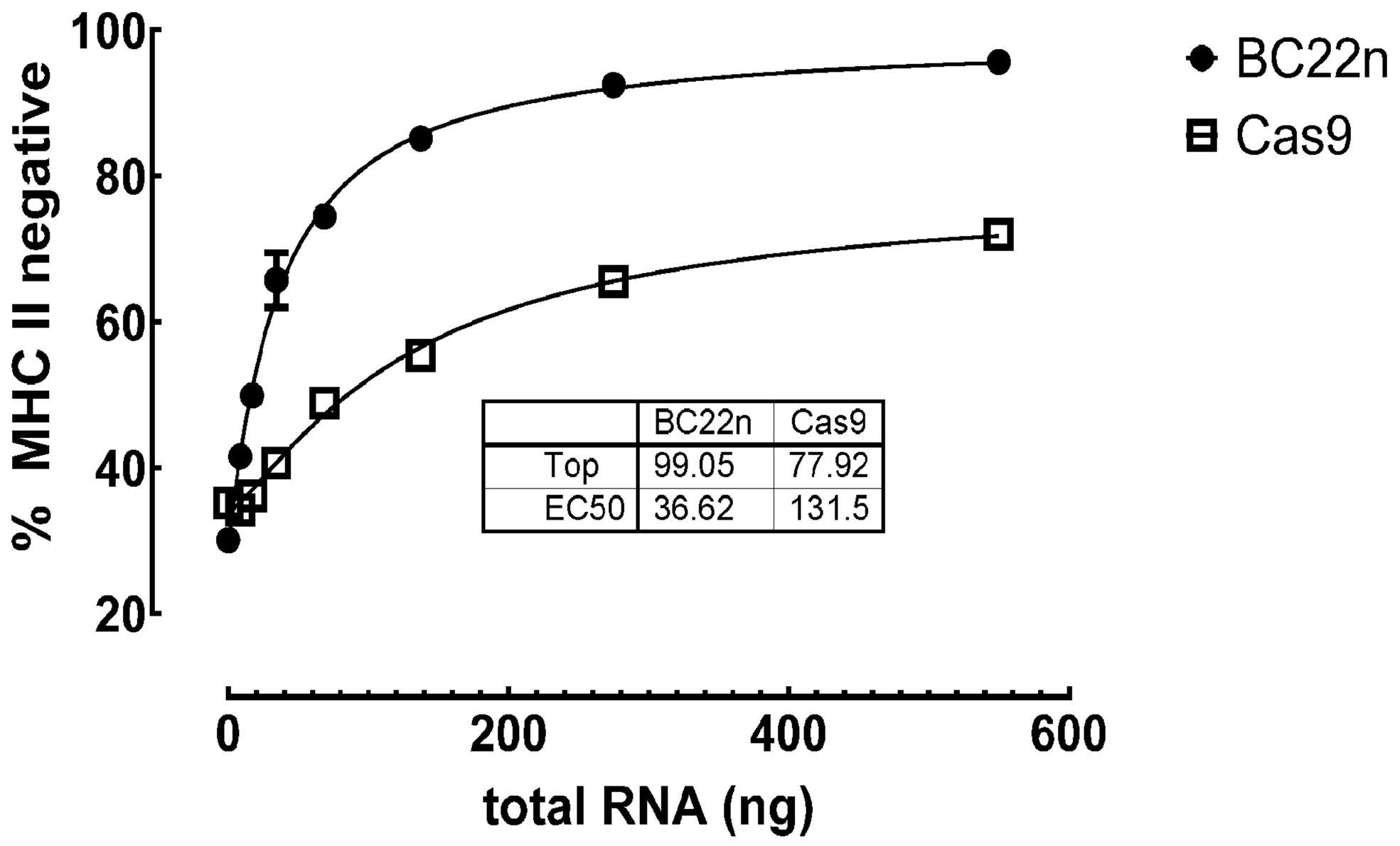


FIG. 19D



	BC22n	Cas9
Top	100.0	100.3
EC50	19.38	104.7

FIG. 19E



	BC22n	Cas9
Top	99.05	77.92
EC50	36.62	131.5

FIG. 19F

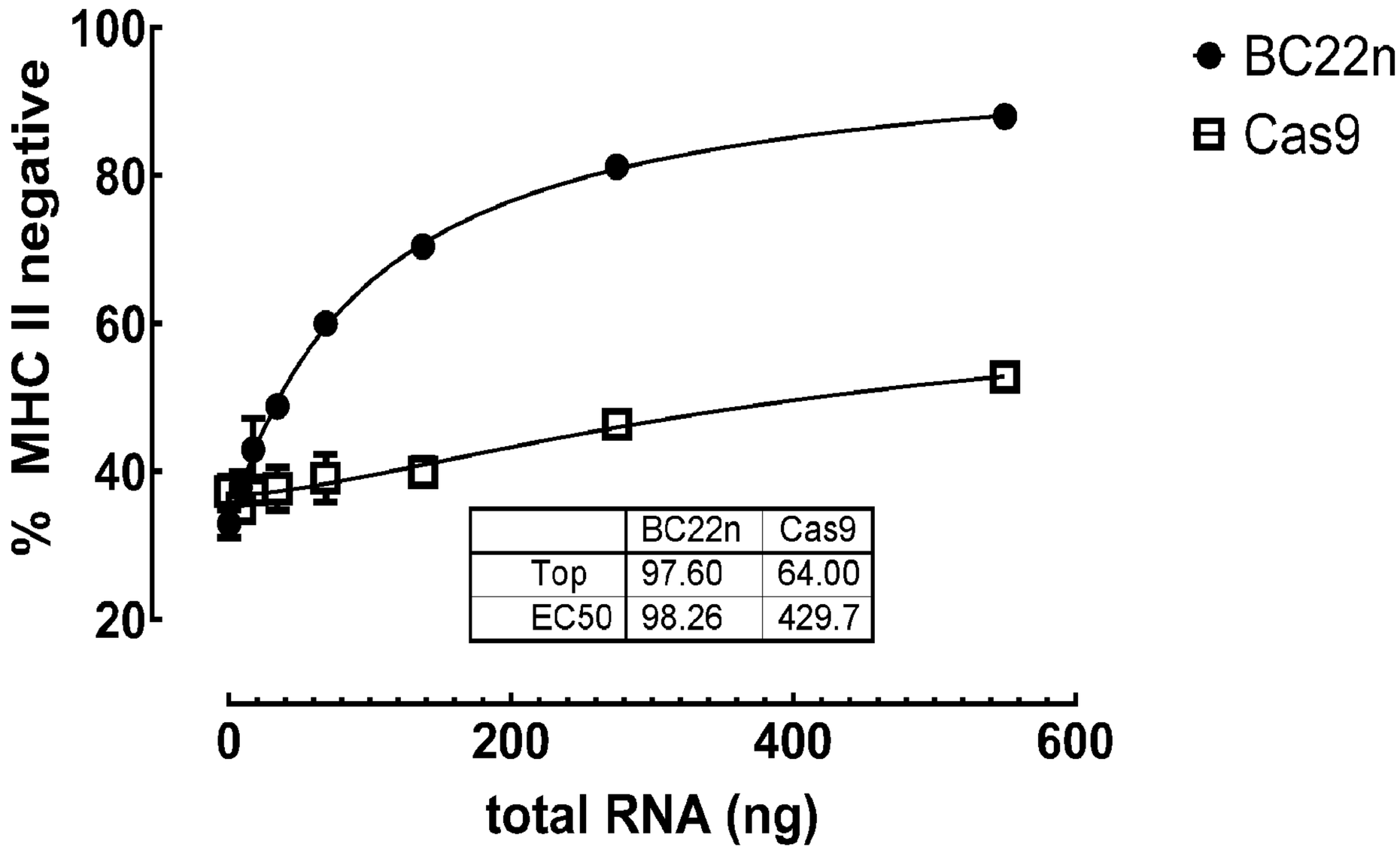


FIG. 19G

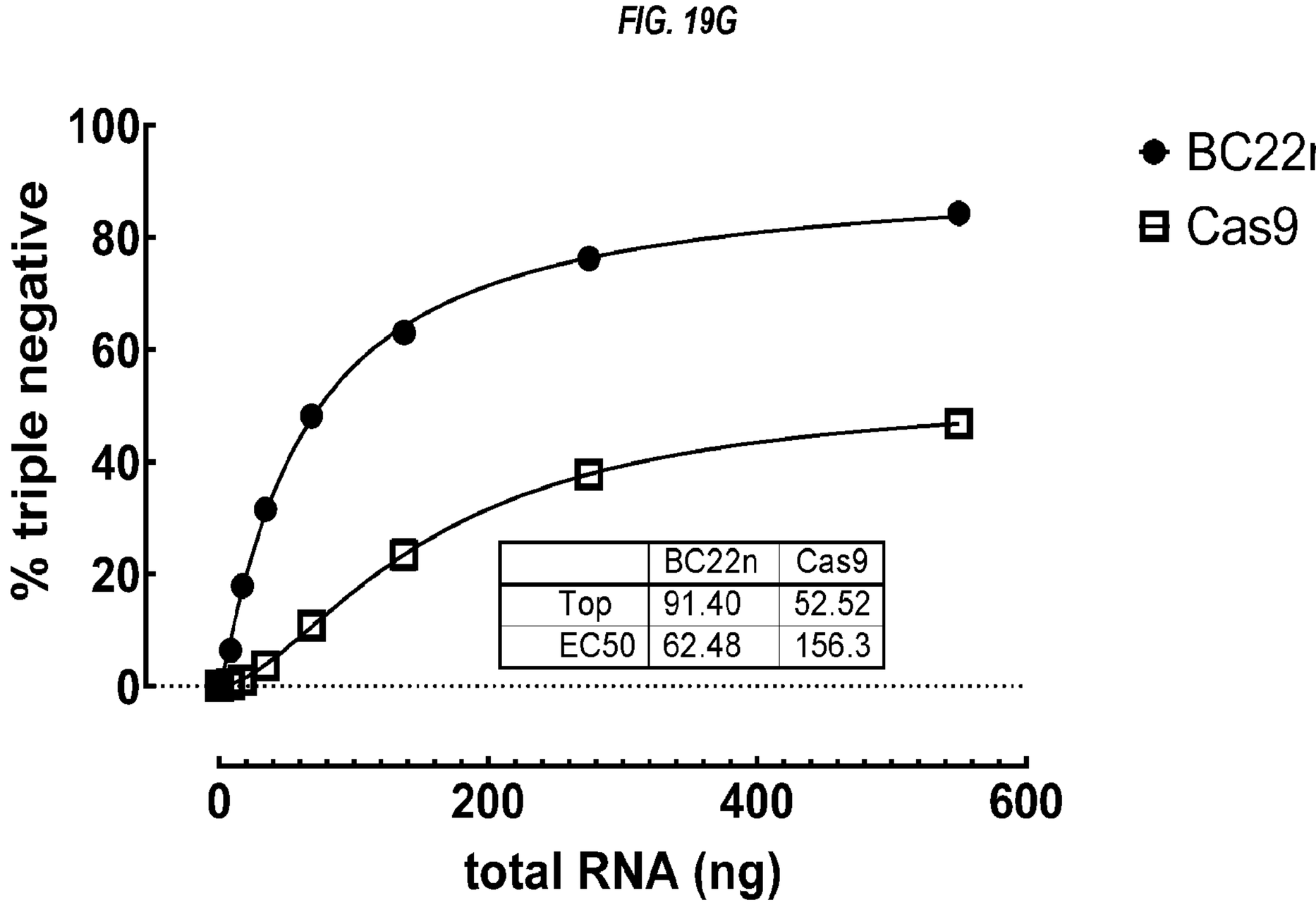


FIG. 19H

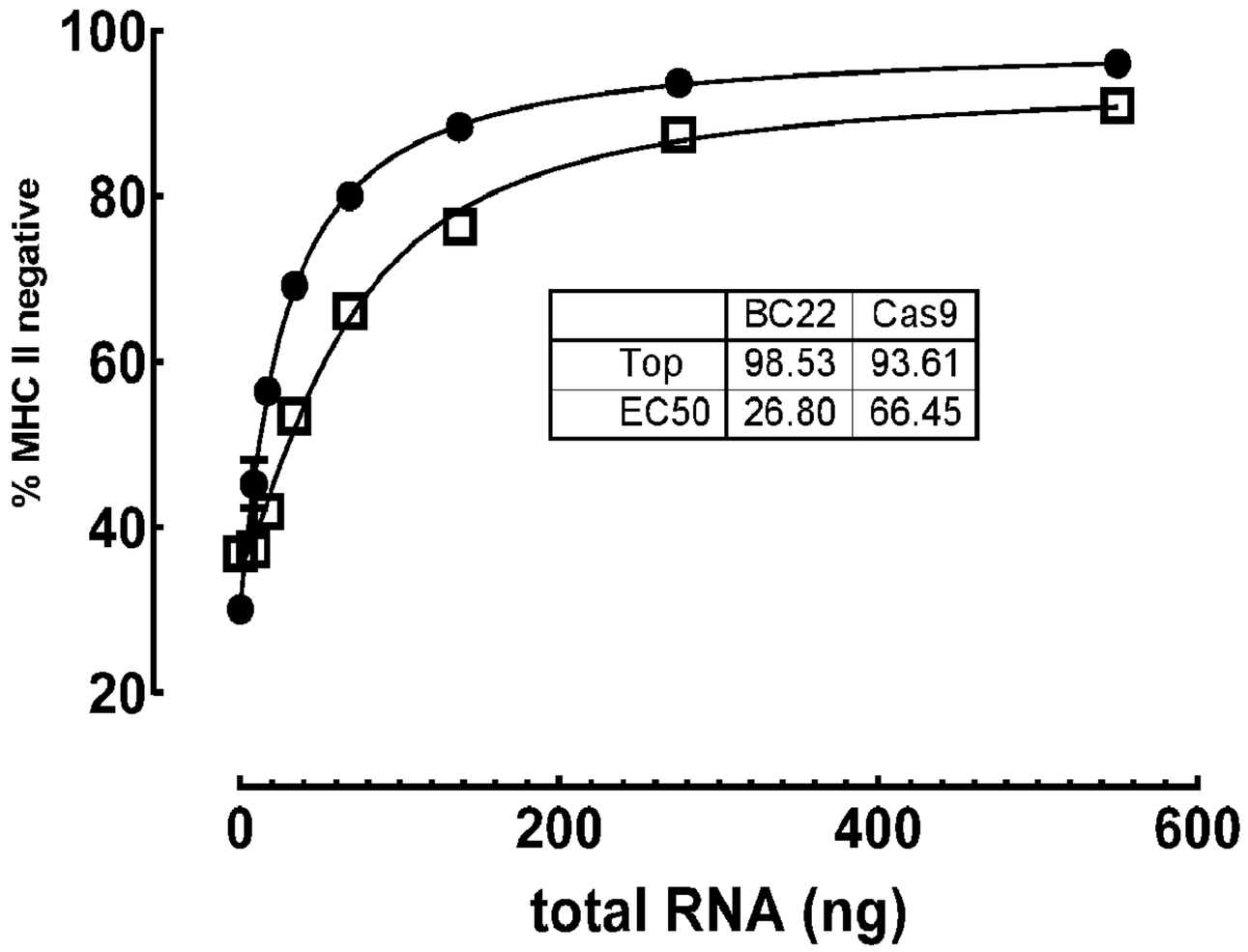

100
80
60
40
20
% MHC II negative
0
200
400
600
total RNA (ng)
BC22
Cas9
BC22 Cas9
Top 98.53 93.61
EC50 26.80 66.45

FIG. 19I

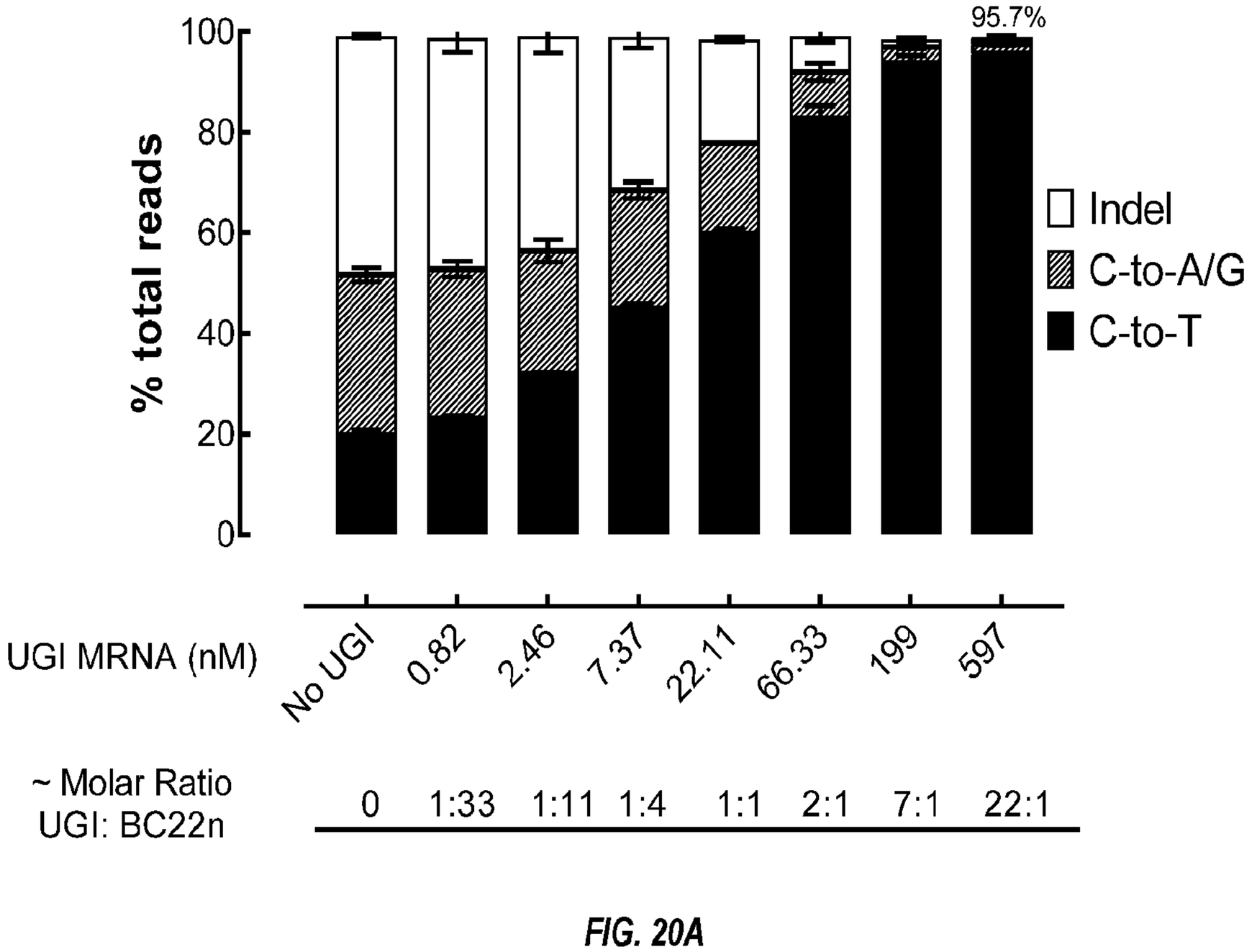


FIG. 20A

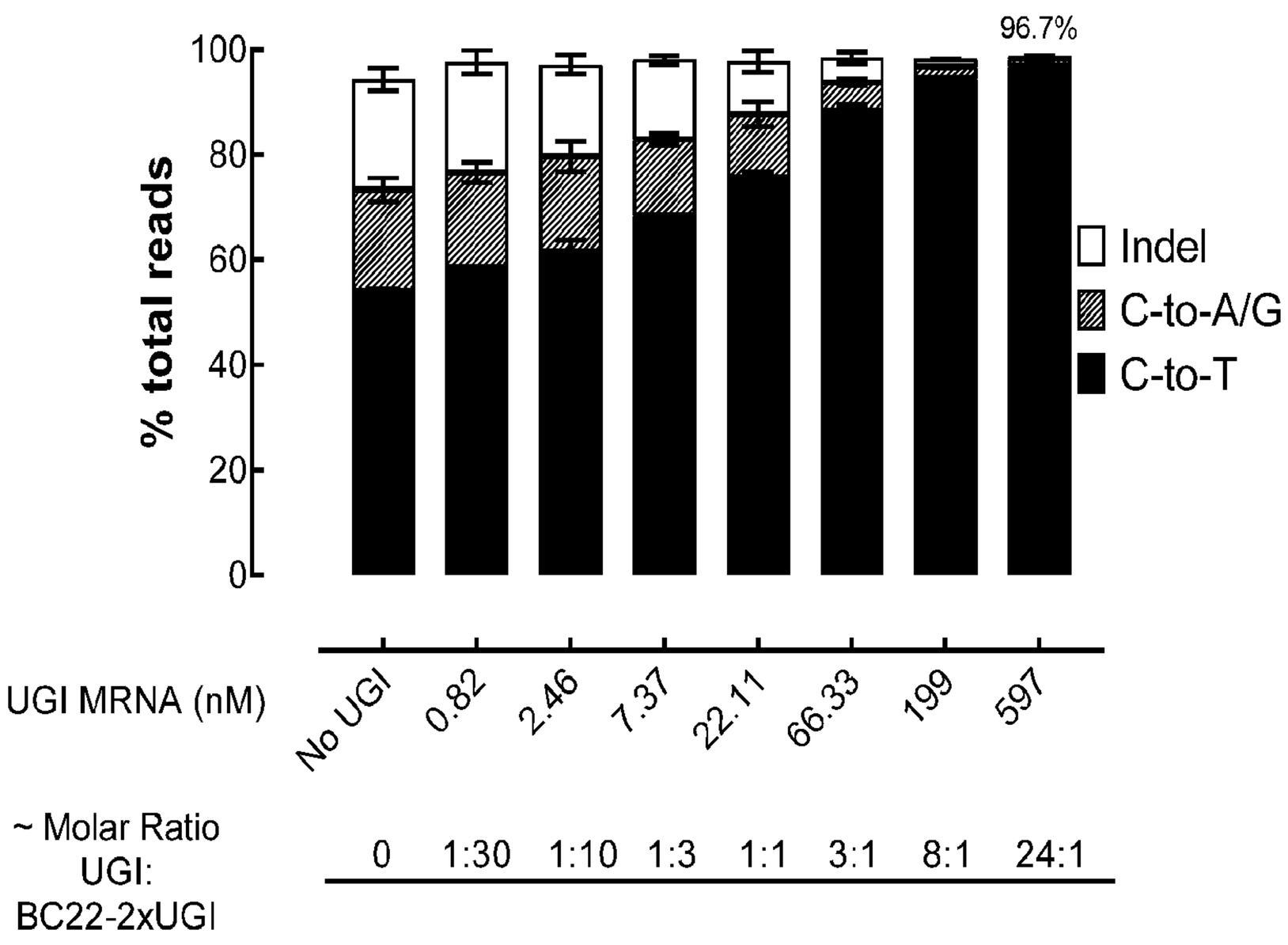


FIG. 20B

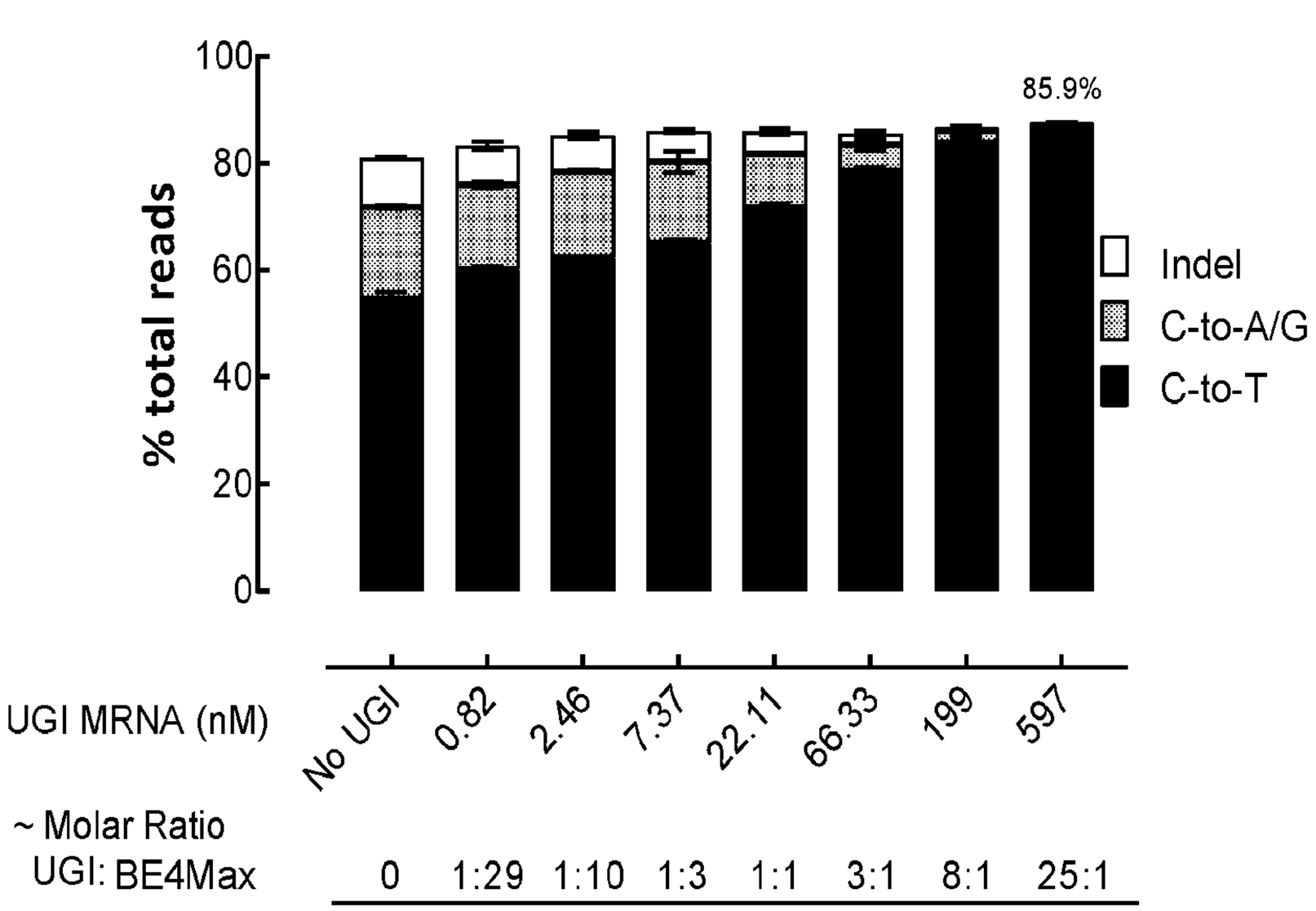

100
80
60
40
20
0
% total reads
85.9%
Indel
C-to-A/G
C-to-T
UGI MRNA (nM)
No UGI
0.82
2.46
7.37
22.11
66.33
199
597
~ Molar Ratio
UGI: BE4Max
0
1:29
1:10
1:3
1:1
3:1
8:1
25:1

FIG. 20C

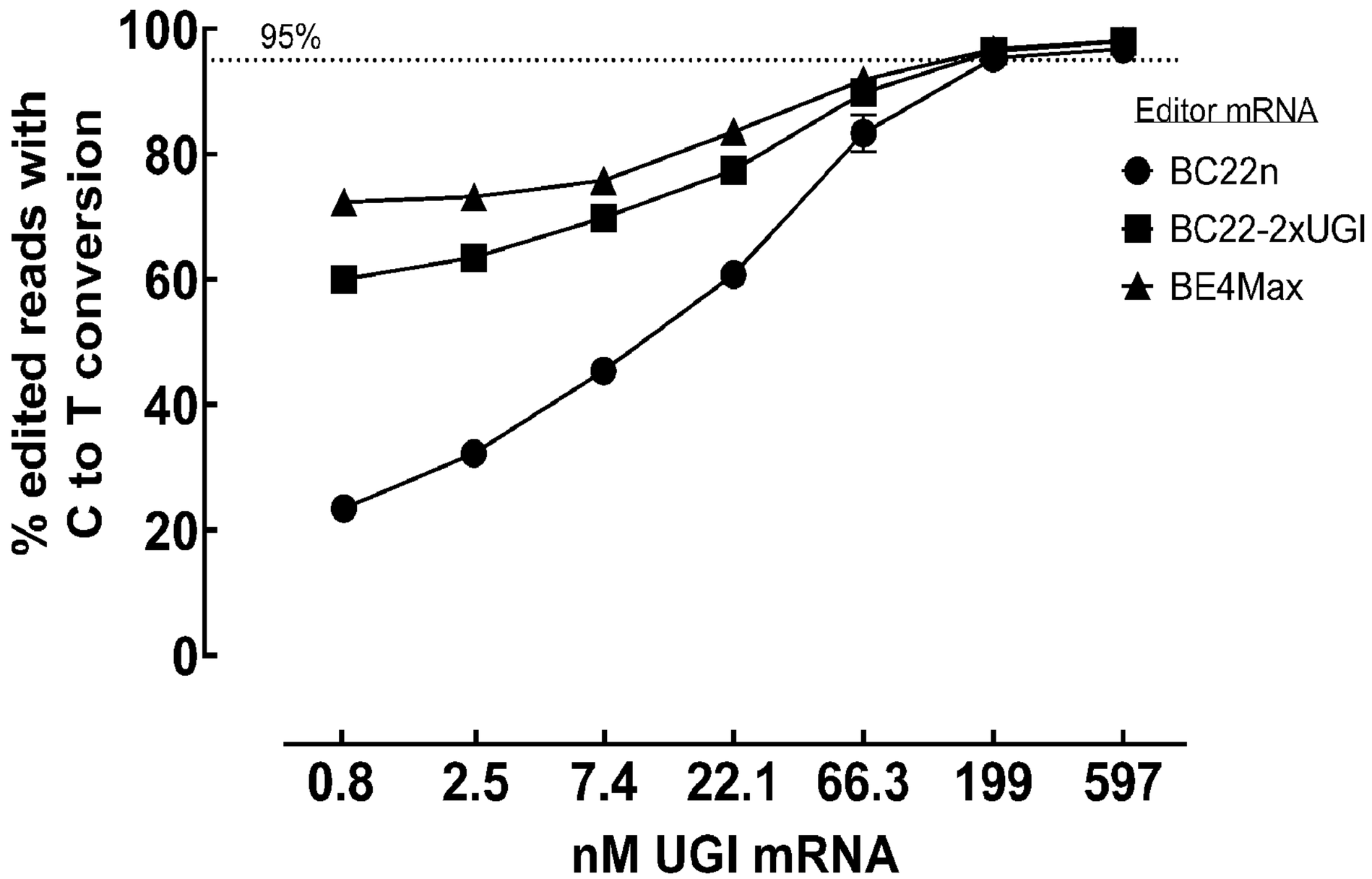

100
95%
80
60
40
20
0
% edited reads with C to T conversion
Editor mRNA
BC22n
BC22-2xUGI
BE4Max
0.8
2.5
7.4
22.1
66.3
199
597
nM UGI mRNA

FIG. 21

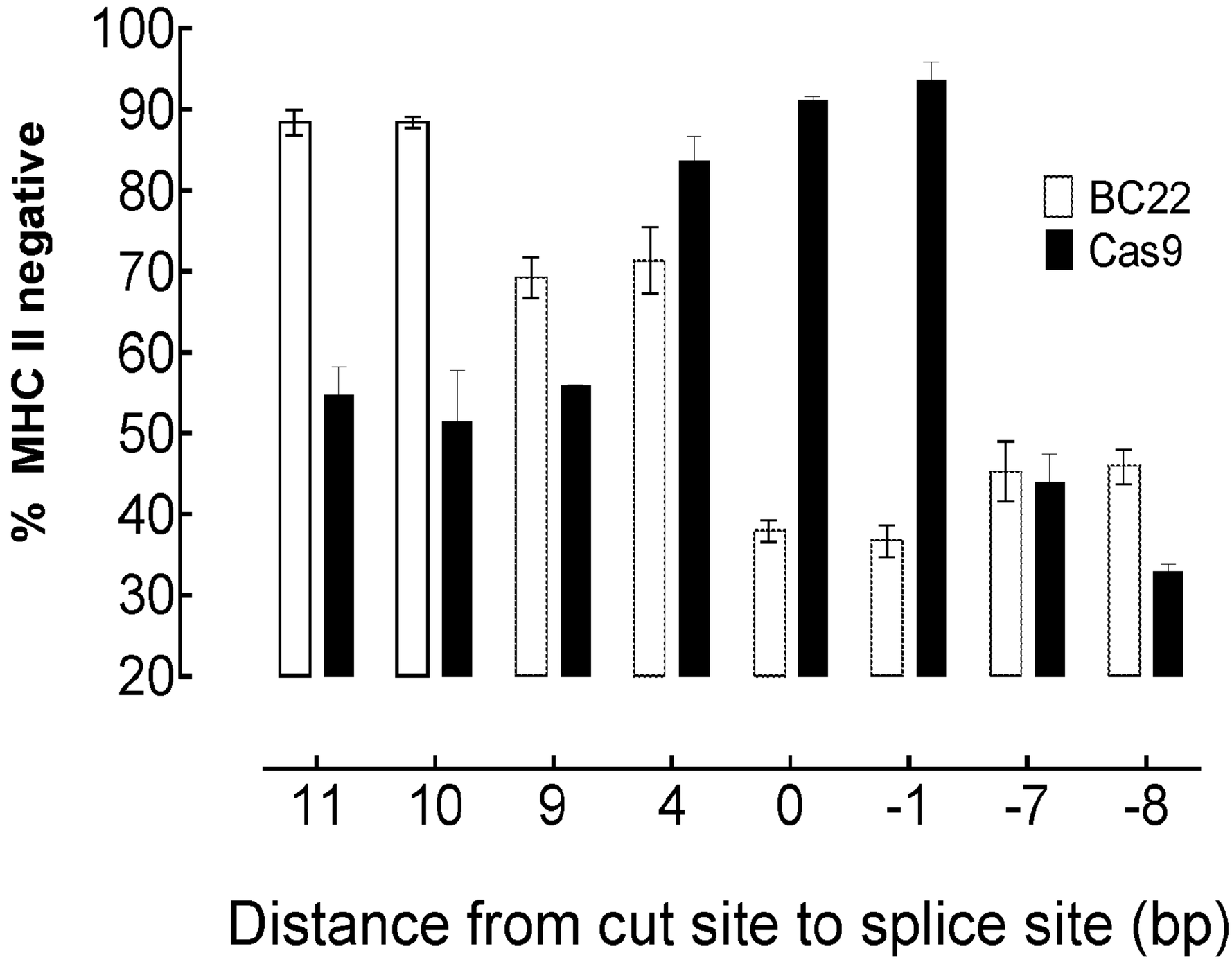

% MHC II negative
100
90
80
70
60
50
40
30
20
BC22
Cas9
11
10
9
4
0
-1
-7
-8
Distance from cut site to splice site (bp)

FIG. 22

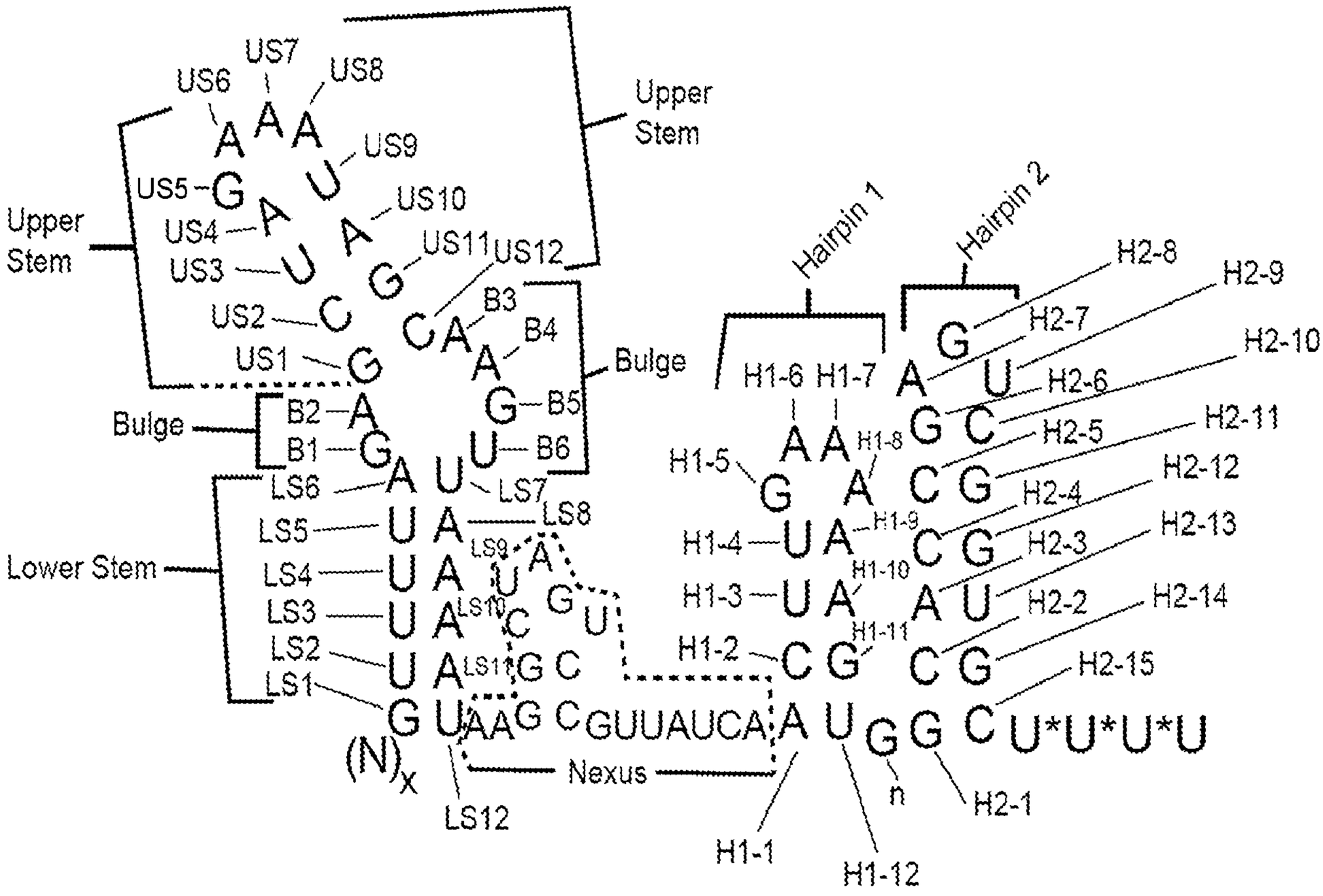


FIG. 23A

Scaffold Y/A

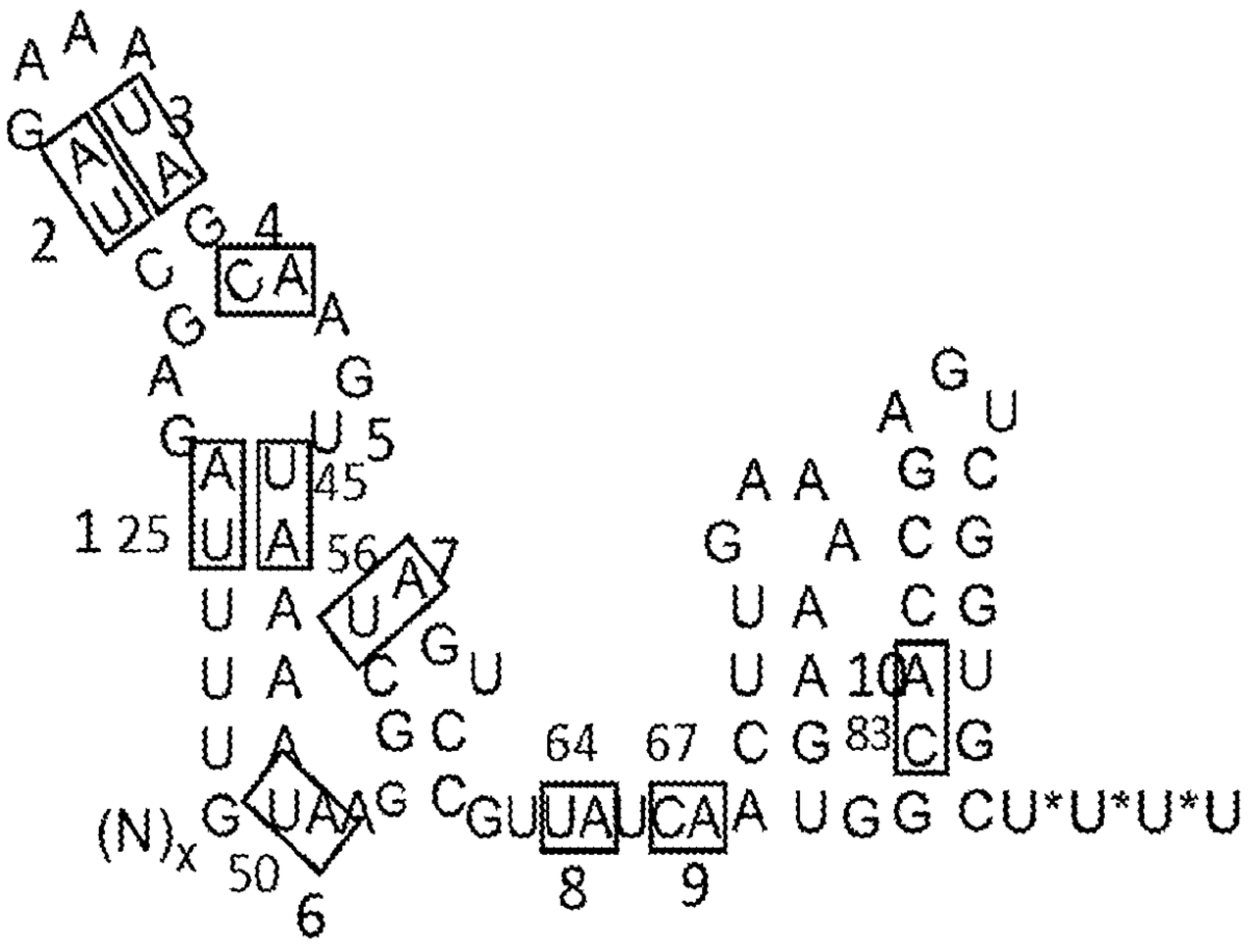


FIG. 23B

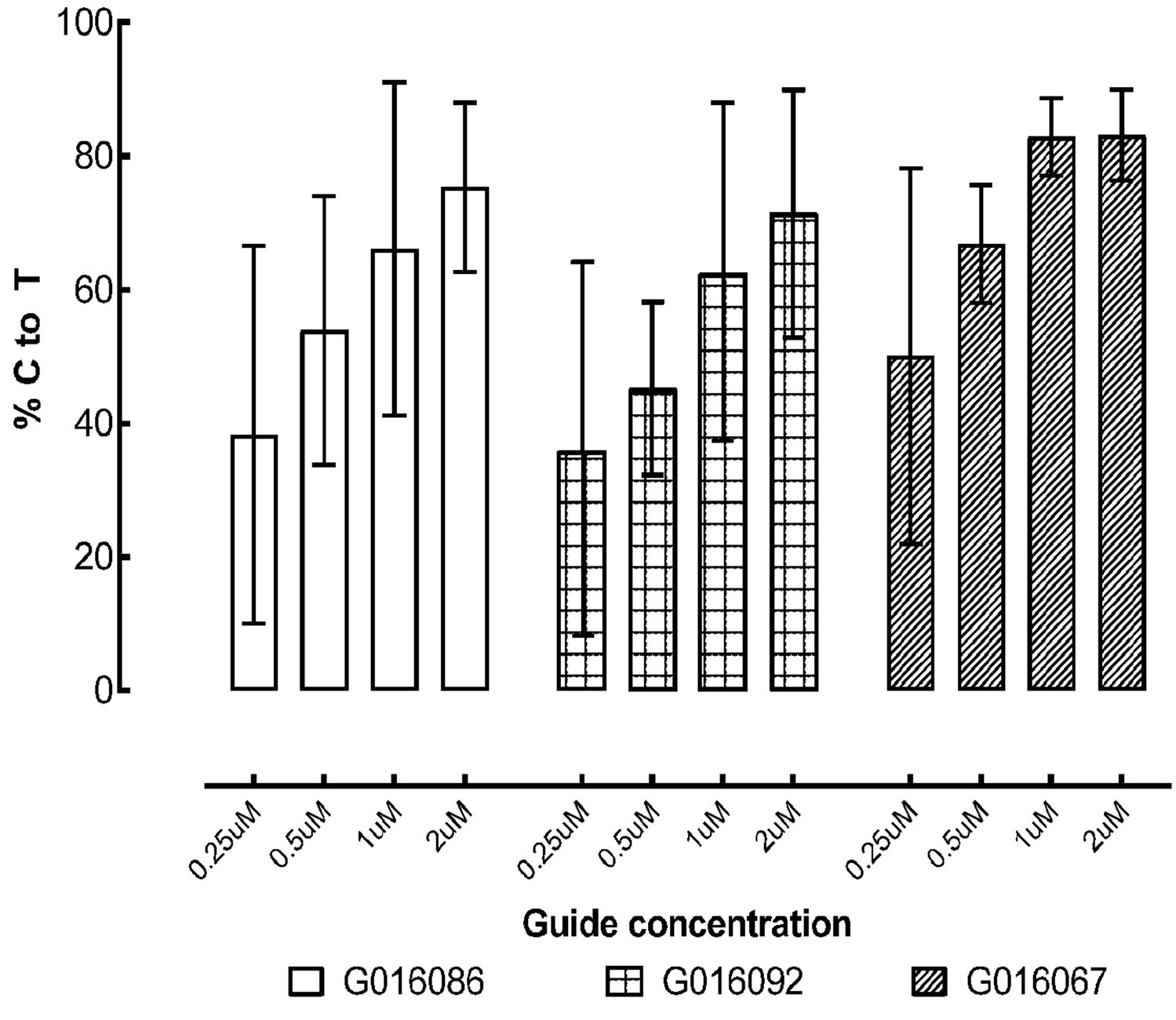


Fig. 24A

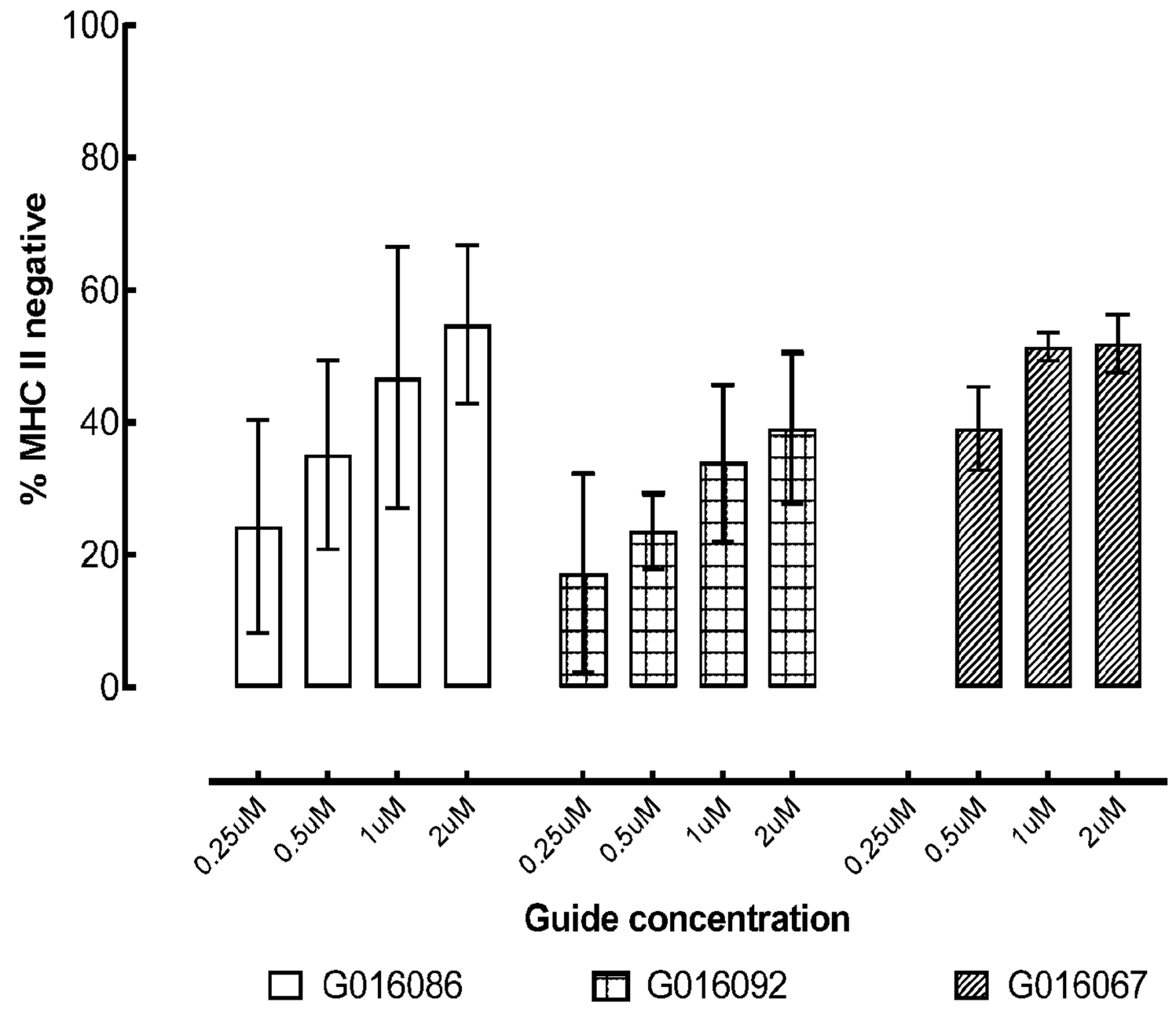


FIG. 24B

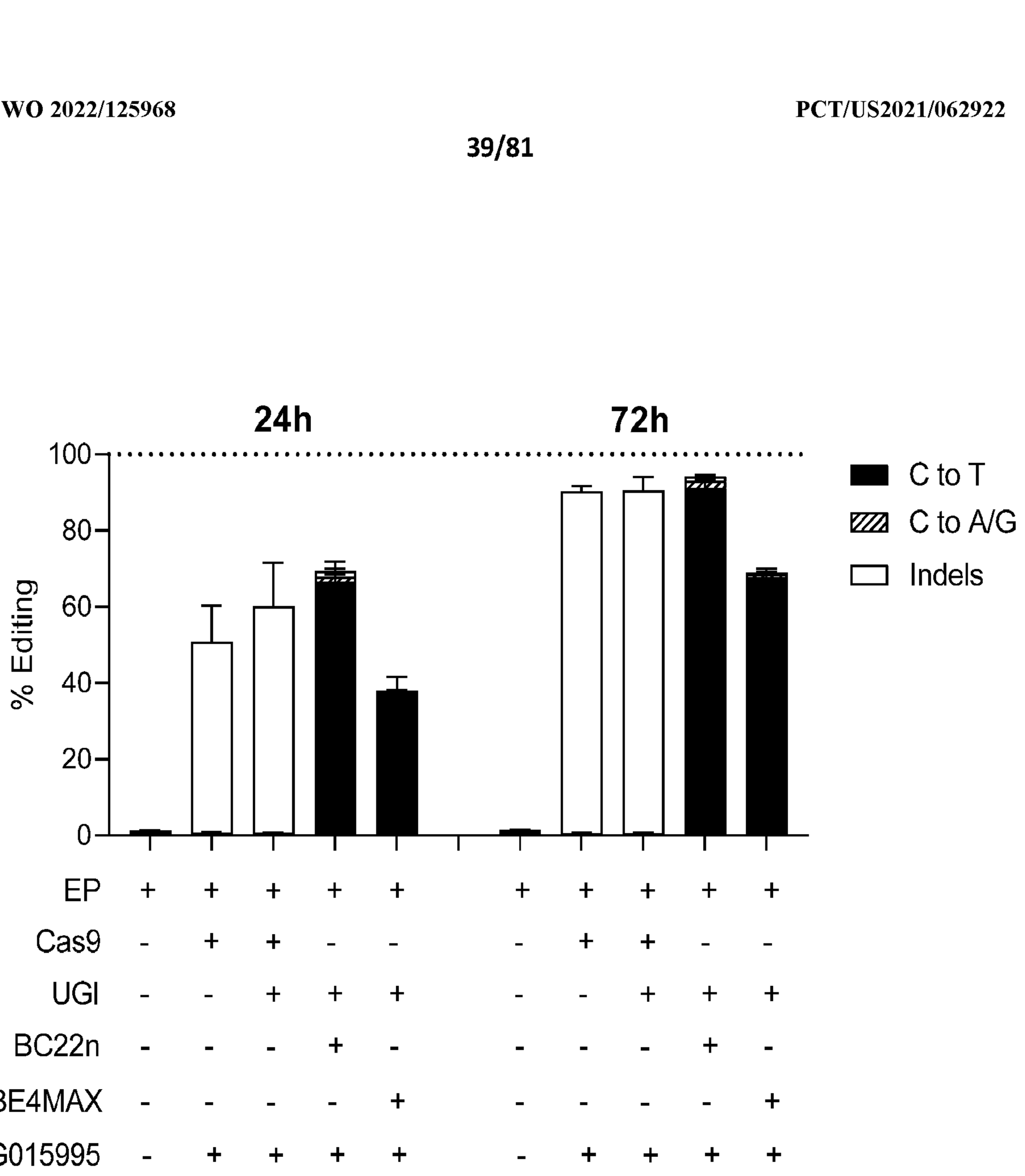

24h
72h
100
80
60
40
20
0
% Editing
C to T
C to A/G
Indels
EP + + + + + + + + + +
Cas9 - + + - - - + + - -
UGI - - + + + - - + + +
BC22n - - - + - - - - + -
BE4MAX - - - - + - - - - +
G015995 - + + + + - + + + +

Fig. 25

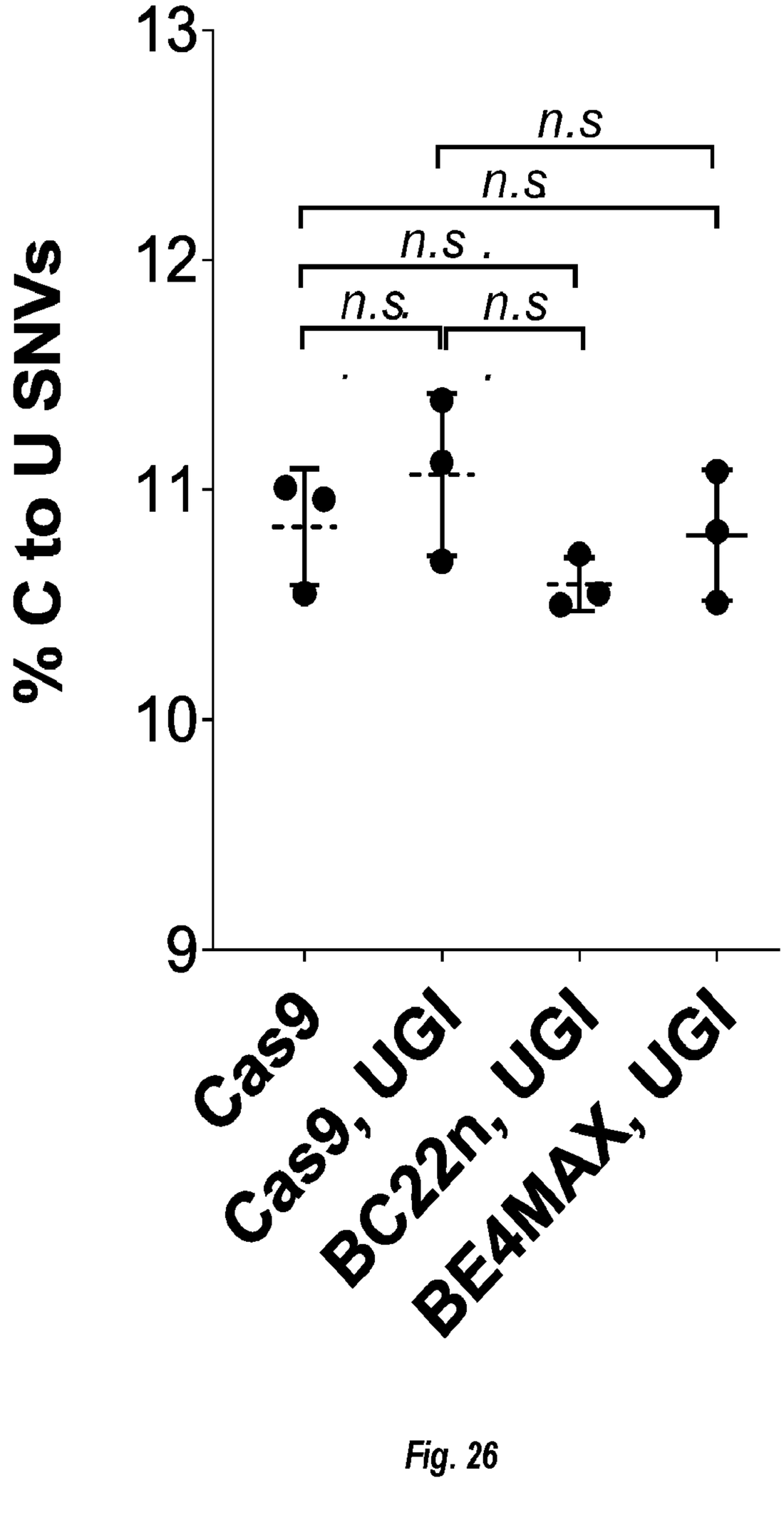

13
12
11
10
9
% C to U SNVs
n.s
n.s
n.s .
n.s.
n.s
Cas9
Cas9, UGI
BC22n, UGI
BE4MAX, UGI

Fig. 26

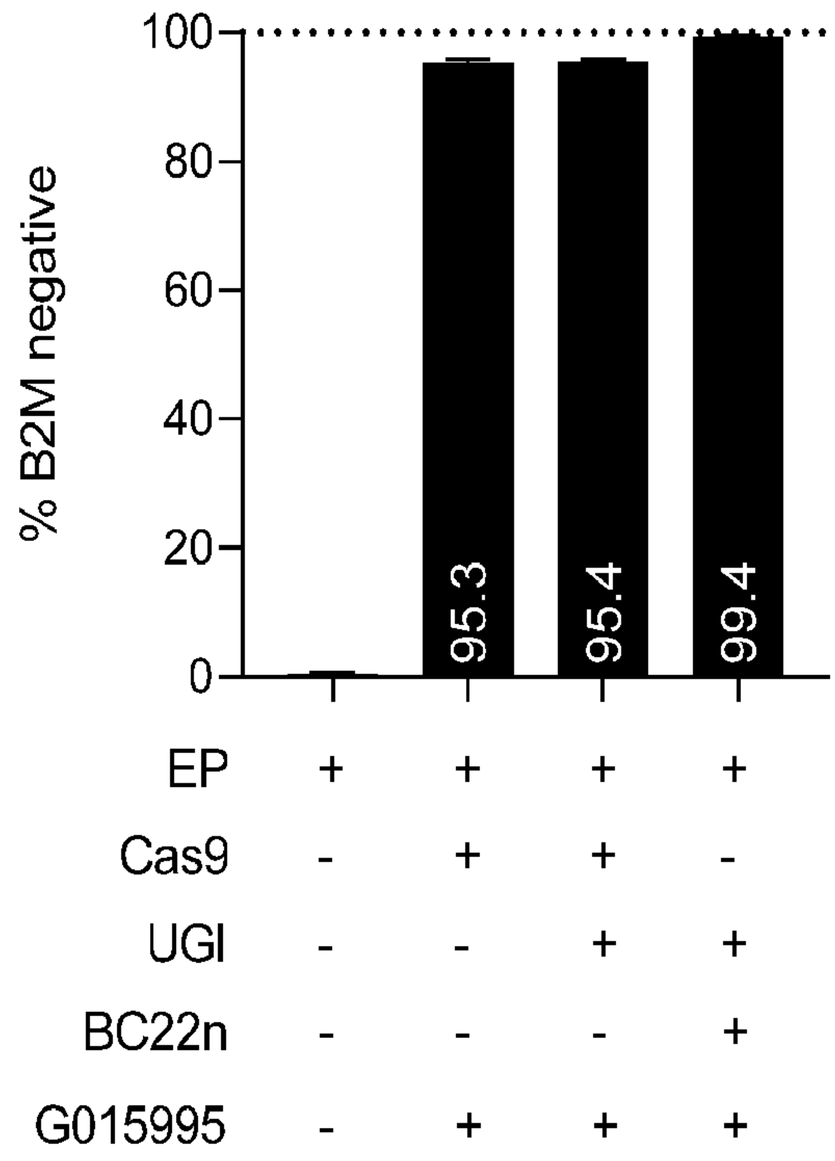

% B2M negative
100
80
60
40
20
0
95.3
95.4
99.4
EP + + + +
Cas9 - + + -
UGI - - + +
BC22n - - - +
G015995 - + + +

Fig. 27

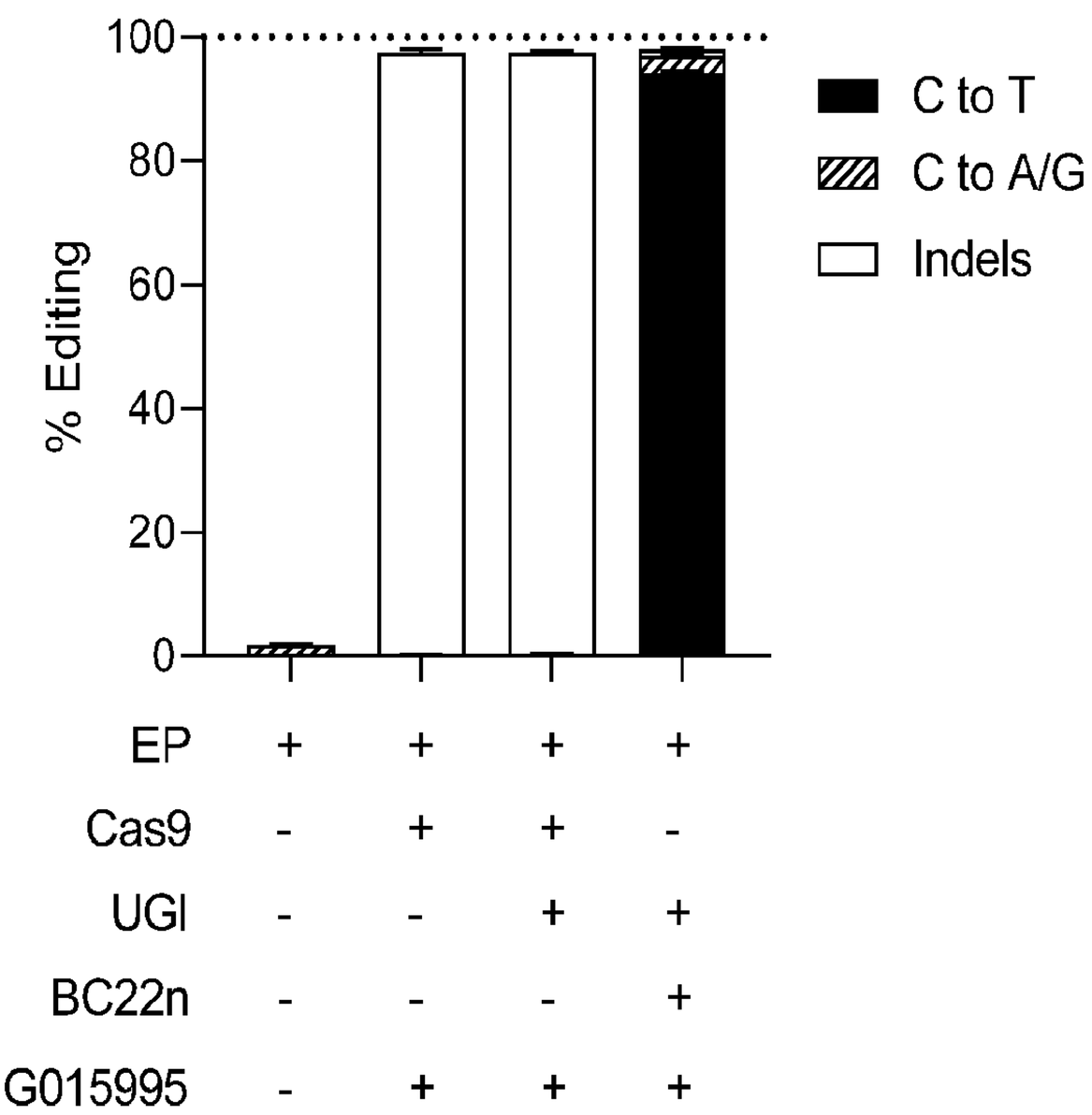

% Editing
100
80
60
40
20
0
C to T
C to A/G
Indels
EP + + + +
Cas9 - + + -
UGI - - + +
BC22n - - - +
G015995 - + + +

FIG. 28

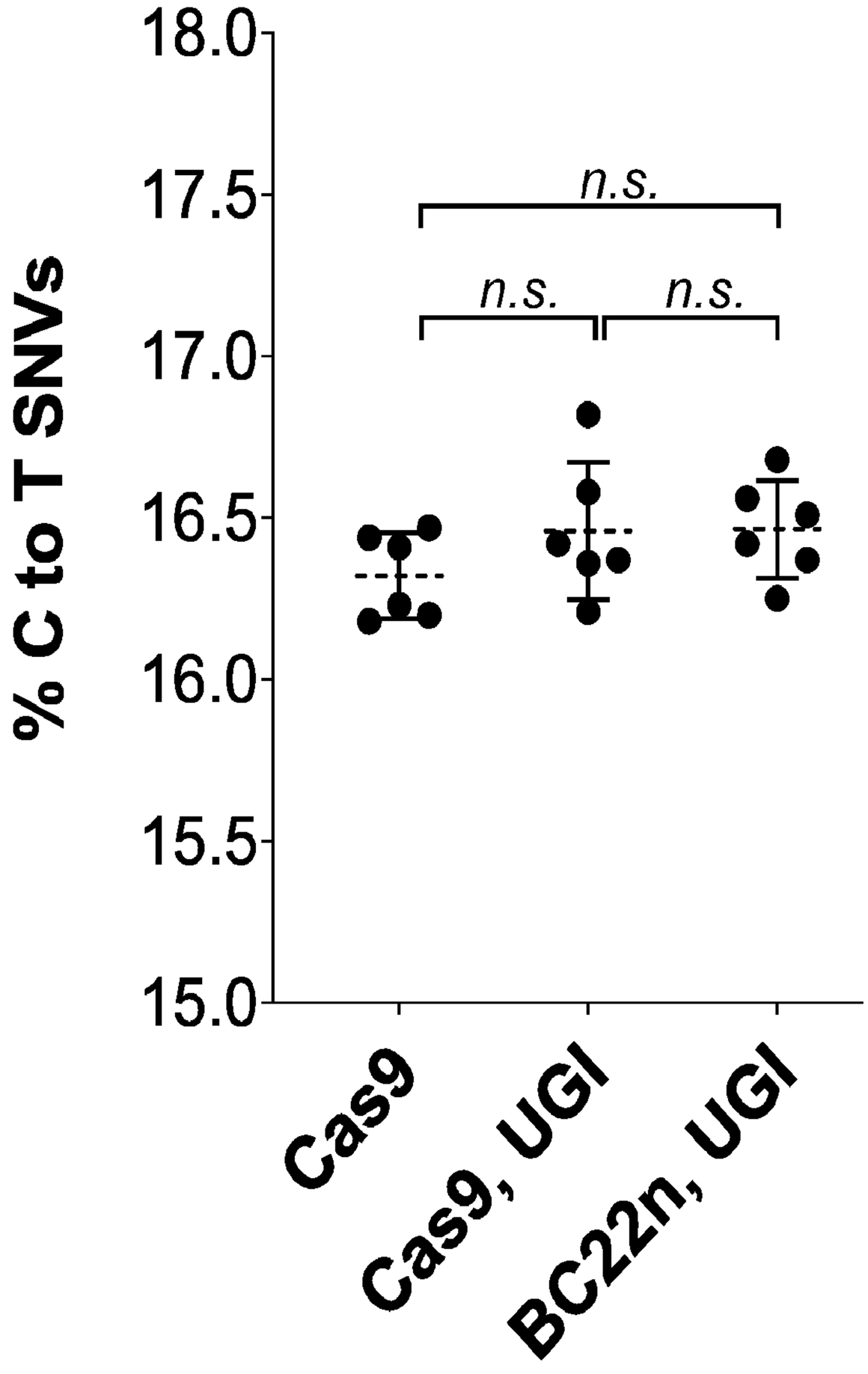

% C to T SNVs
18.0
17.5
17.0
16.5
16.0
15.5
15.0
n.s.
n.s.
n.s.
Cas9
Cas9, UGI
BC22n, UGI

Fig. 29

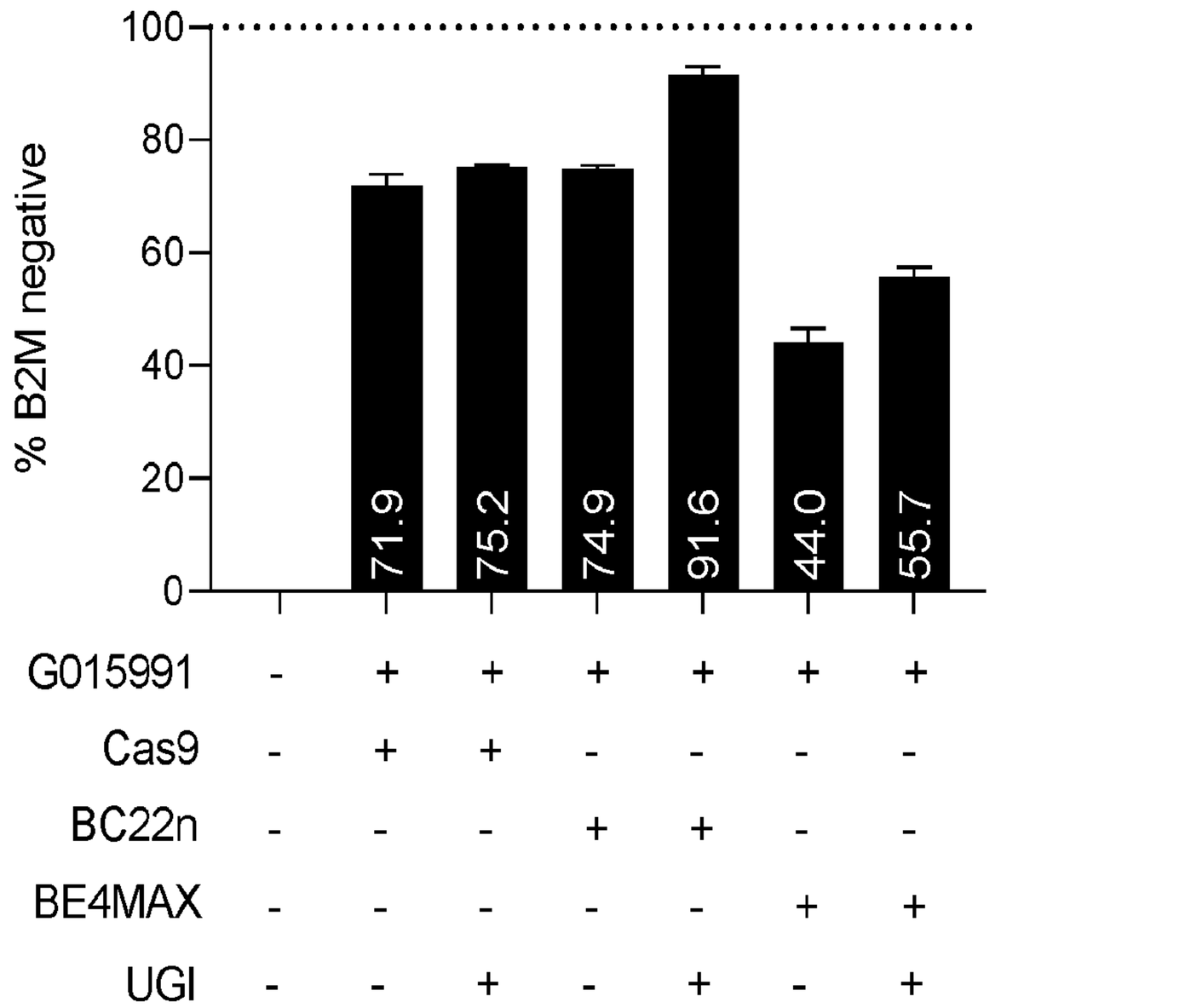

% B2M negative
100
80
60
40
20
0
71.9
75.2
74.9
91.6
44.0
55.7
G015991 - + + + + + +
Cas9 - + + - - - -
BC22n - - - + + - -
BE4MAX - - - - - + +
UGI - - + - + - +

Fig. 30

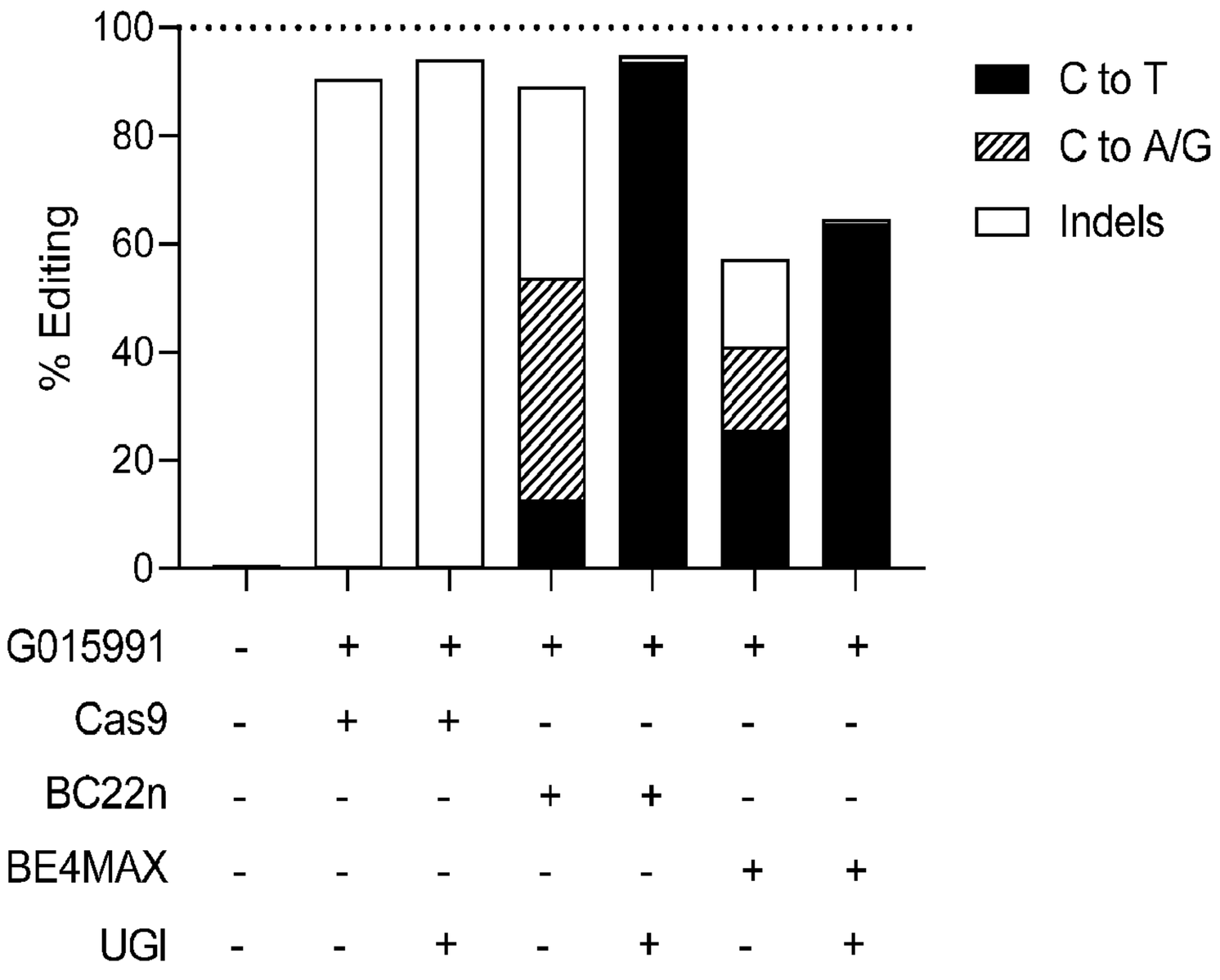

% Editing
100
80
60
40
20
0
C to T
C to A/G
Indels
G015991 - + + + + + +
Cas9 - + + - - - -
BC22n - - - + + - -
BE4MAX - - - - - + +
UGI - - + - + - +

FIG. 31

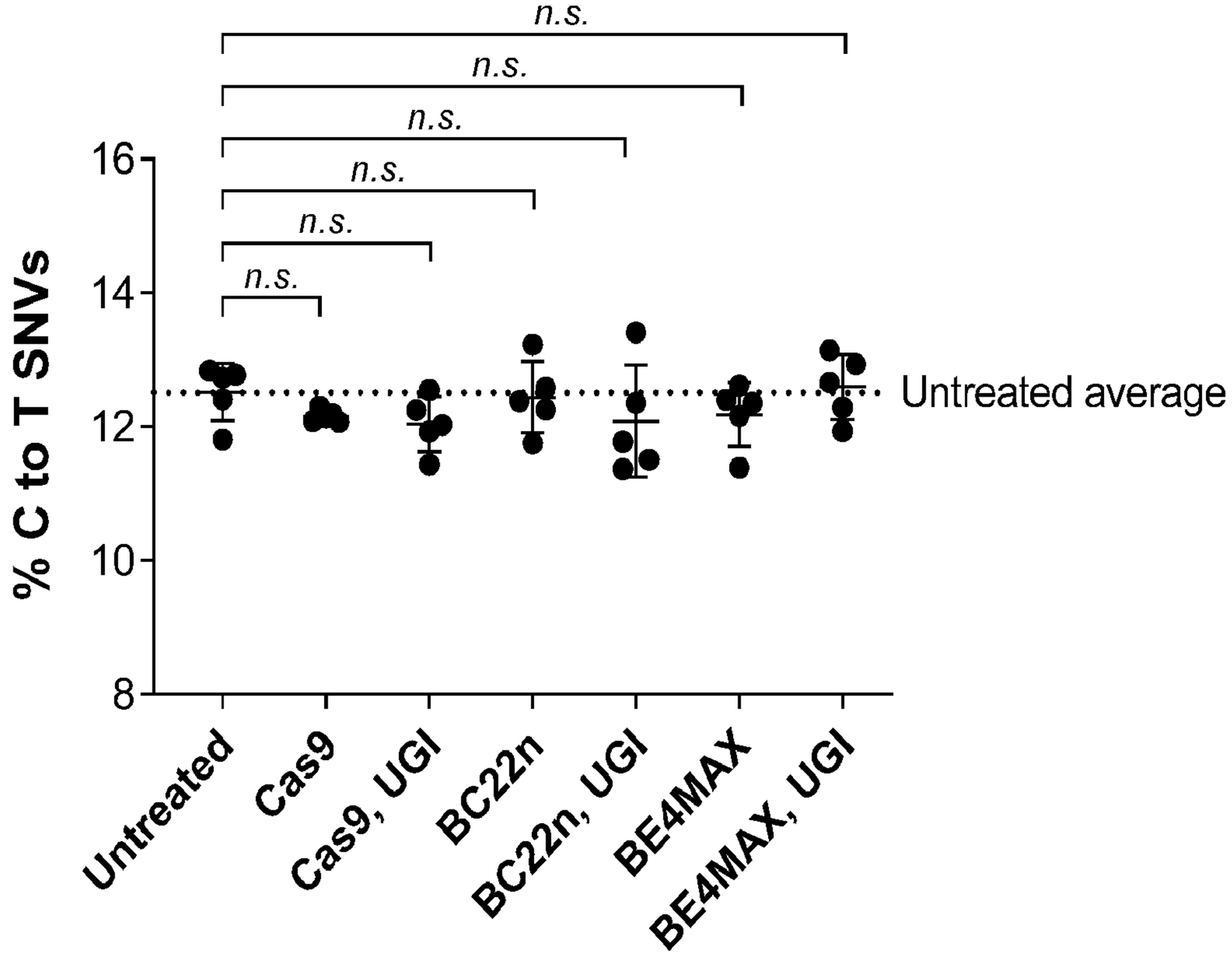

n.s.
n.s.
n.s.
n.s.
n.s.
16
14
12
10
8
% C to T SNVs
Untreated average
Untreated
Cas9
Cas9, UGI
BC22n
BC22n, UGI
BE4MAX
BE4MAX, UGI

Fig. 32

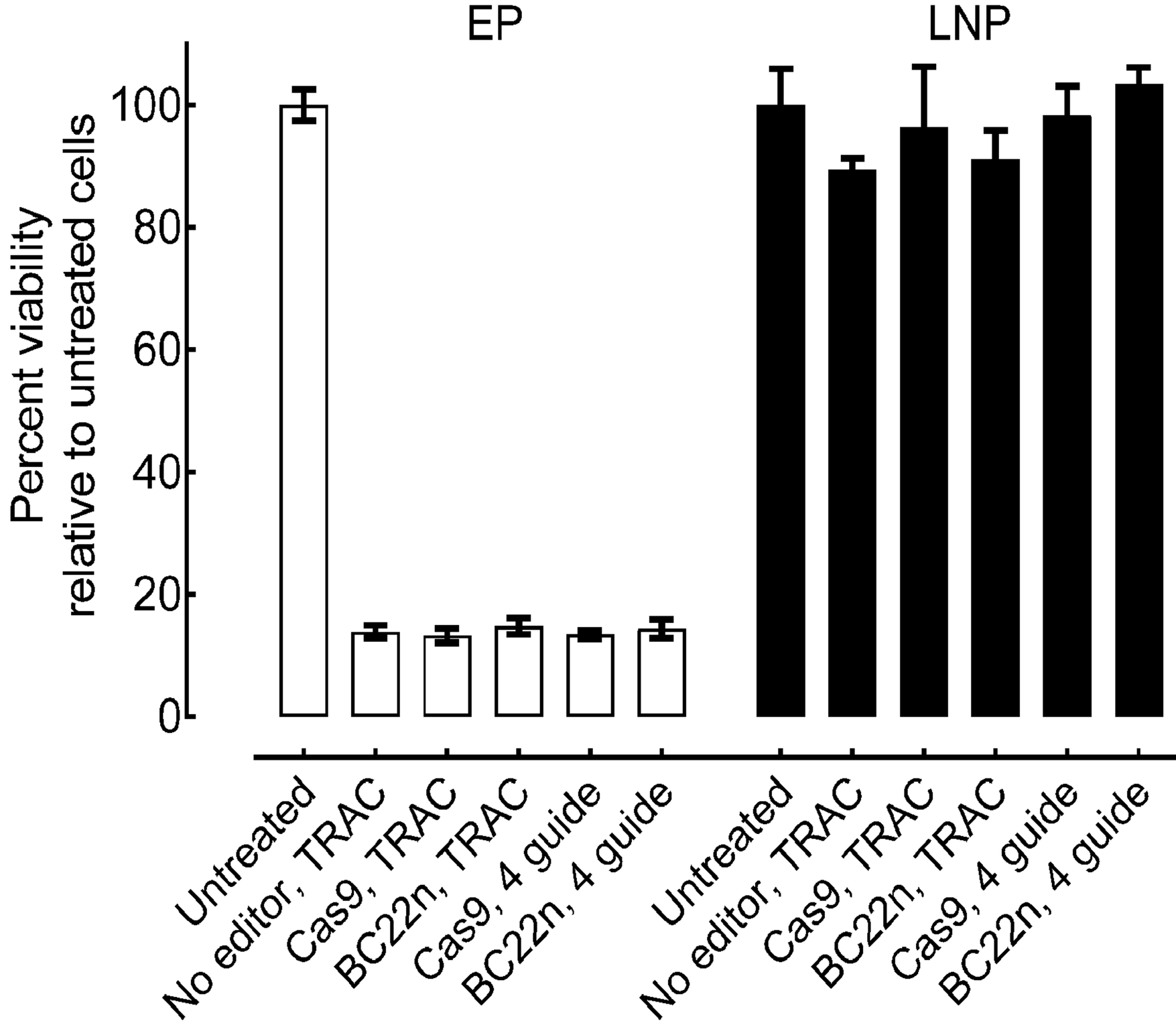

EP
LNP
Percent viability
relative to untreated cells
100
80
60
40
20
0
Untreated
No editor, TRAC
Cas9, TRAC
BC22n, TRAC
Cas9, 4 guide
BC22n, 4 guide
Untreated
No editor, TRAC
Cas9, TRAC
BC22n, TRAC
Cas9, 4 guide
BC22n, 4 guide

Fig. 33

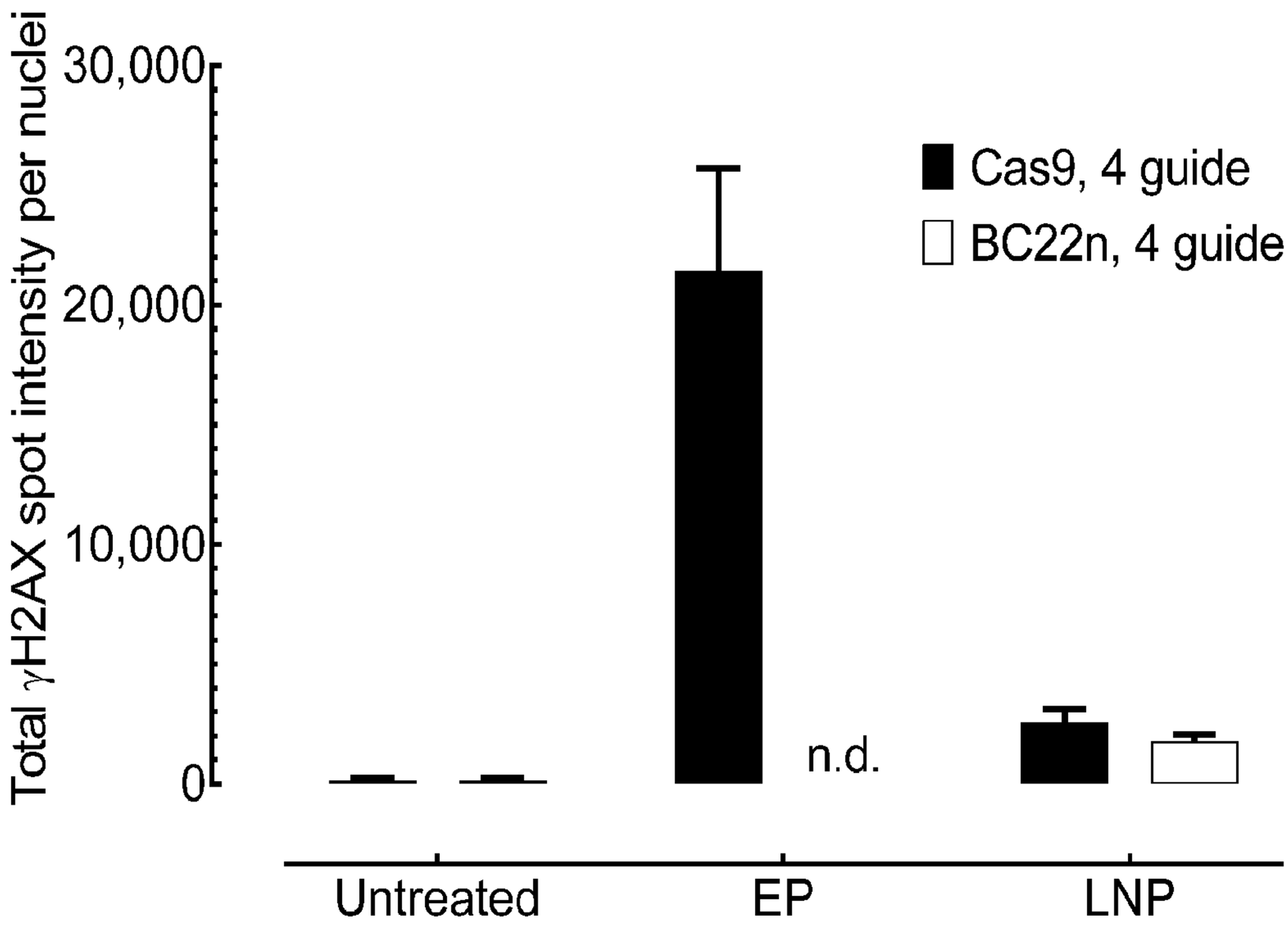

Total γH2AX spot intensity per nuclei
30,000
20,000
10,000
0
Cas9, 4 guide
BC22n, 4 guide
n.d.
Untreated
EP
LNP

Fig. 34

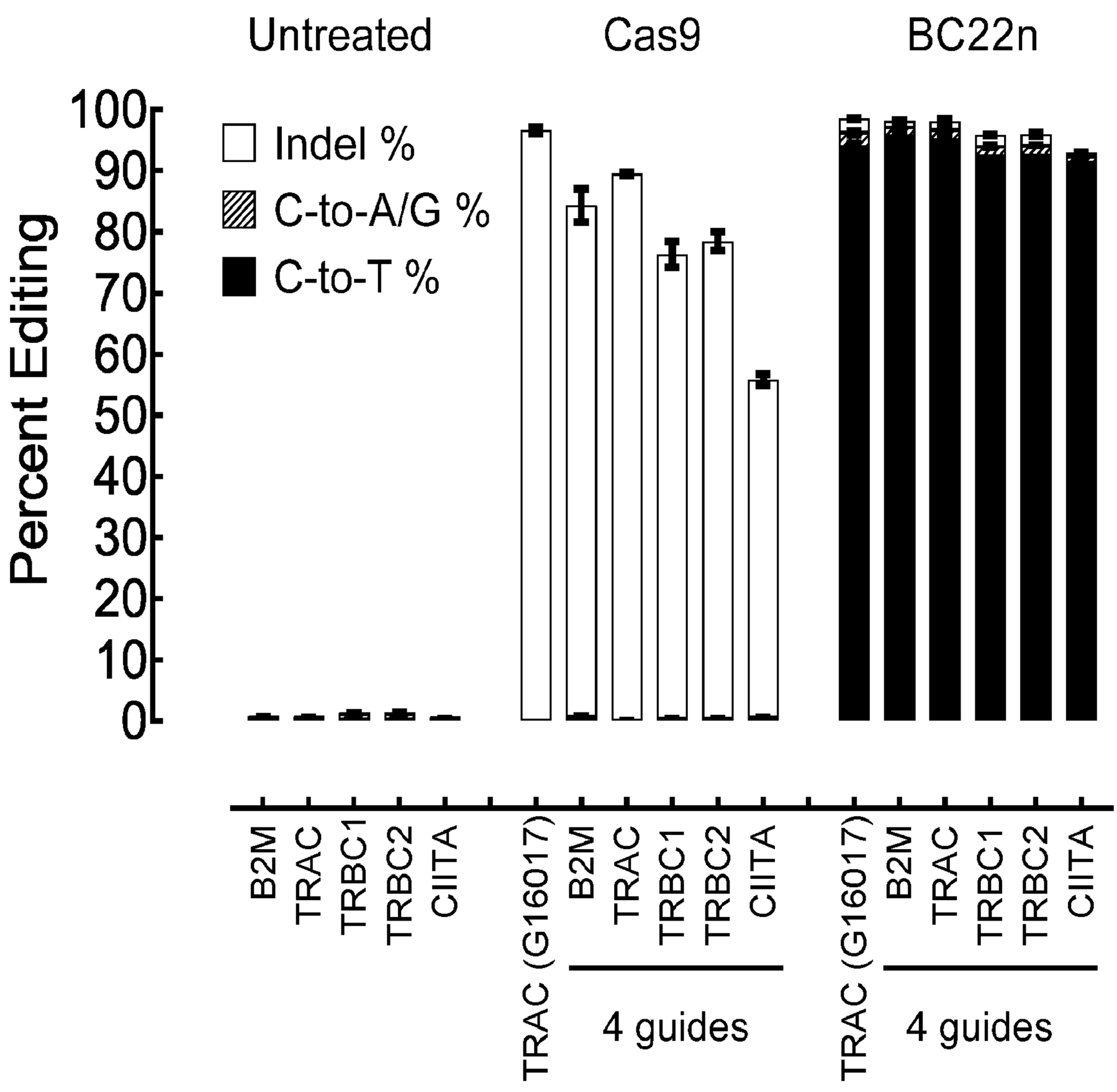

Untreated
Cas9
BC22n
Indel %
C-to-A/G %
C-to-T %
Percent Editing
100
90
80
70
60
50
40
30
20
10
0
B2M
TRAC
TRBC1
TRBC2
CIITA
TRAC (G16017)
B2M
TRAC
TRBC1
TRBC2
CIITA
4 guides
TRAC (G16017)
B2M
TRAC
TRBC1
TRBC2
CIITA
4 guides

FIG. 35

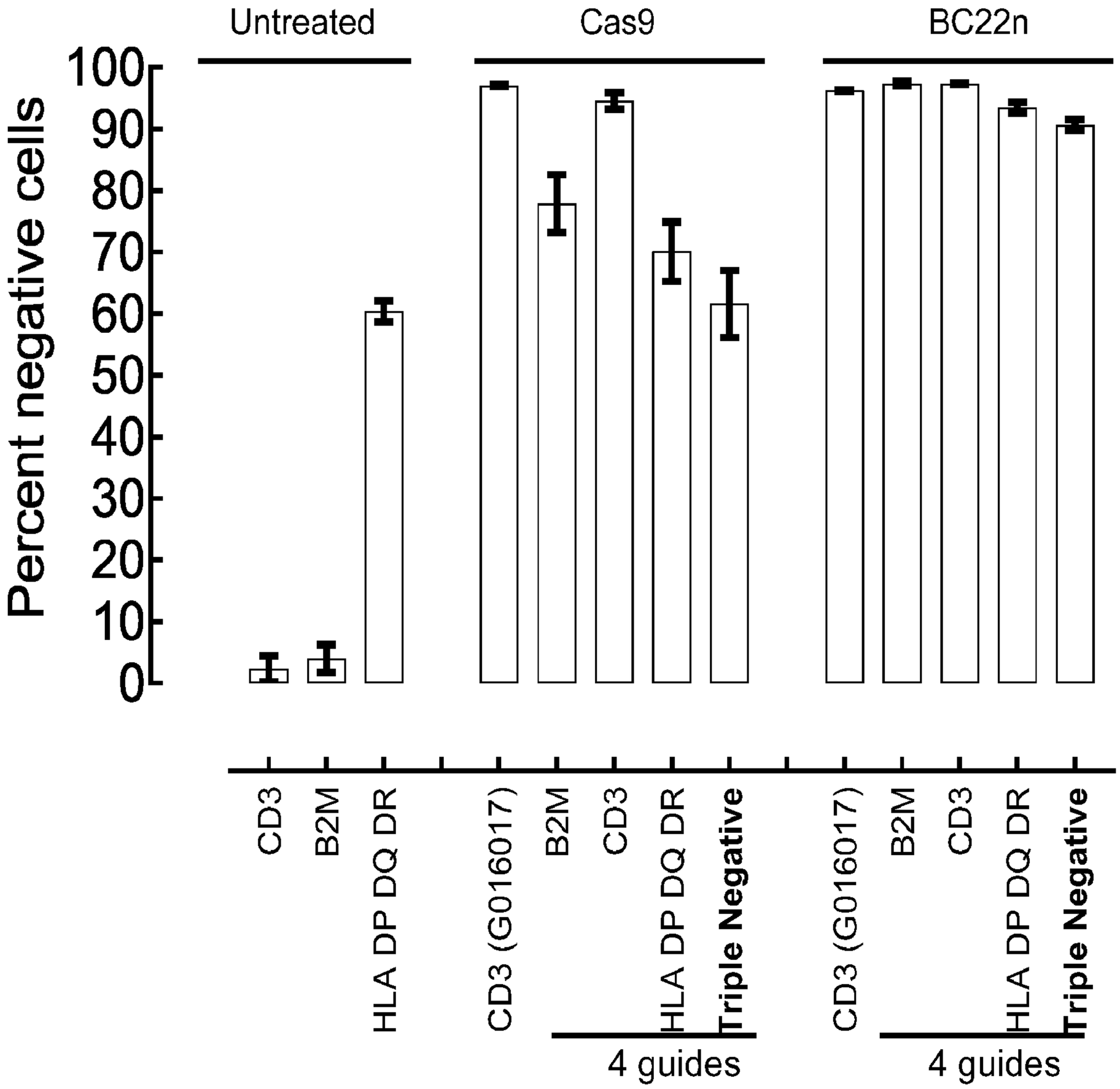

Untreated
Cas9
BC22n
Percent negative cells
100
90
80
70
60
50
40
30
20
10
0
CD3
B2M
HLA DP DQ DR
CD3 (G016017)
B2M
CD3
HLA DP DQ DR
Triple Negative
4 guides
CD3 (G016017)
B2M
CD3
HLA DP DQ DR
Triple Negative
4 guides

Fig. 36

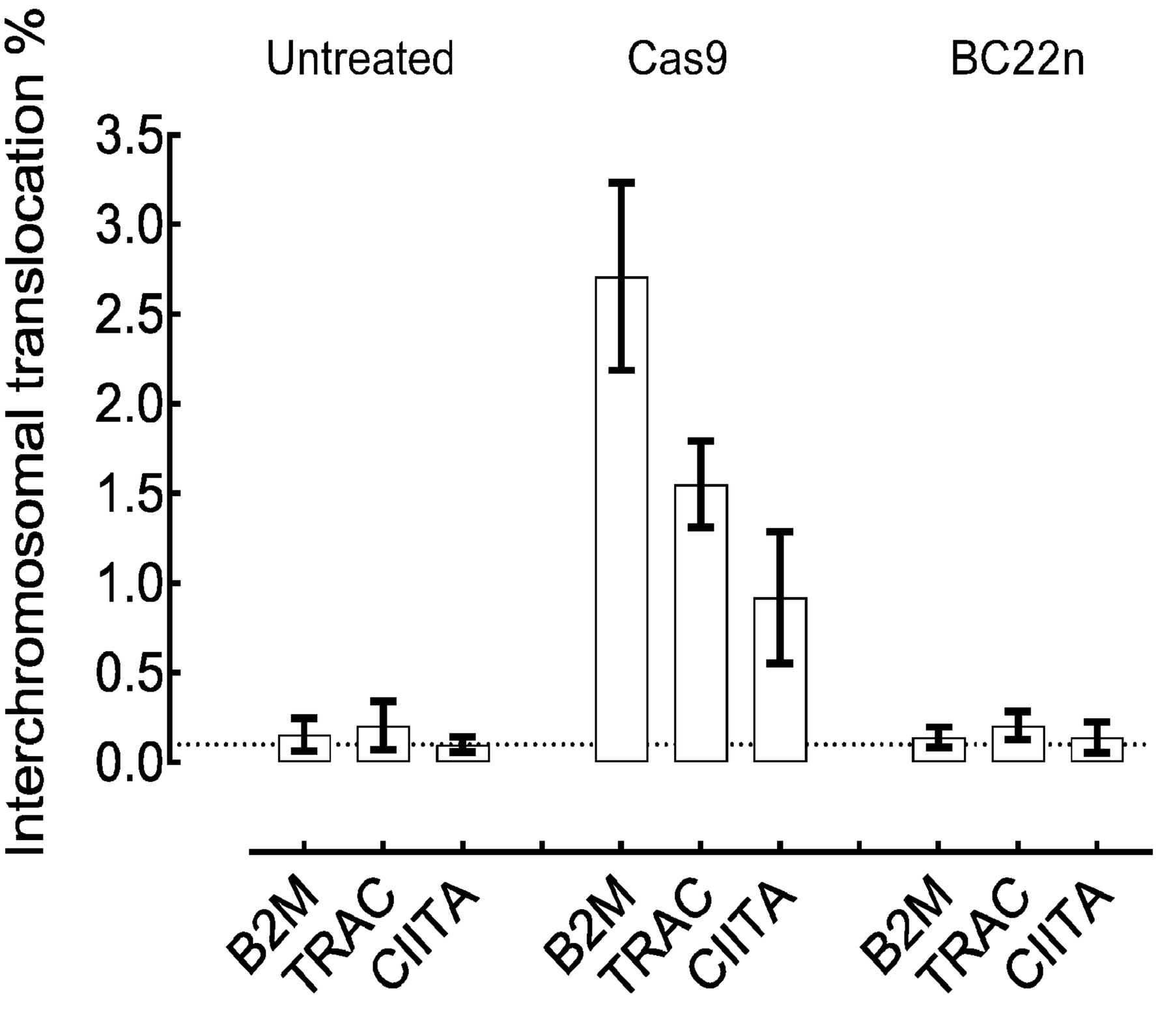

Untreated
Cas9
BC22n
Interchromosomal translocation %
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
B2M
TRAC
CIITA
B2M
TRAC
CIITA
B2M
TRAC
CIITA

Fig. 37

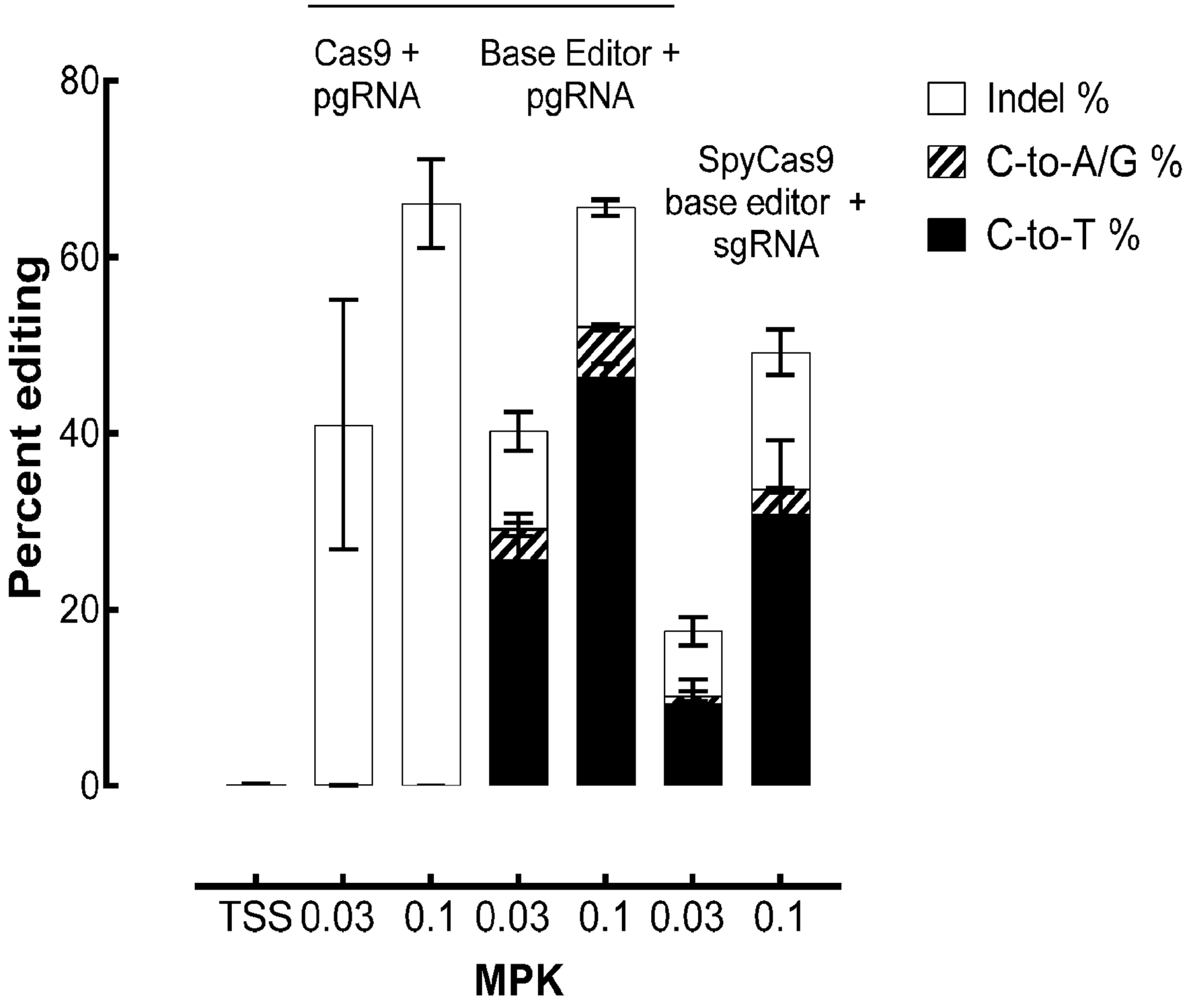

Nme2Cas9
Cas9 + pgRNA
Base Editor + pgRNA
SpyCas9 base editor + sgRNA
Indel %
C-to-A/G %
C-to-T %
Percent editing
80
60
40
20
0
TSS 0.03 0.1 0.03 0.1 0.03 0.1
MPK

FIG. 38

sg005883

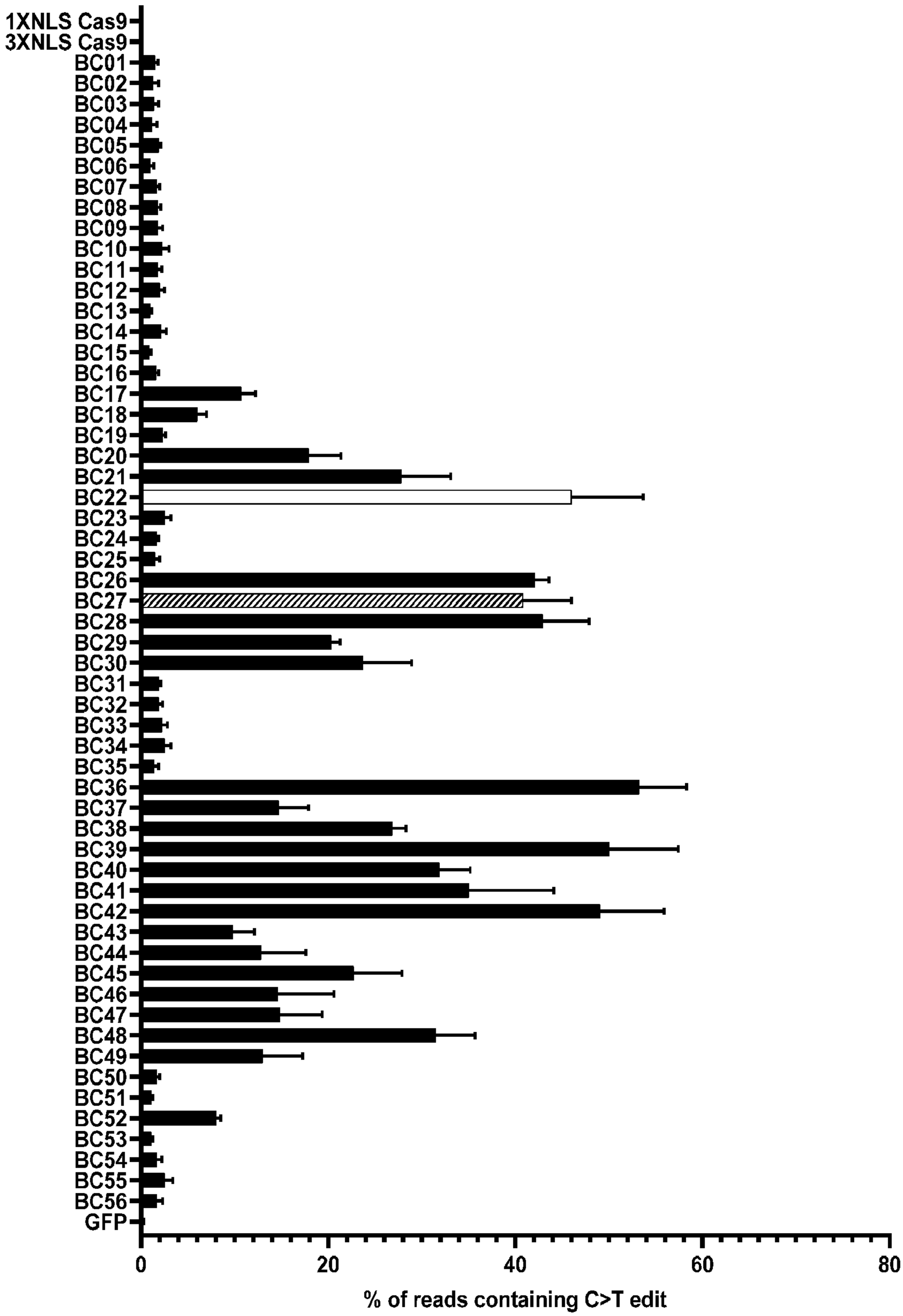


Fig. 39

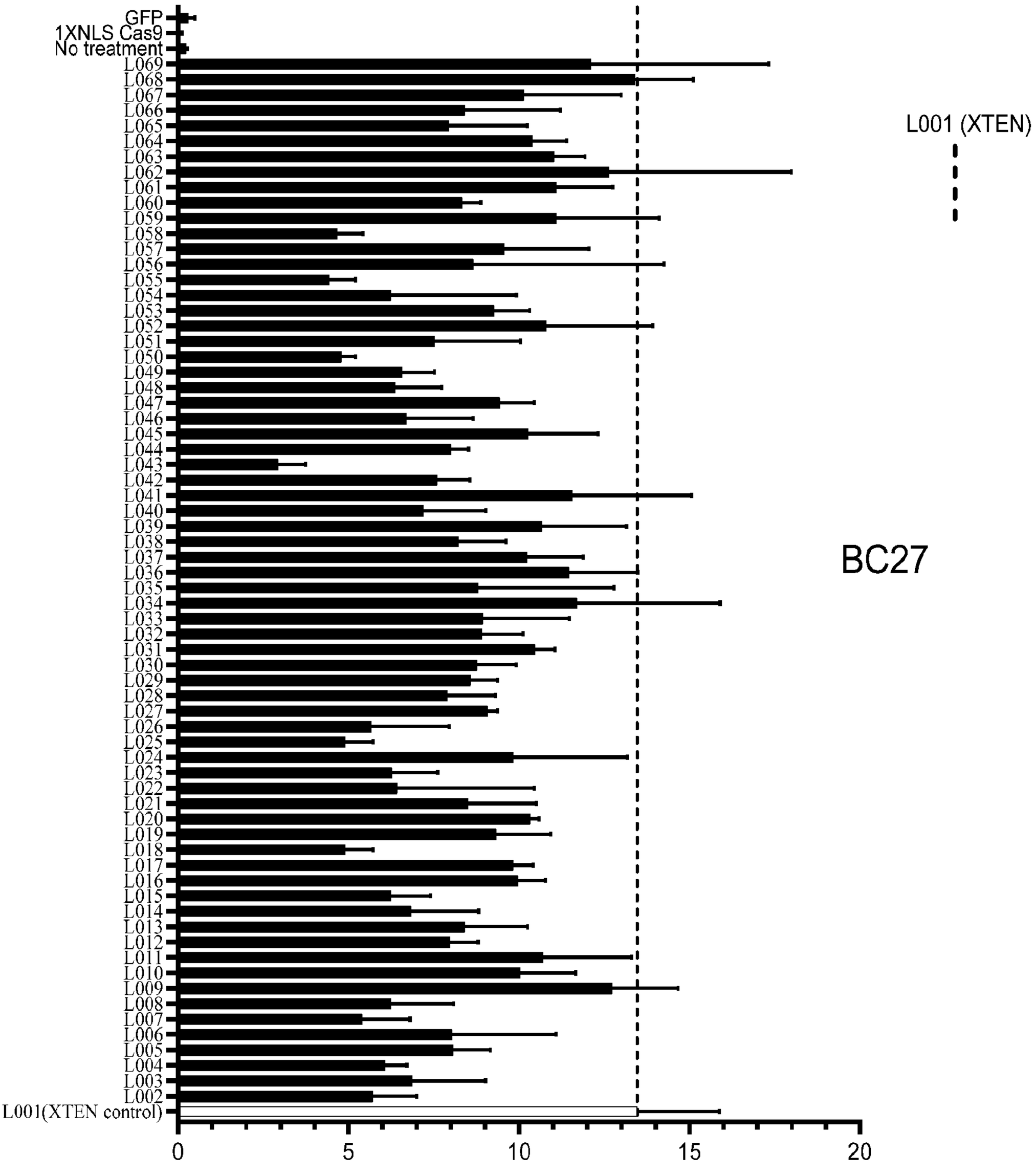

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC27
0
5
10
15
20
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40A

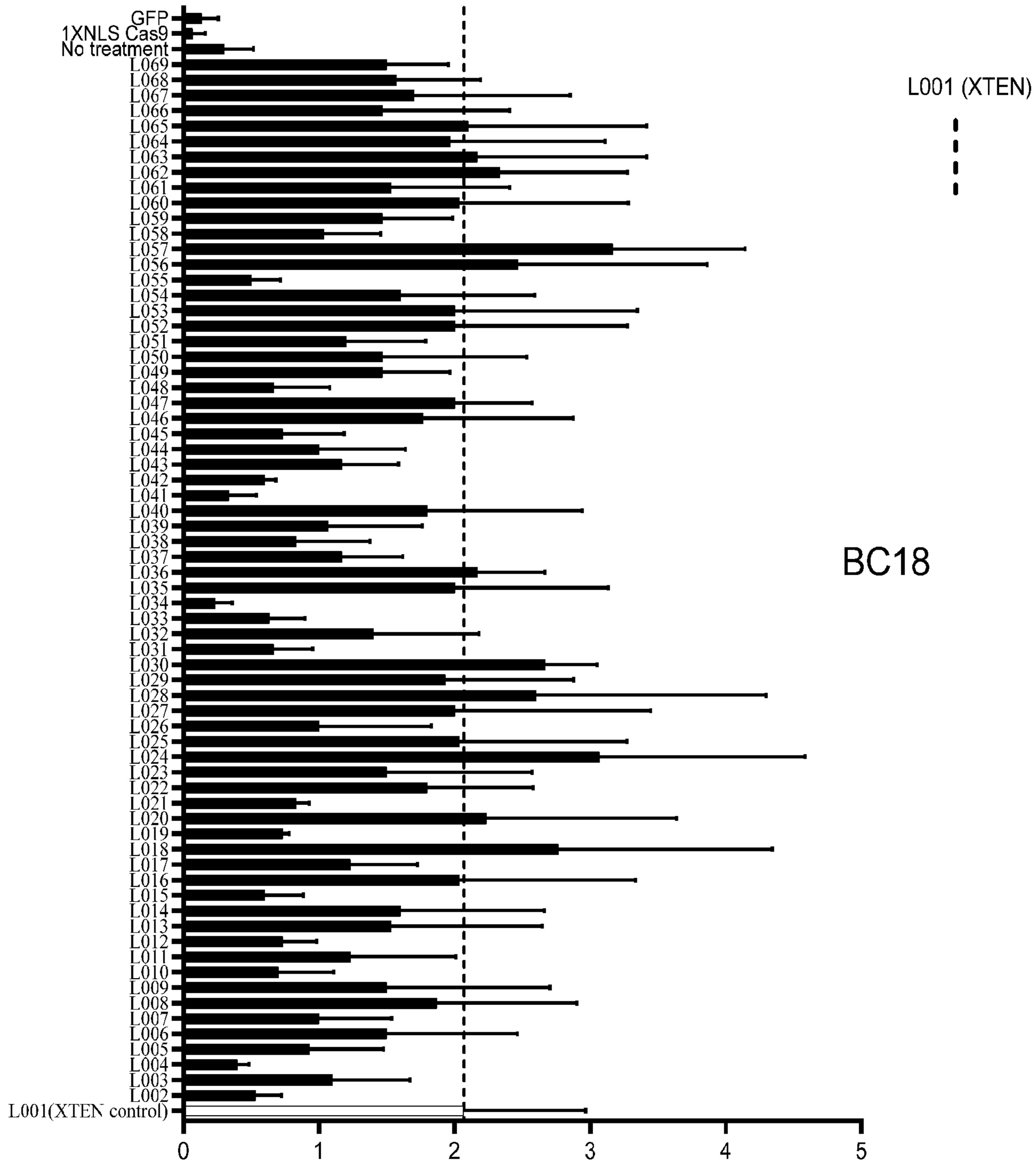

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC18
0
1
2
3
4
5
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40B

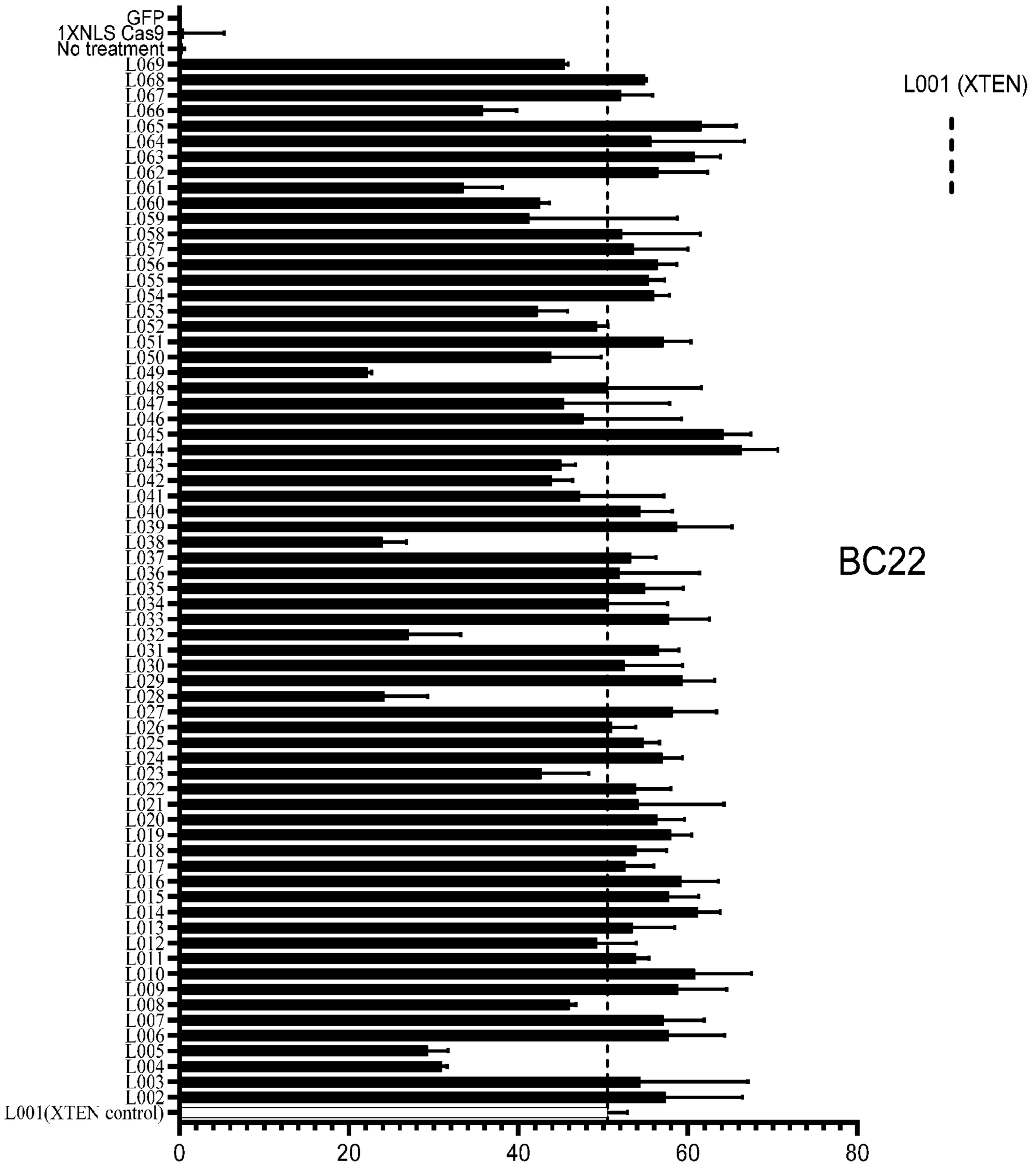

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC22
0
20
40
60
80
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40C

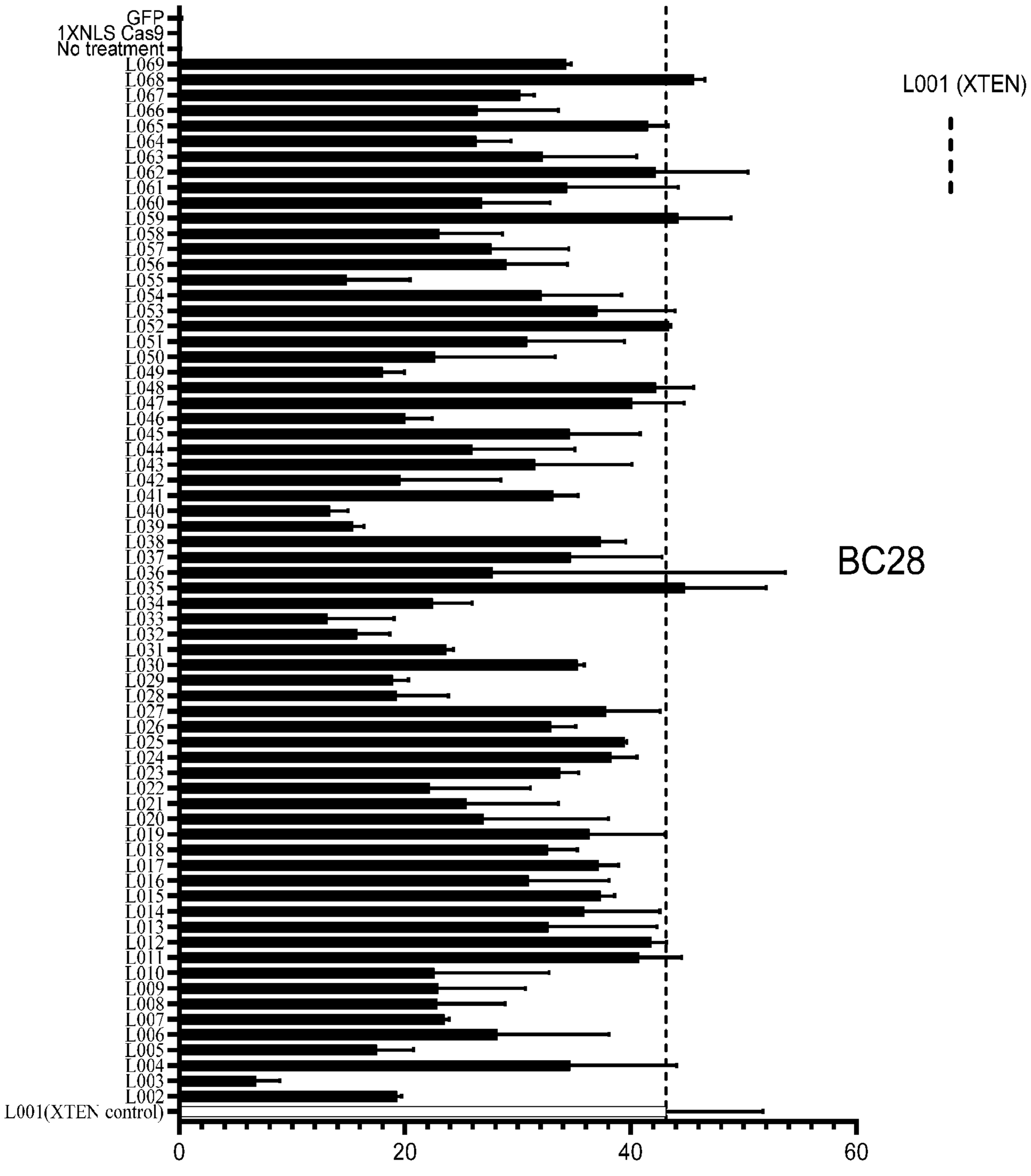

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC28
0
20
40
60
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40D

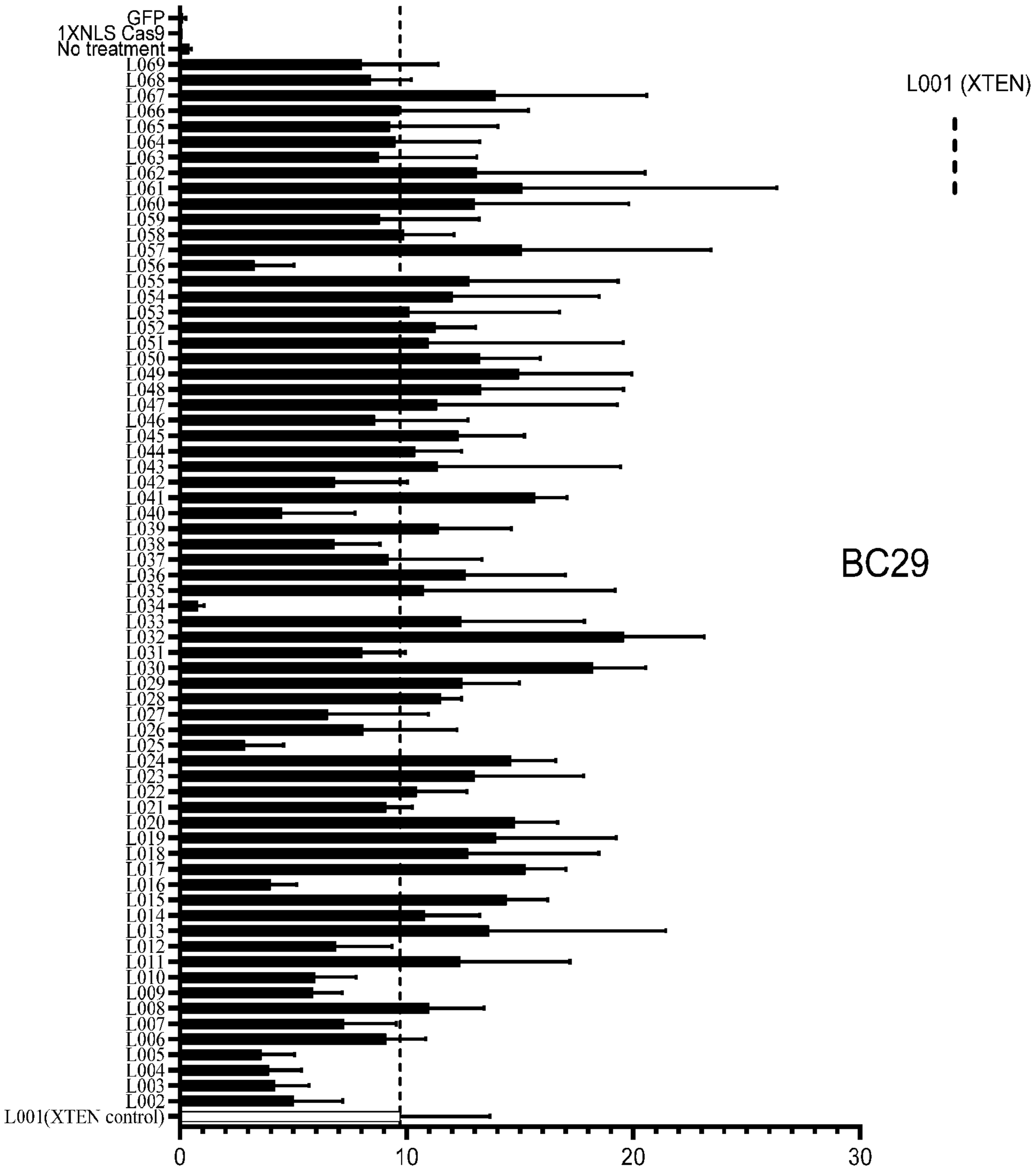

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC29
0
10
20
30
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40E

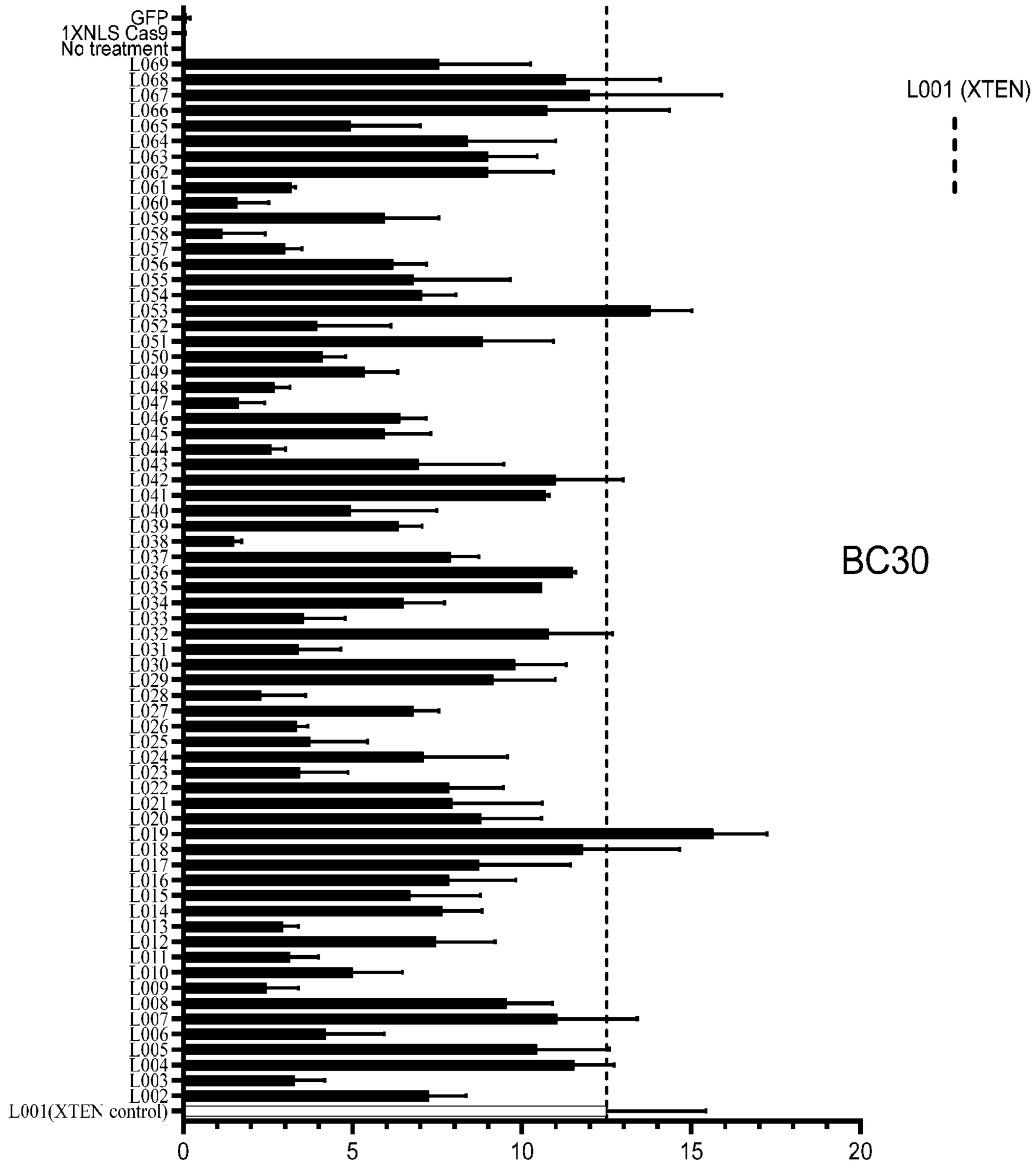

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC30
0
5
10
15
20
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40F

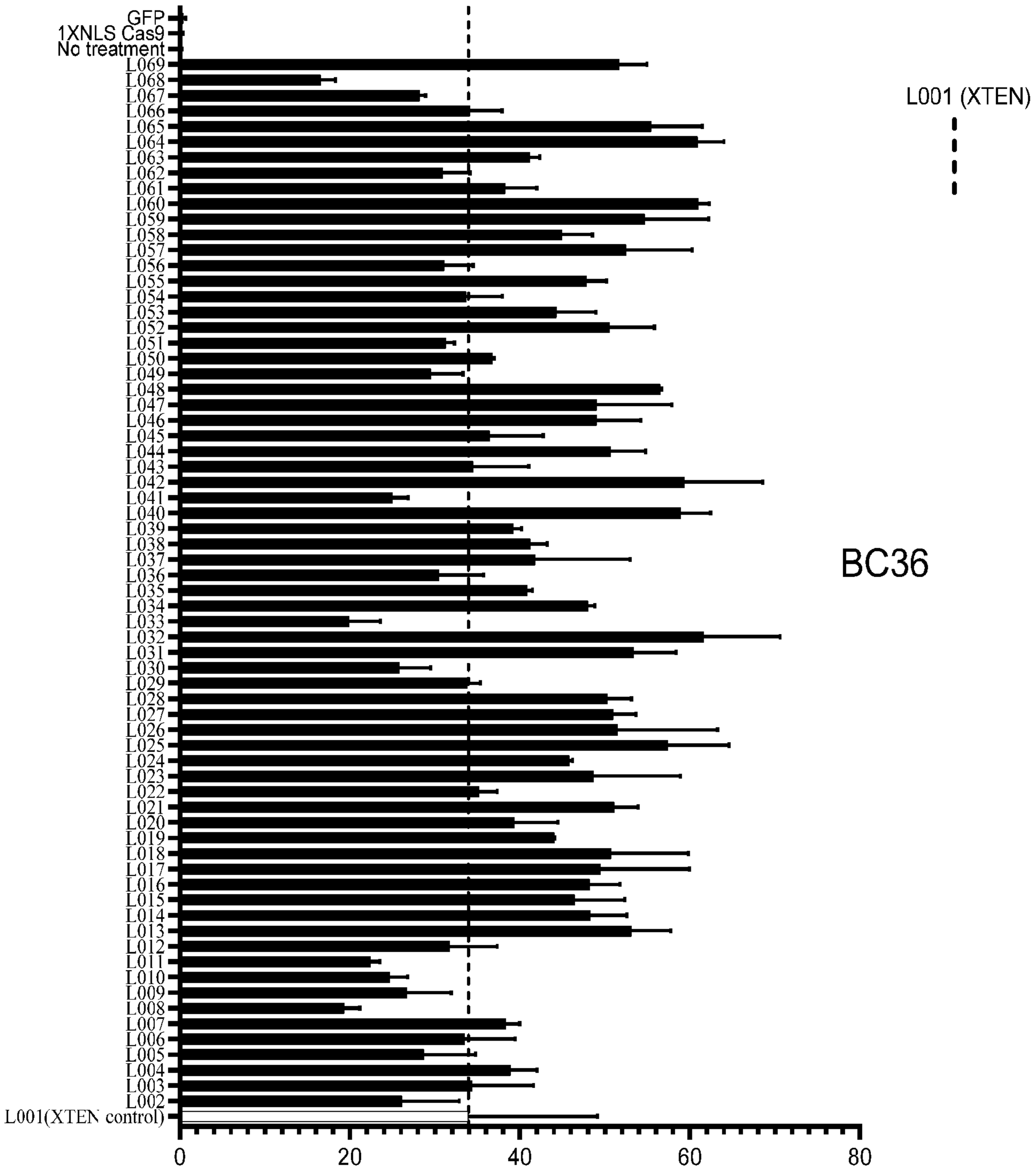

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC36
0
20
40
60
80
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40G

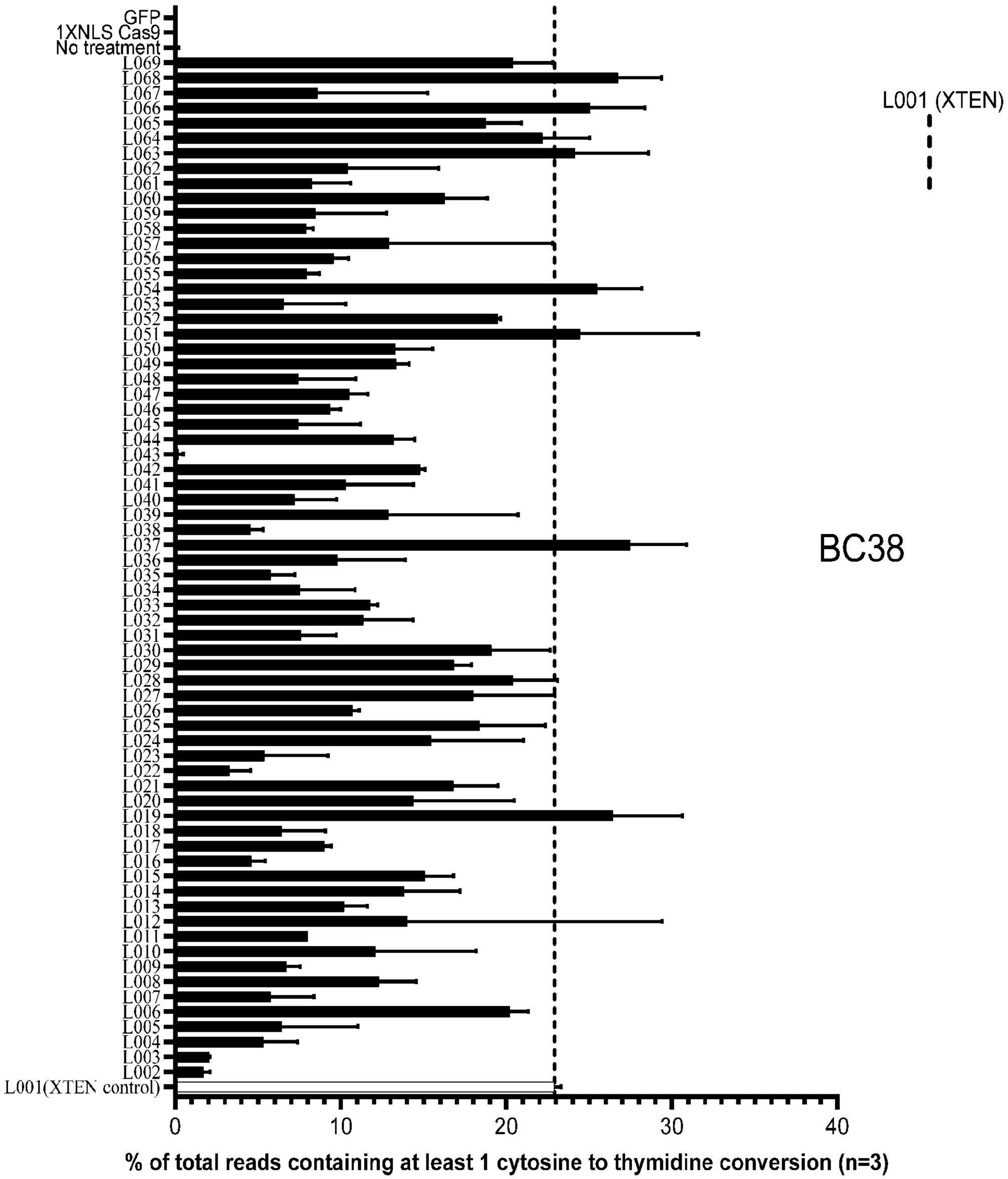

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC38
0
10
20
30
40
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40H

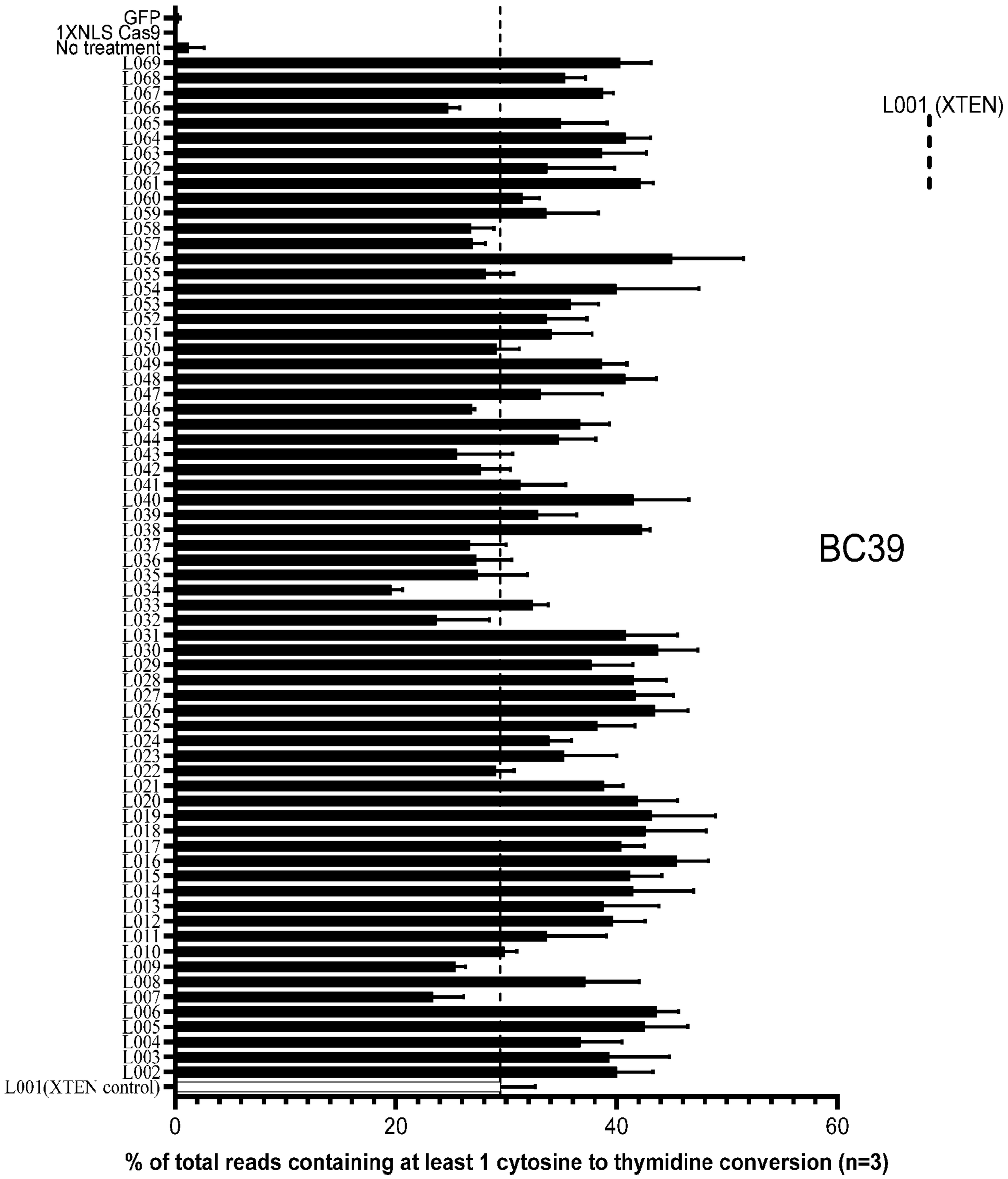

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001 (XTEN)
BC39
0
20
40
60
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40I

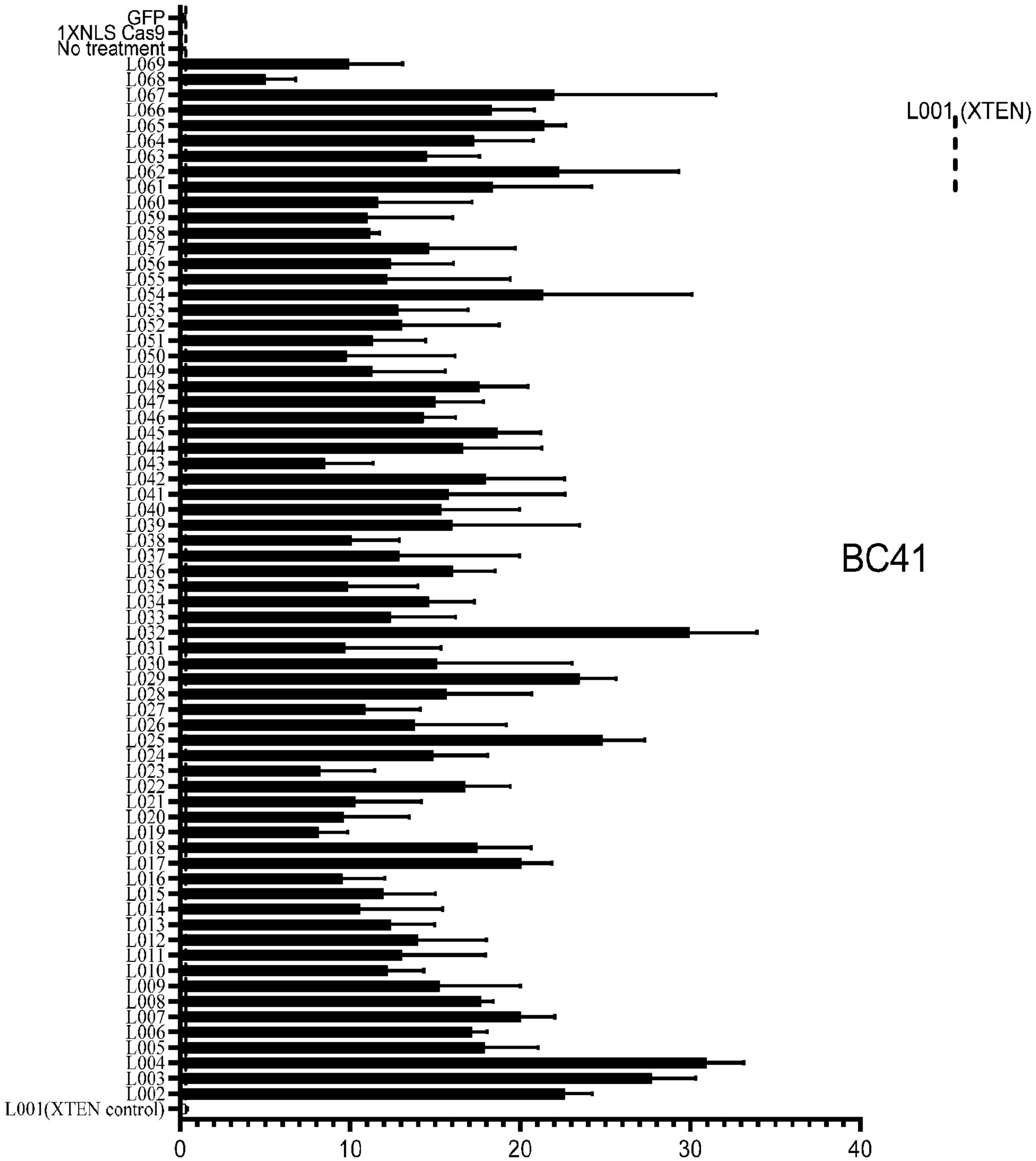

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001(XTEN)
BC41
0
10
20
30
40
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40J

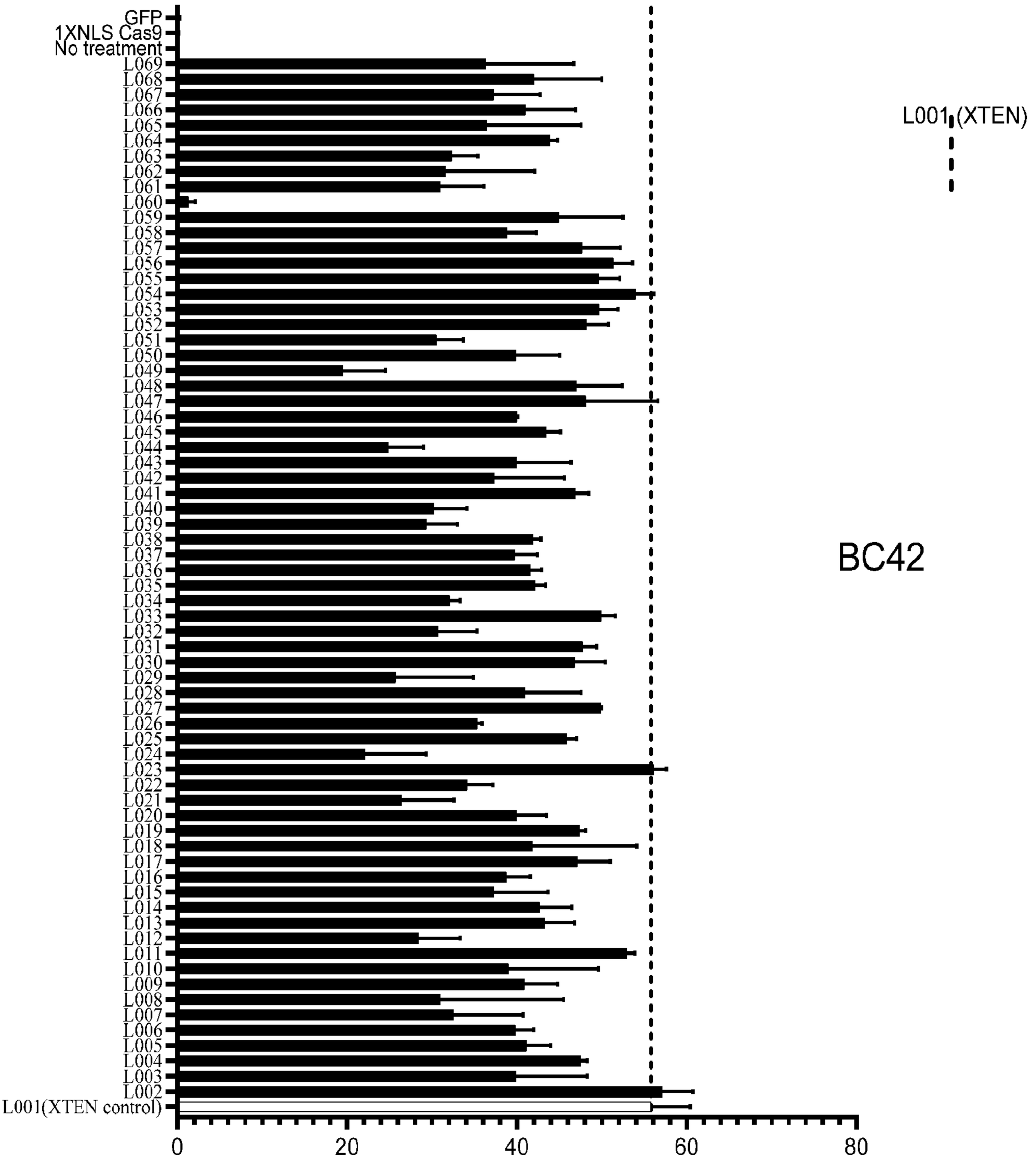

GFP
1XNLS Cas9
No treatment
L069
L068
L067
L066
L065
L064
L063
L062
L061
L060
L059
L058
L057
L056
L055
L054
L053
L052
L051
L050
L049
L048
L047
L046
L045
L044
L043
L042
L041
L040
L039
L038
L037
L036
L035
L034
L033
L032
L031
L030
L029
L028
L027
L026
L025
L024
L023
L022
L021
L020
L019
L018
L017
L016
L015
L014
L013
L012
L011
L010
L009
L008
L007
L006
L005
L004
L003
L002
L001(XTEN control)
L001(XTEN)
BC42
0
20
40
60
80
% of total reads containing at least 1 cytosine to thymidine conversion (n=3)

Fig. 40K

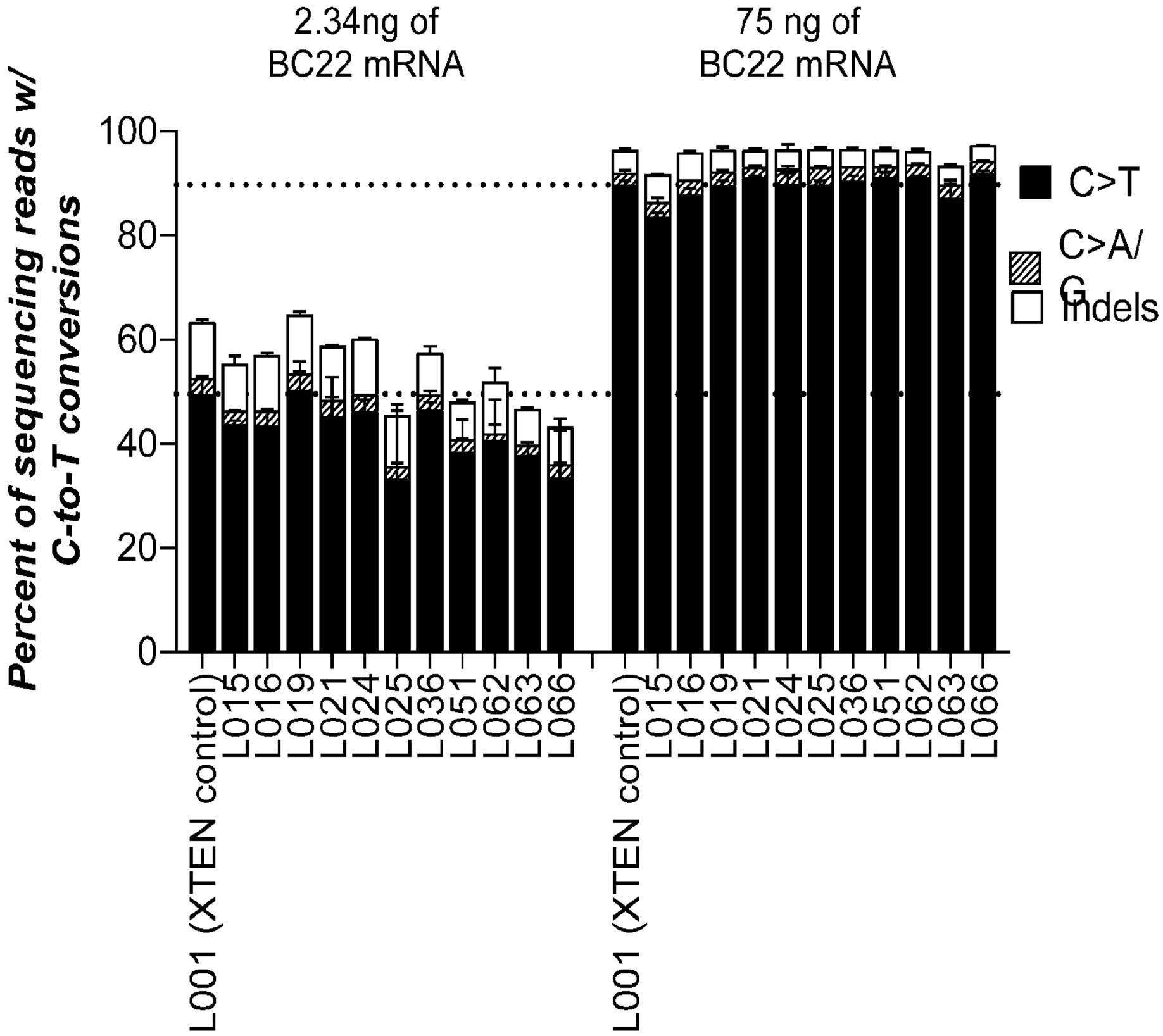

2.34ng of BC22 mRNA
75 ng of BC22 mRNA
Percent of sequencing reads C-to-T conversions
100
80
60
40
20
0
C>T
C>A/G
Indels
L001 (XTEN control)
L015
L016
L019
L021
L024
L025
L036
L051
L062
L063
L066
BC22 mRNA

Fig. 41

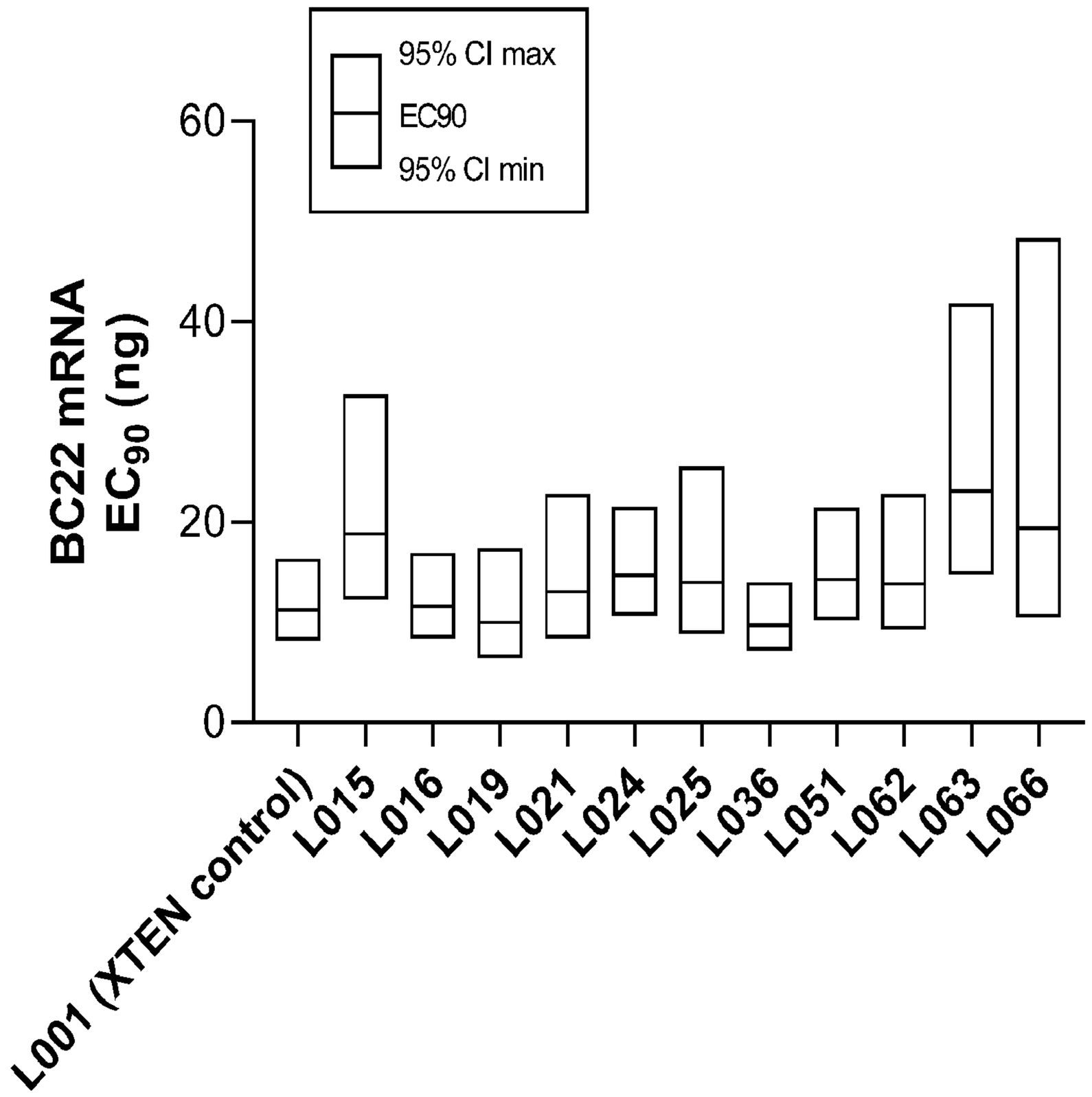

95% CI max
EC90
95% CI min
BC22 mRNA
EC90 (ng)
60
40
20
0
L001 (XTEN control)
L015
L016
L019
L021
L024
L025
L036
L051
L062
L063
L066

Fig. 42

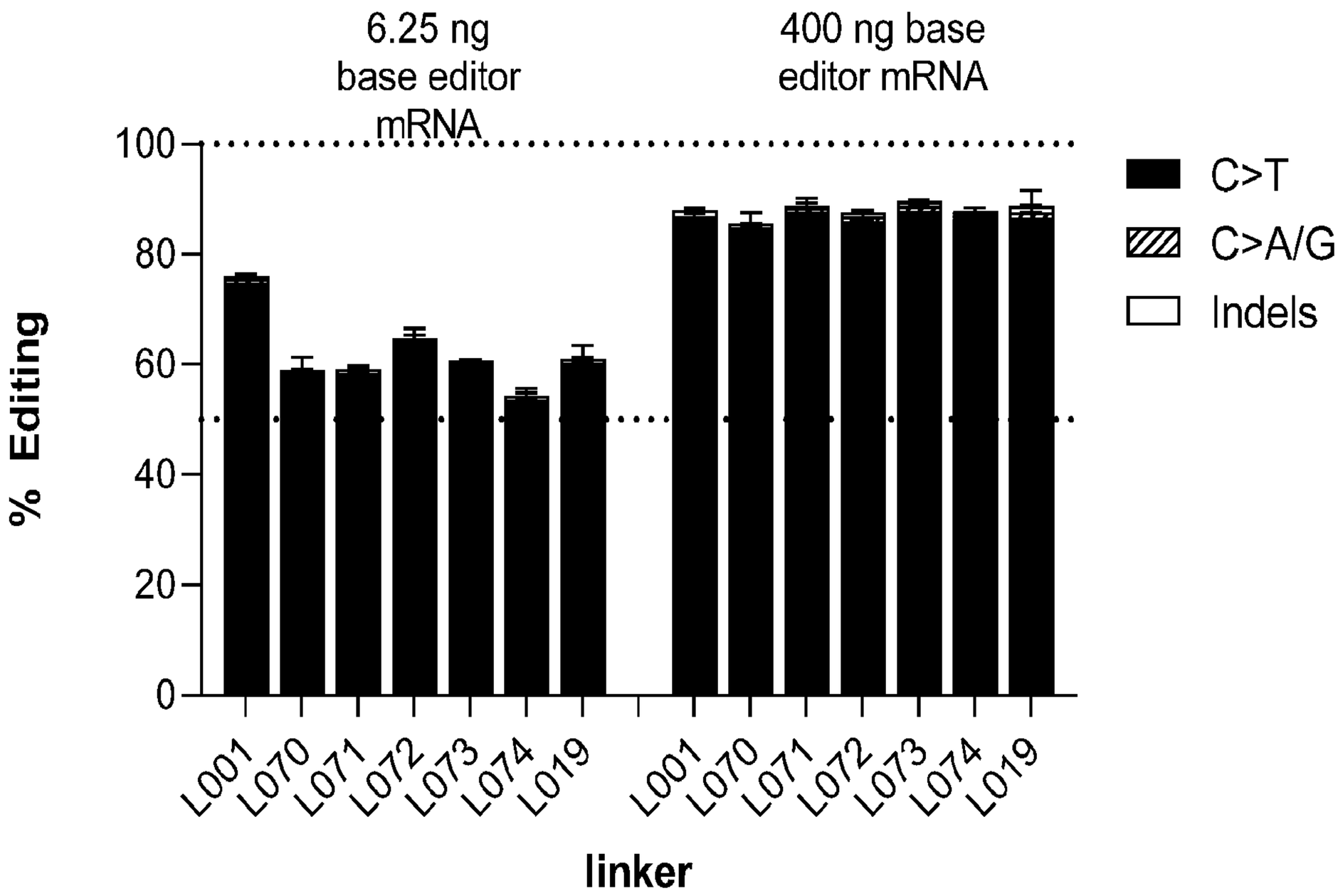

6.25 ng base editor mRNA
400 ng base editor mRNA
C>T
C>A/G
Indels
% Editing
100
80
60
40
20
0
L001
L070
L071
L072
L073
L074
L019
linker

Fig. 43

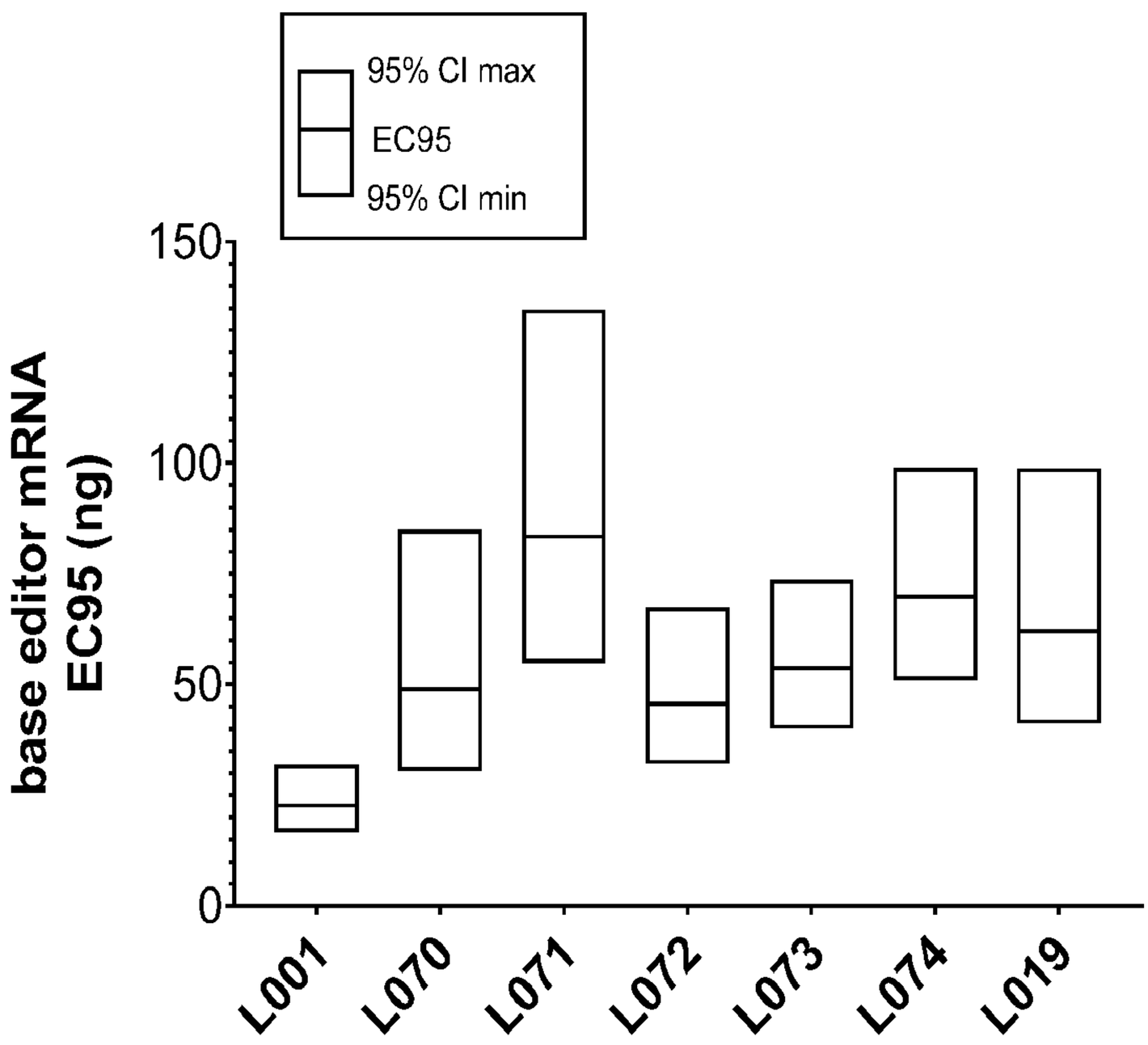

95% CI max
EC95
95% CI min
base editor mRNA EC95 (ng)
150
100
50
0
L001
L070
L071
L072
L073
L074
L019

Fig. 44

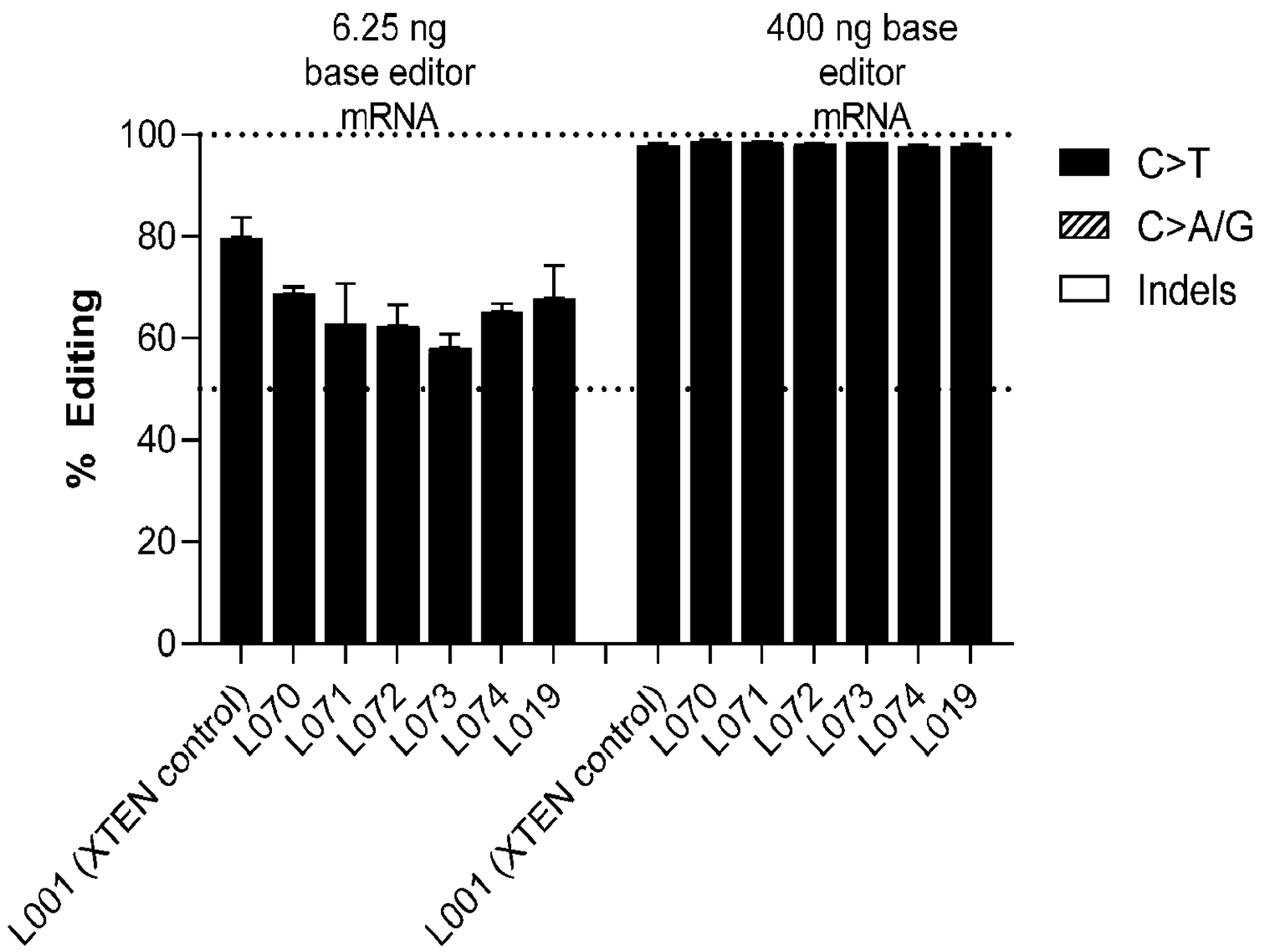

6.25 ng base editor mRNA
400 ng base editor mRNA
% Editing
100
80
60
40
20
0
C>T
C>A/G
Indels
L001 (XTEN control)
L070
L071
L072
L073
L074
L019

Fig. 45

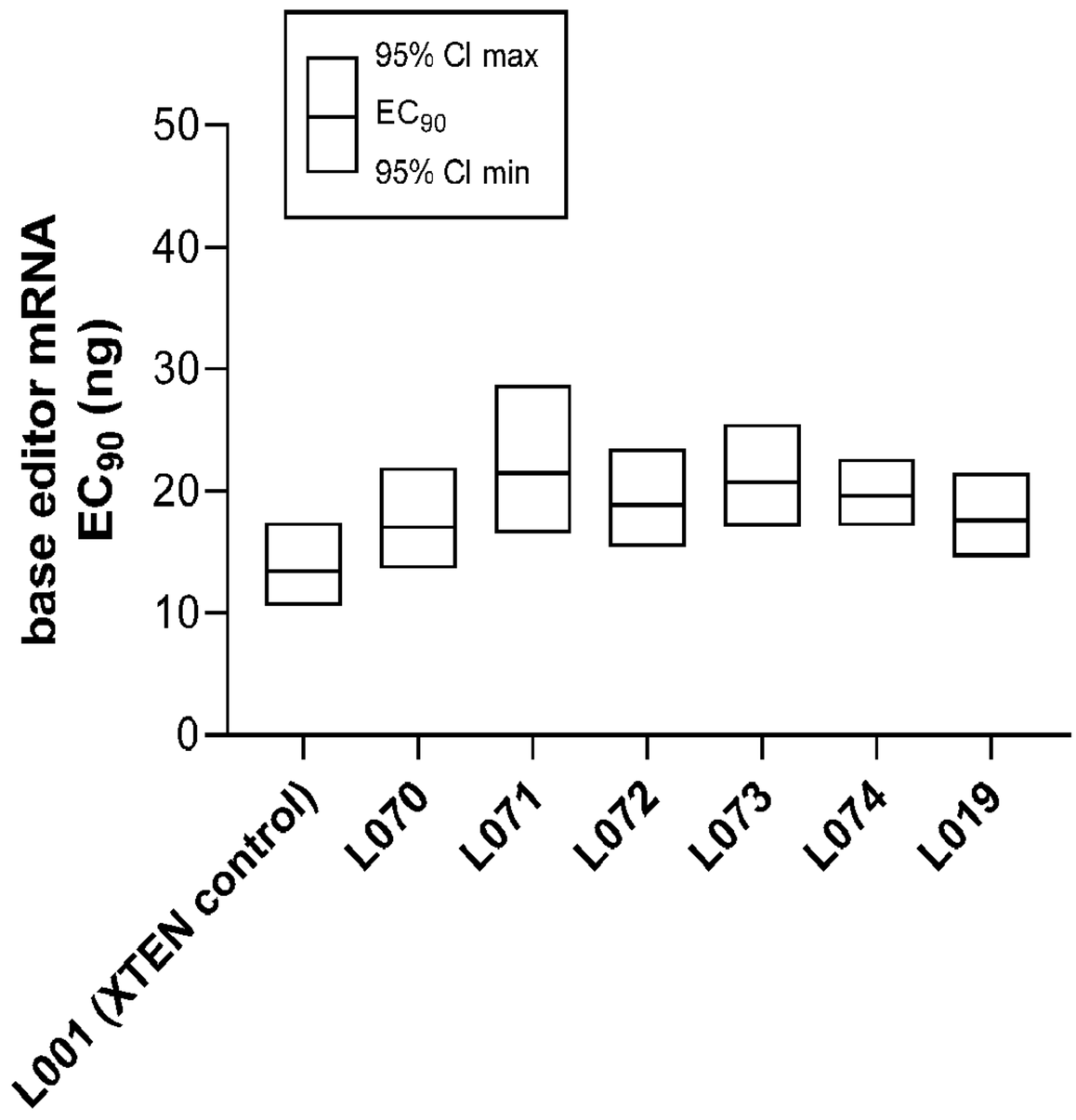

95% CI max
EC90
95% CI min
base editor mRNA EC90 (ng)
50
40
30
20
10
0
L001 (XTEN control)
L070
L071
L072
L073
L074
L019

Fig. 46

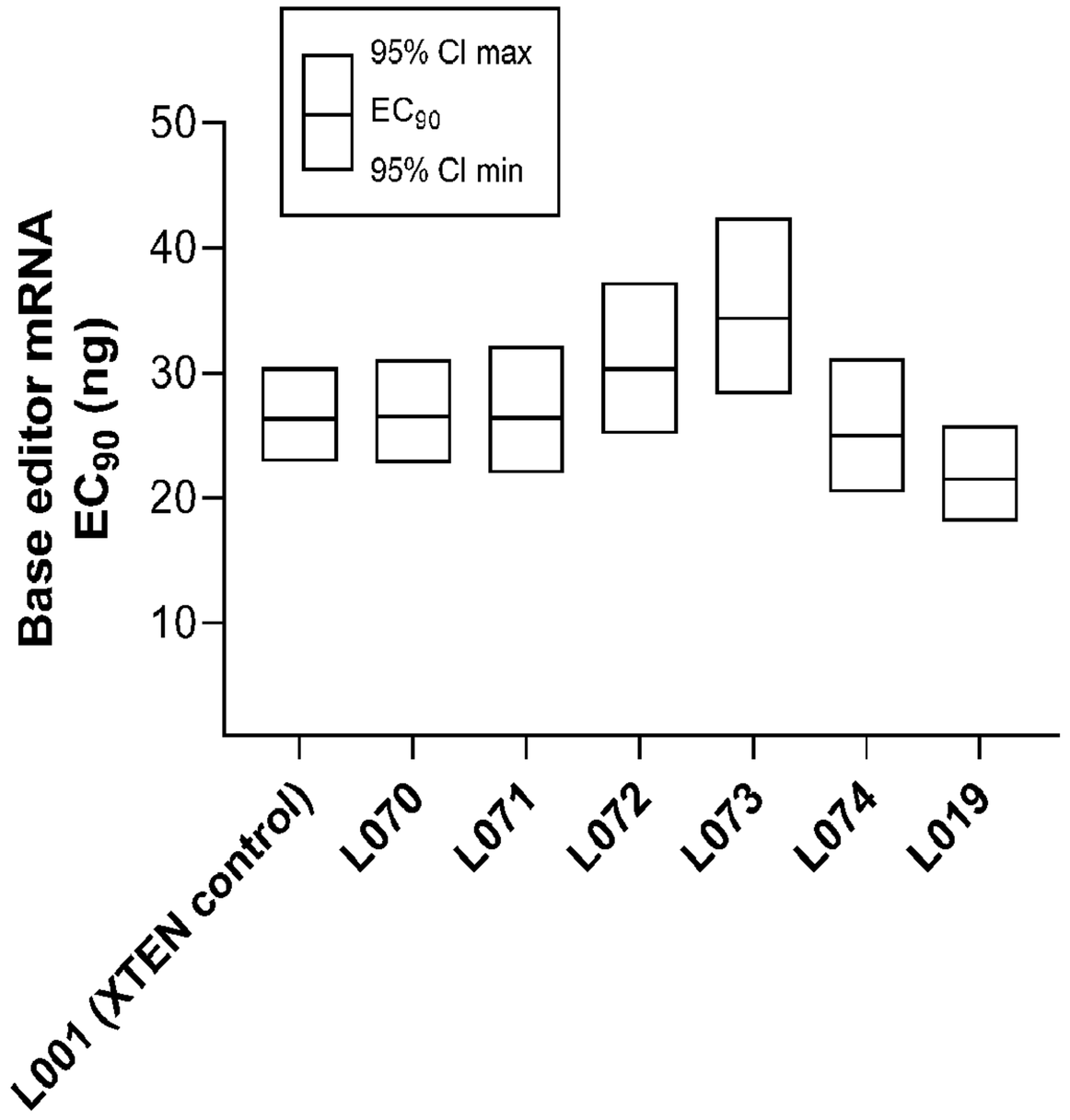

95% CI max
EC90
95% CI min
Base editor mRNA EC90 (ng)
50
40
30
20
10
L001 (XTEN control)
L070
L071
L072
L073
L074
L019

Fig. 47

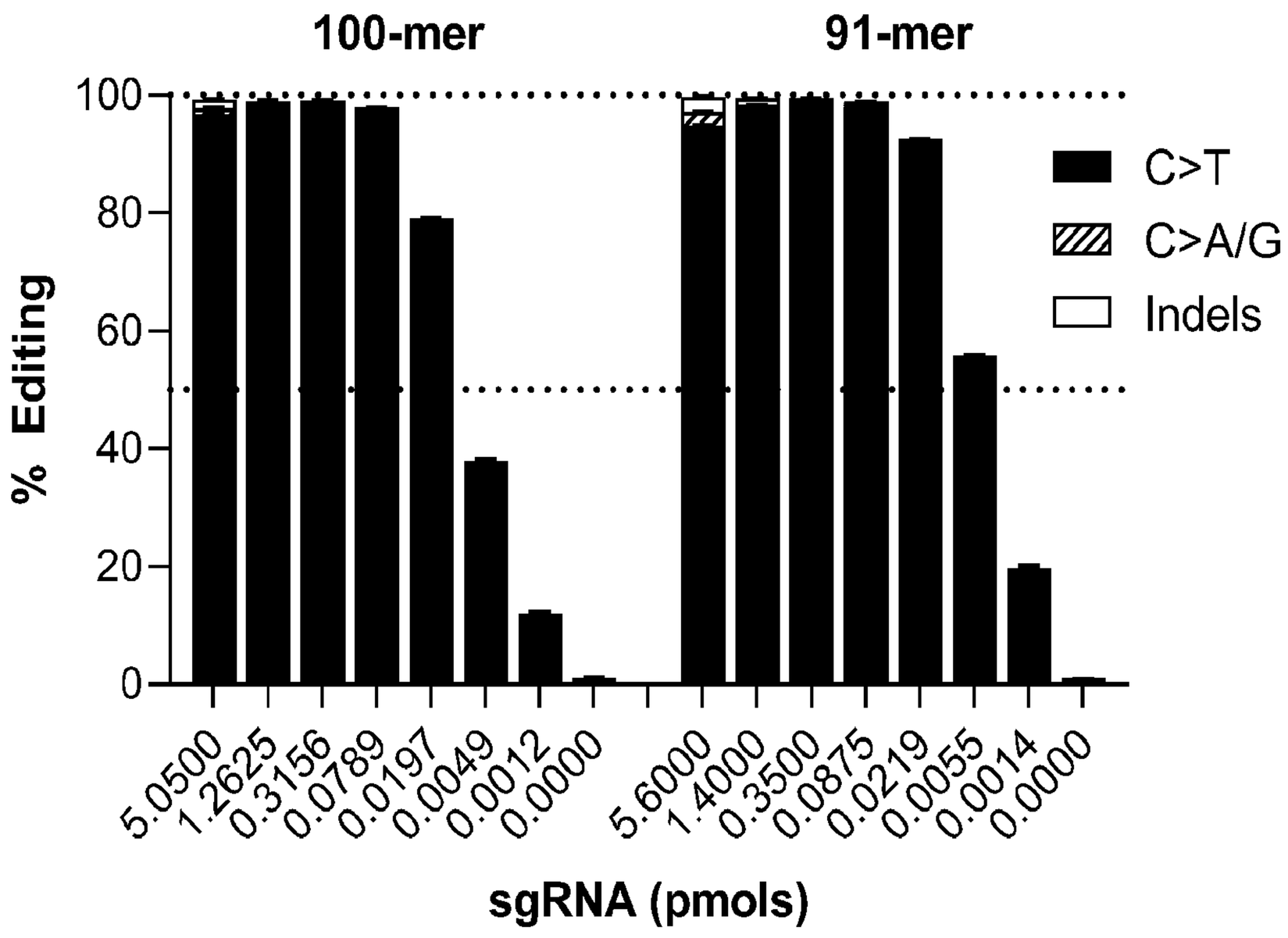

100-mer
91-mer
% Editing
100
80
60
40
20
0
C>T
C>A/G
Indels
5.0500
1.2625
0.3156
0.0789
0.0197
0.0049
0.0012
0.0000
5.6000
1.4000
0.3500
0.0875
0.0219
0.0055
0.0014
0.0000
sgRNA (pmols)

Fig. 48

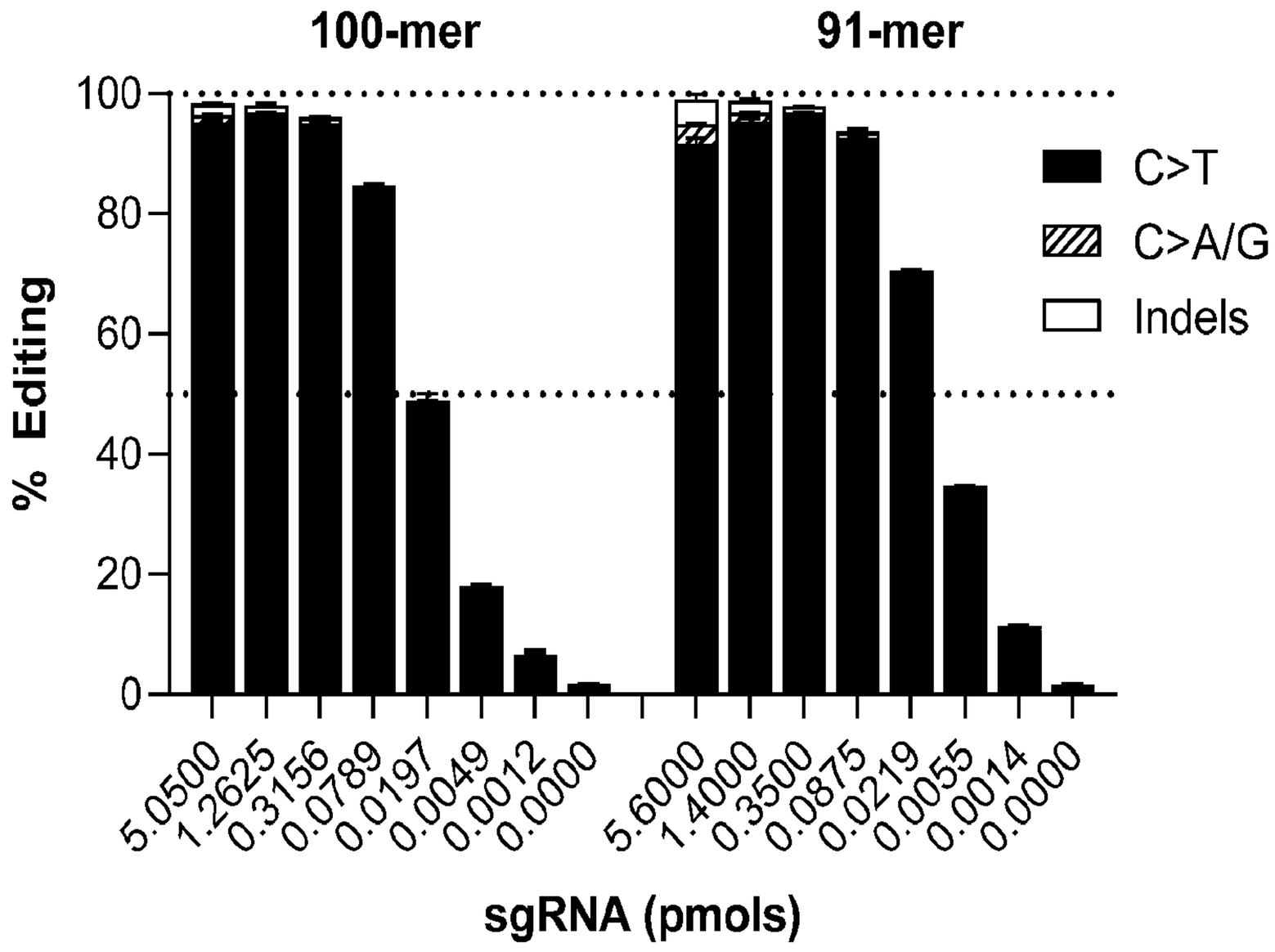


Fig. 49A

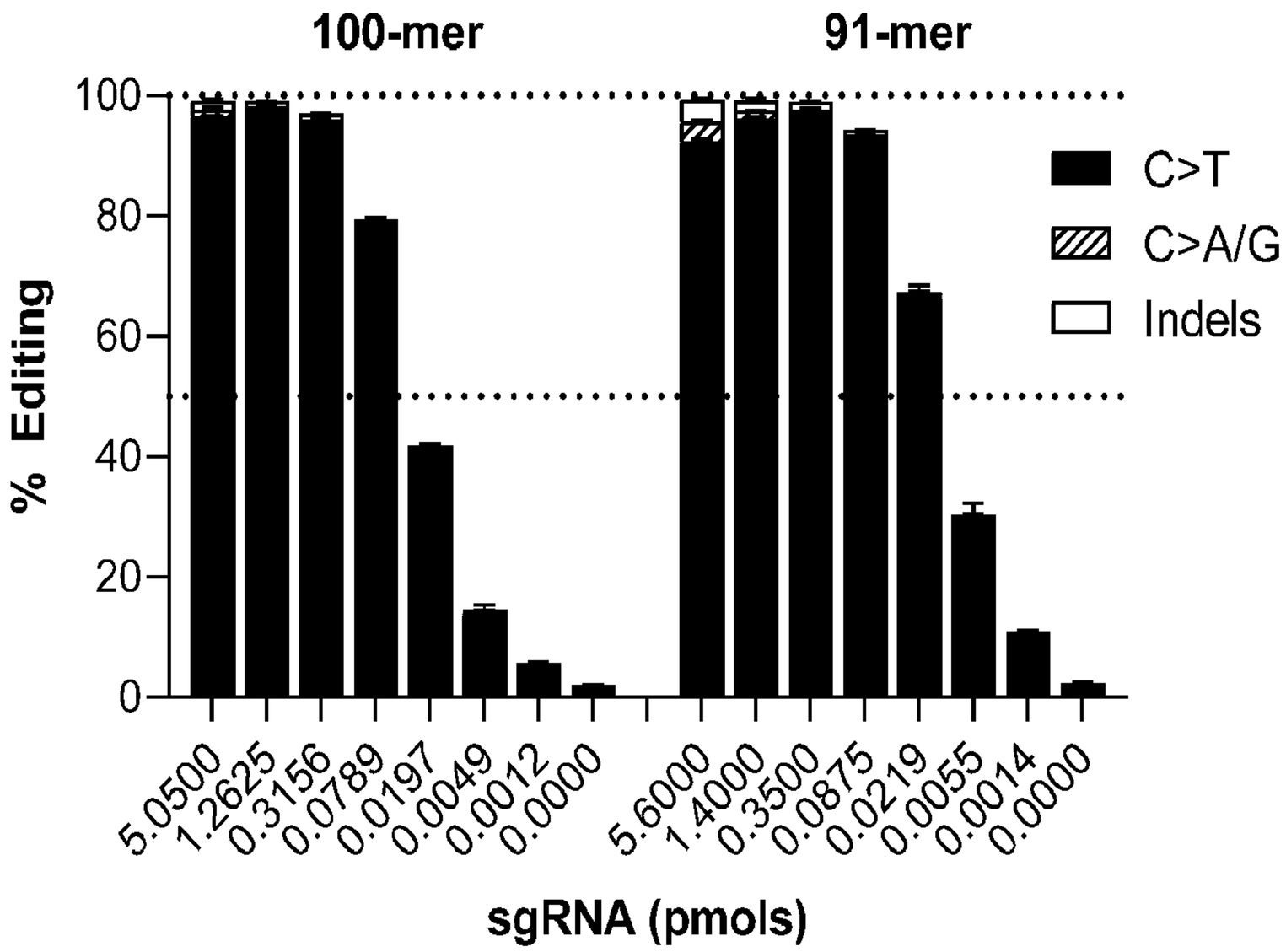


Fig. 49B

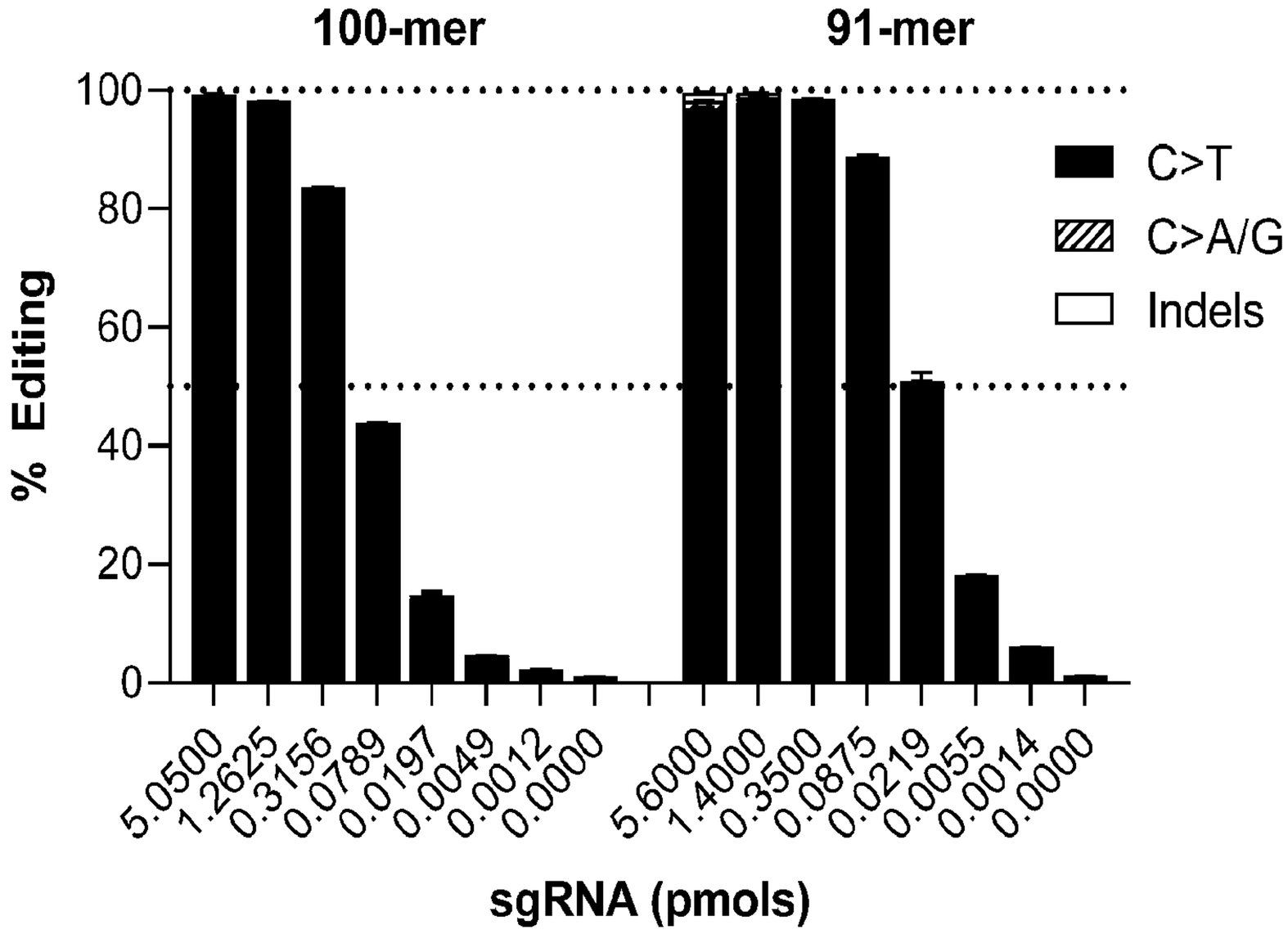

100-mer
91-mer
% Editing
100
80
60
40
20
0
C>T
C>A/G
Indels
5.0500
1.2625
0.3156
0.0789
0.0197
0.0049
0.0012
0.0000
5.6000
1.4000
0.3500
0.0875
0.0219
0.0055
0.0014
0.0000
sgRNA (pmols)

Fig. 50

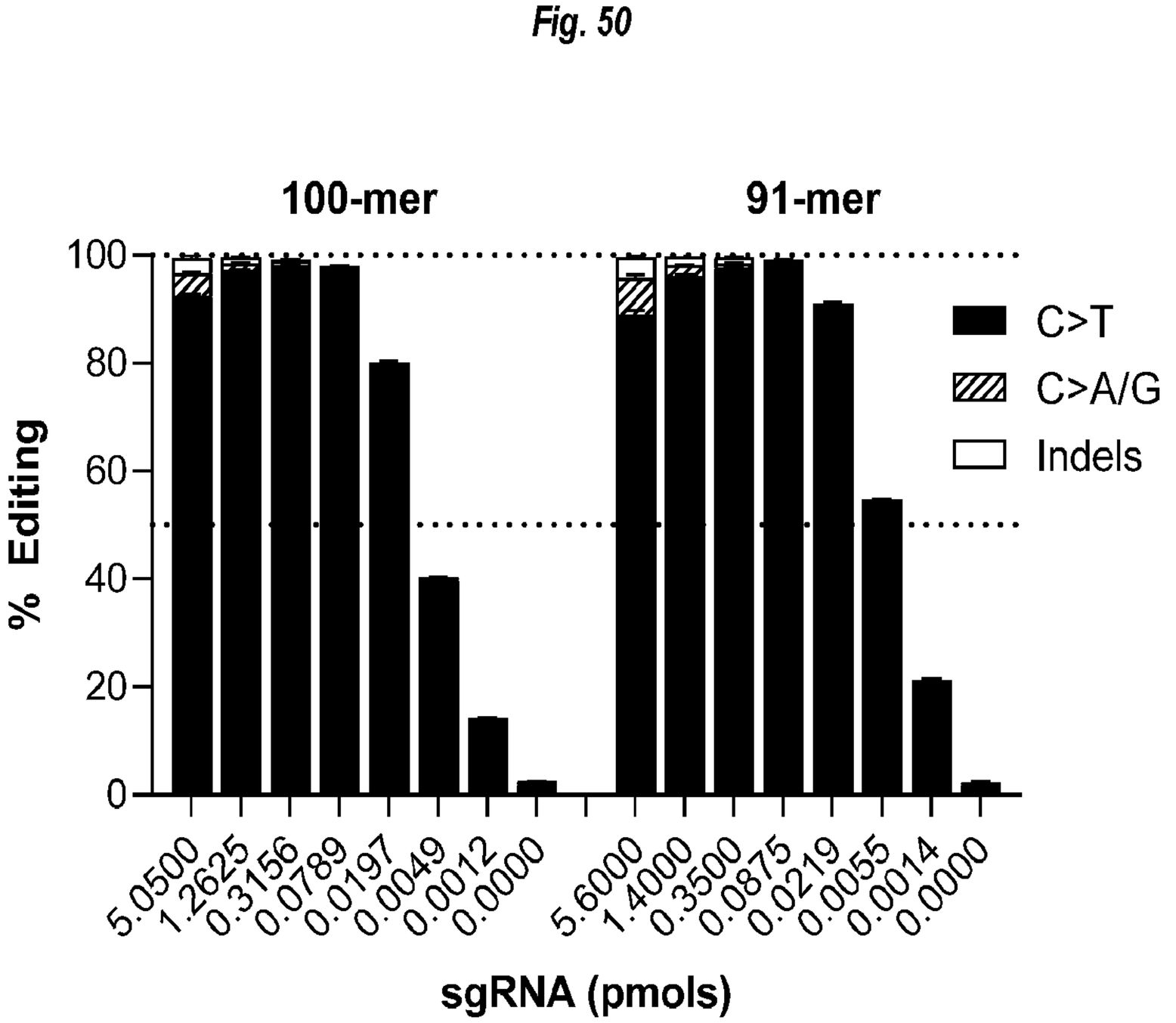

100-mer
91-mer
% Editing
100
80
60
40
20
0
C>T
C>A/G
Indels
5.0500
1.2625
0.3156
0.0789
0.0197
0.0049
0.0012
0.0000
5.6000
1.4000
0.3500
0.0875
0.0219
0.0055
0.0014
0.0000
sgRNA (pmols)

Fig. 51

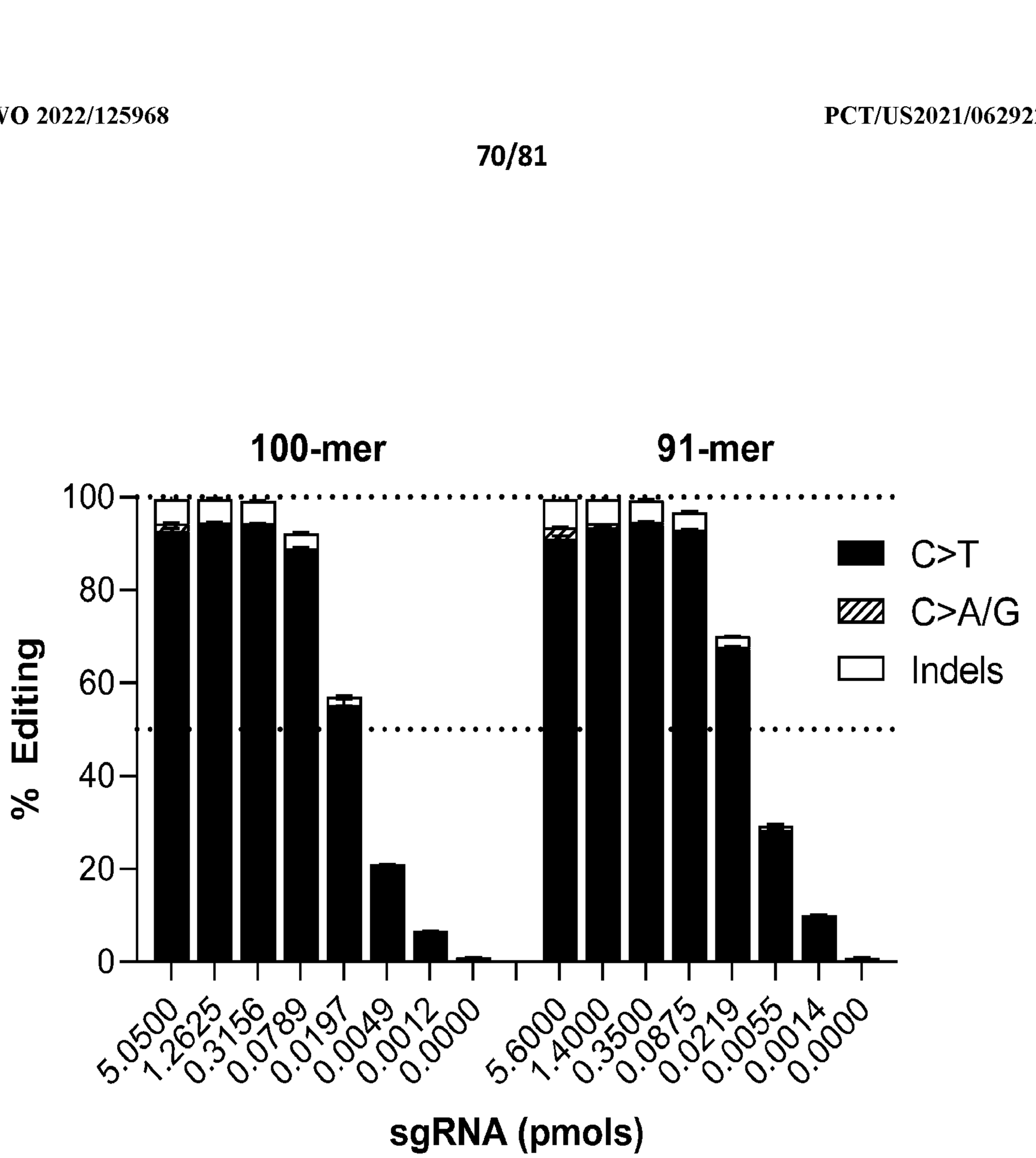

100-mer
91-mer
% Editing
100
80
60
40
20
0
C>T
C>A/G
Indels
5.0500
1.2625
0.3156
0.0789
0.0197
0.0049
0.0012
0.0000
5.6000
1.4000
0.3500
0.0875
0.0219
0.0055
0.0014
0.0000
sgRNA (pmols)

Fig. 52

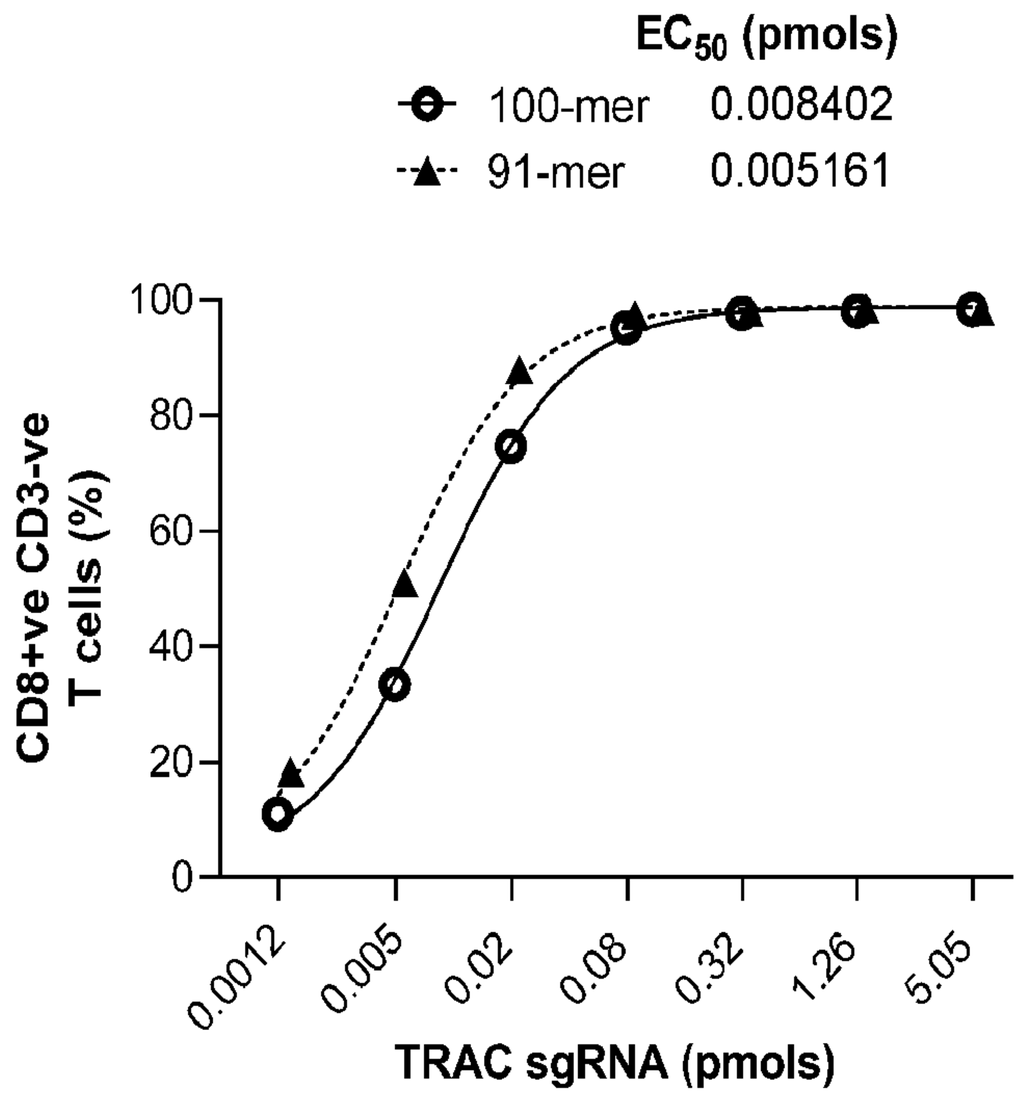


Fig. 53A

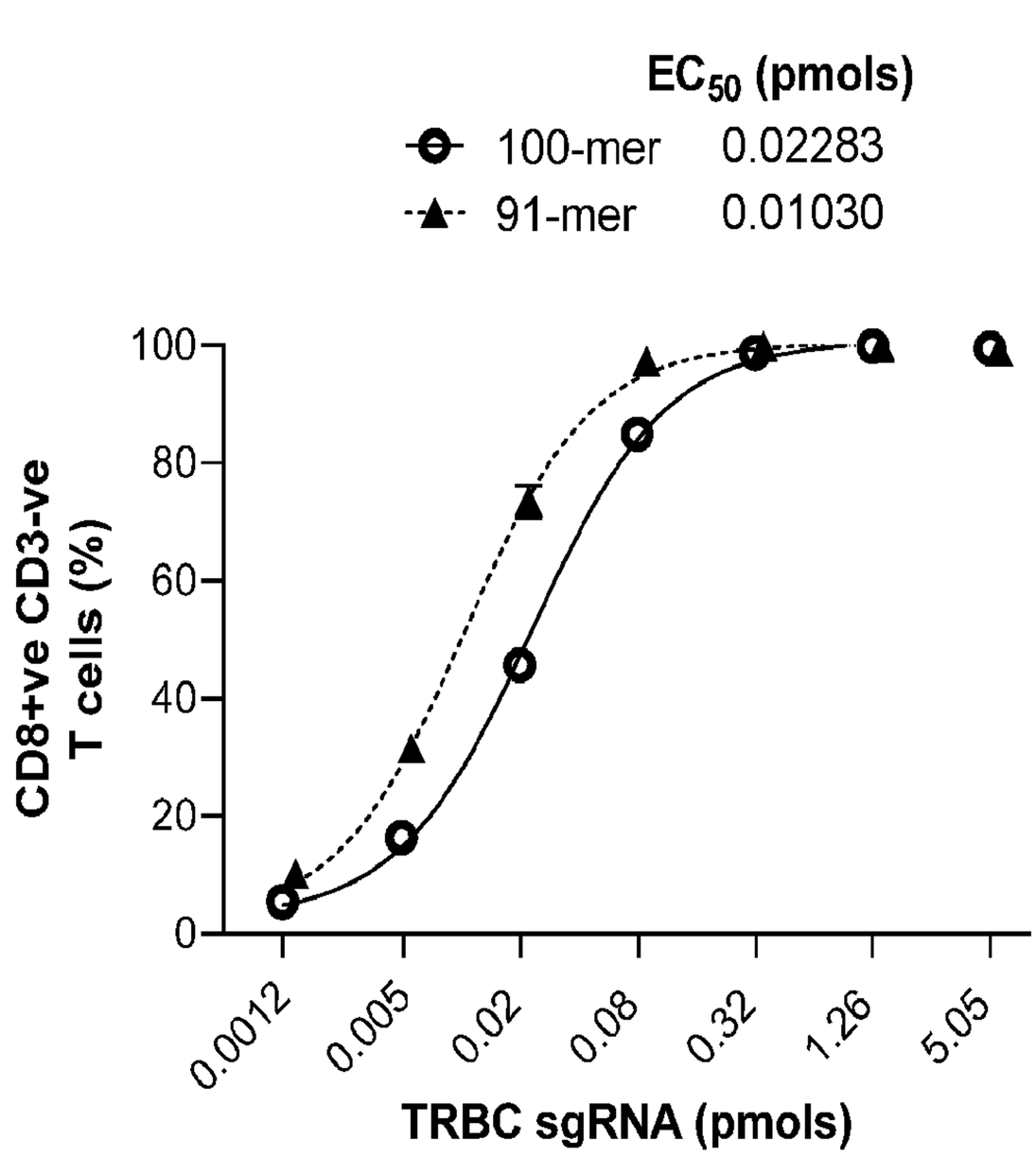


Fig. 53B

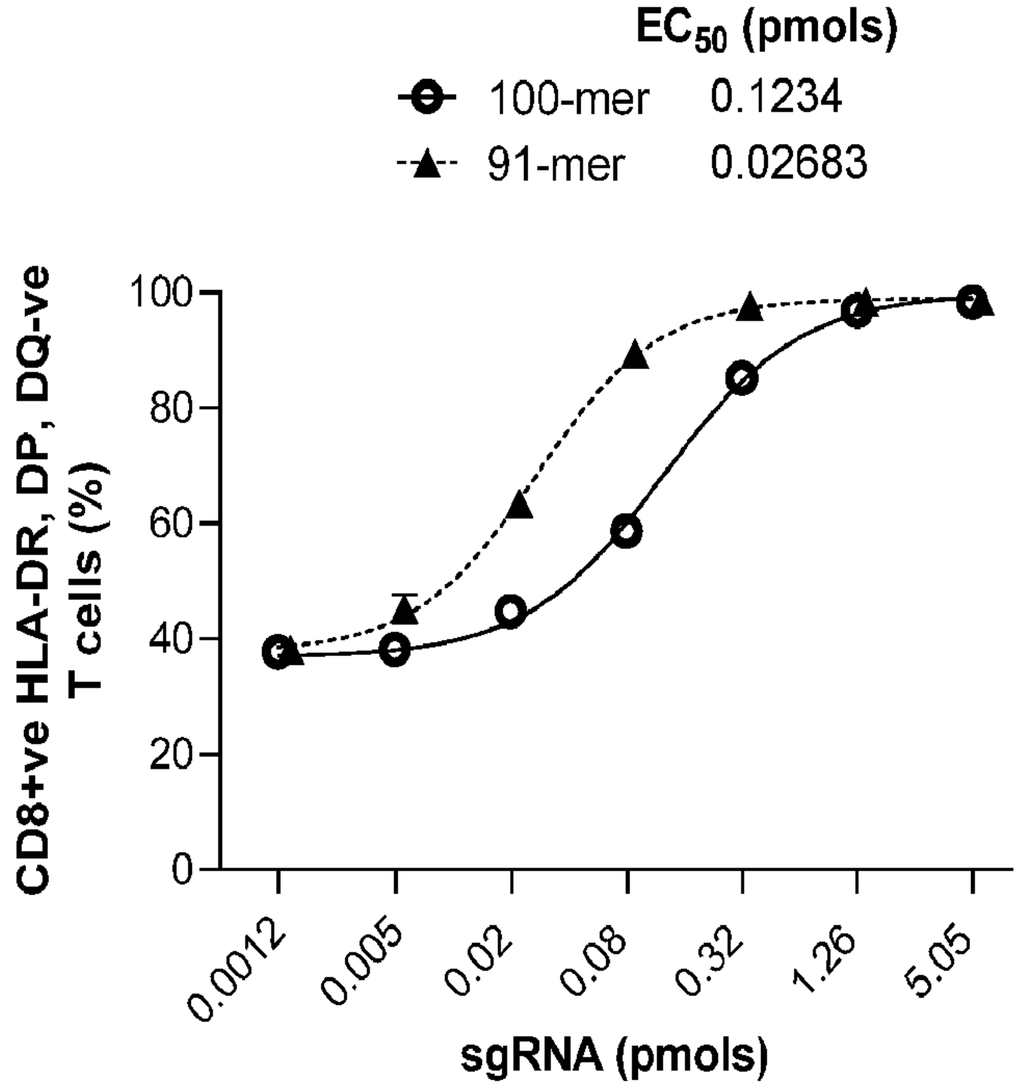


Fig. 54A

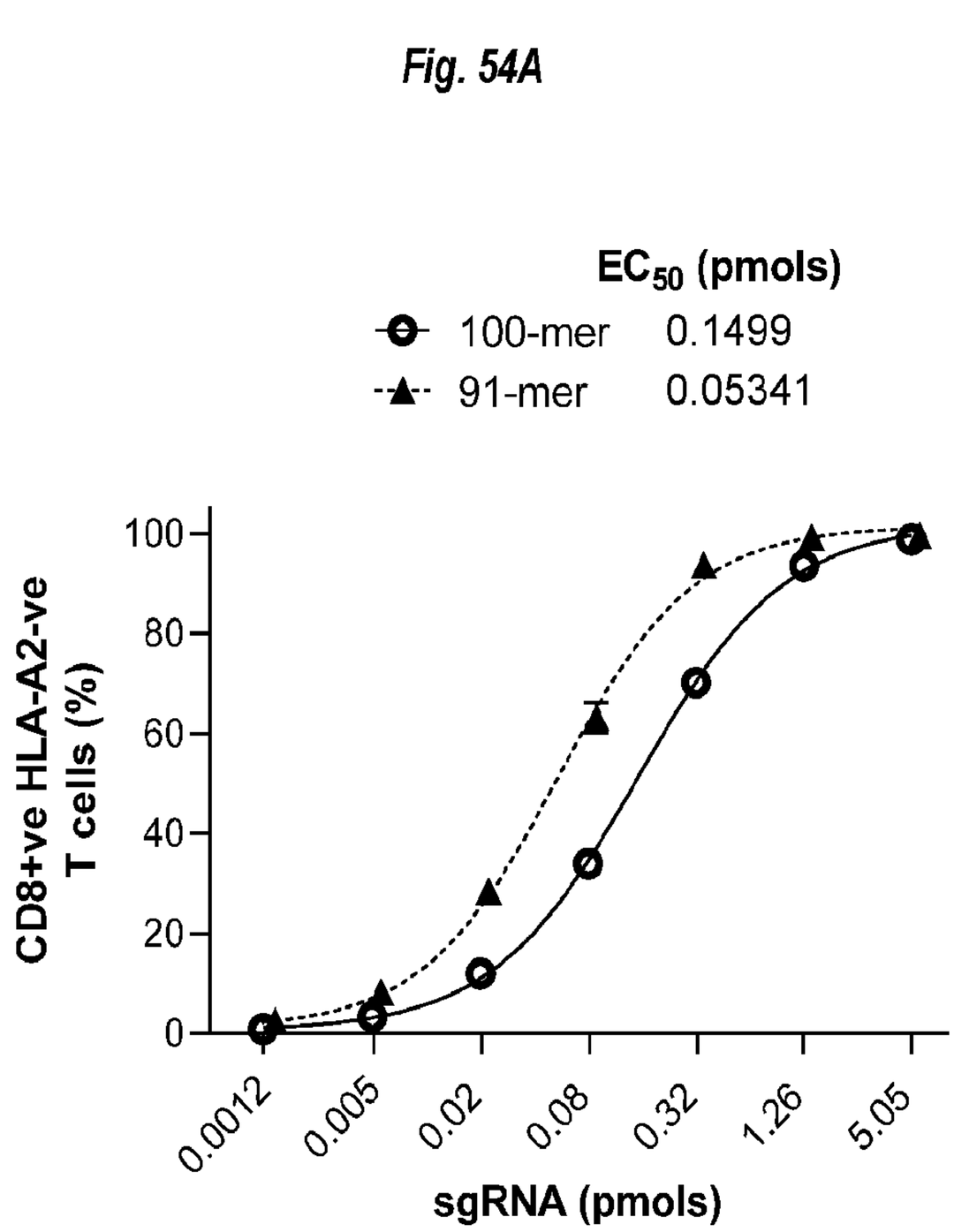


Fig. 54B

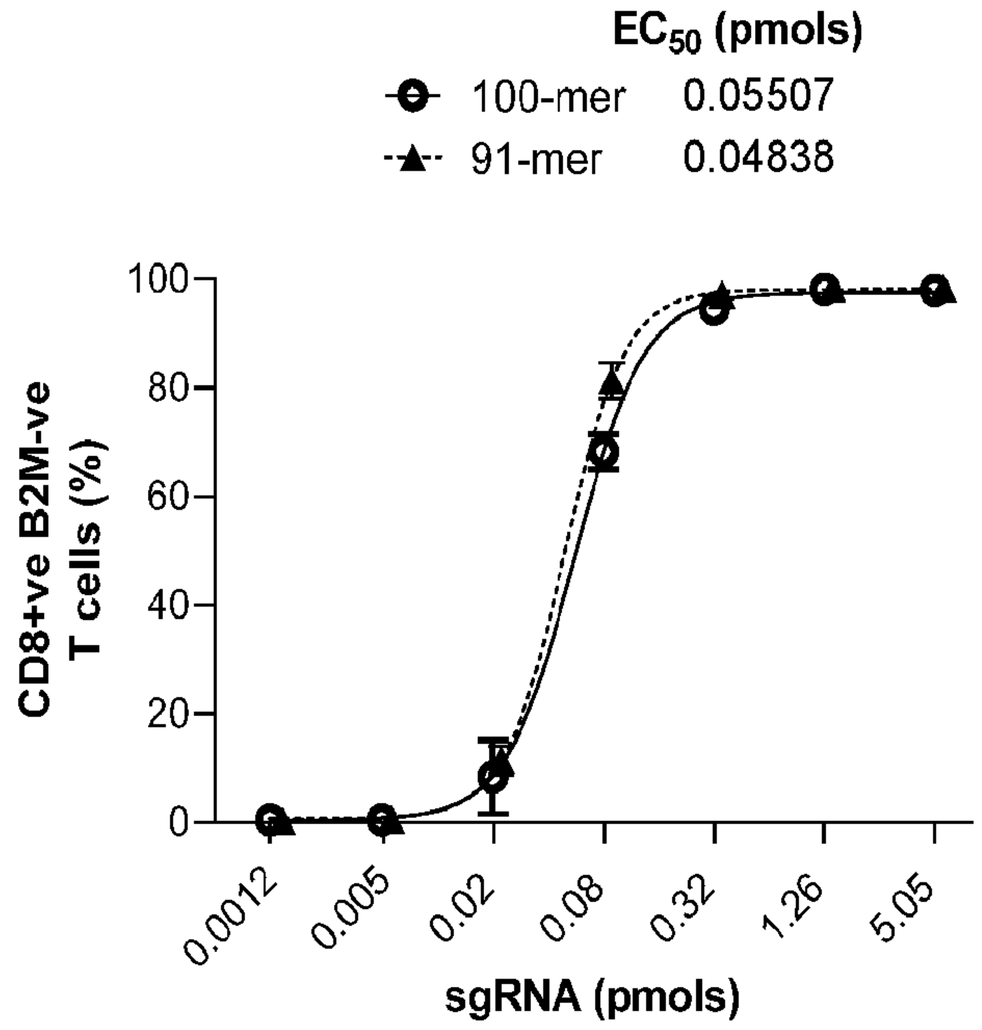


Fig. 55A

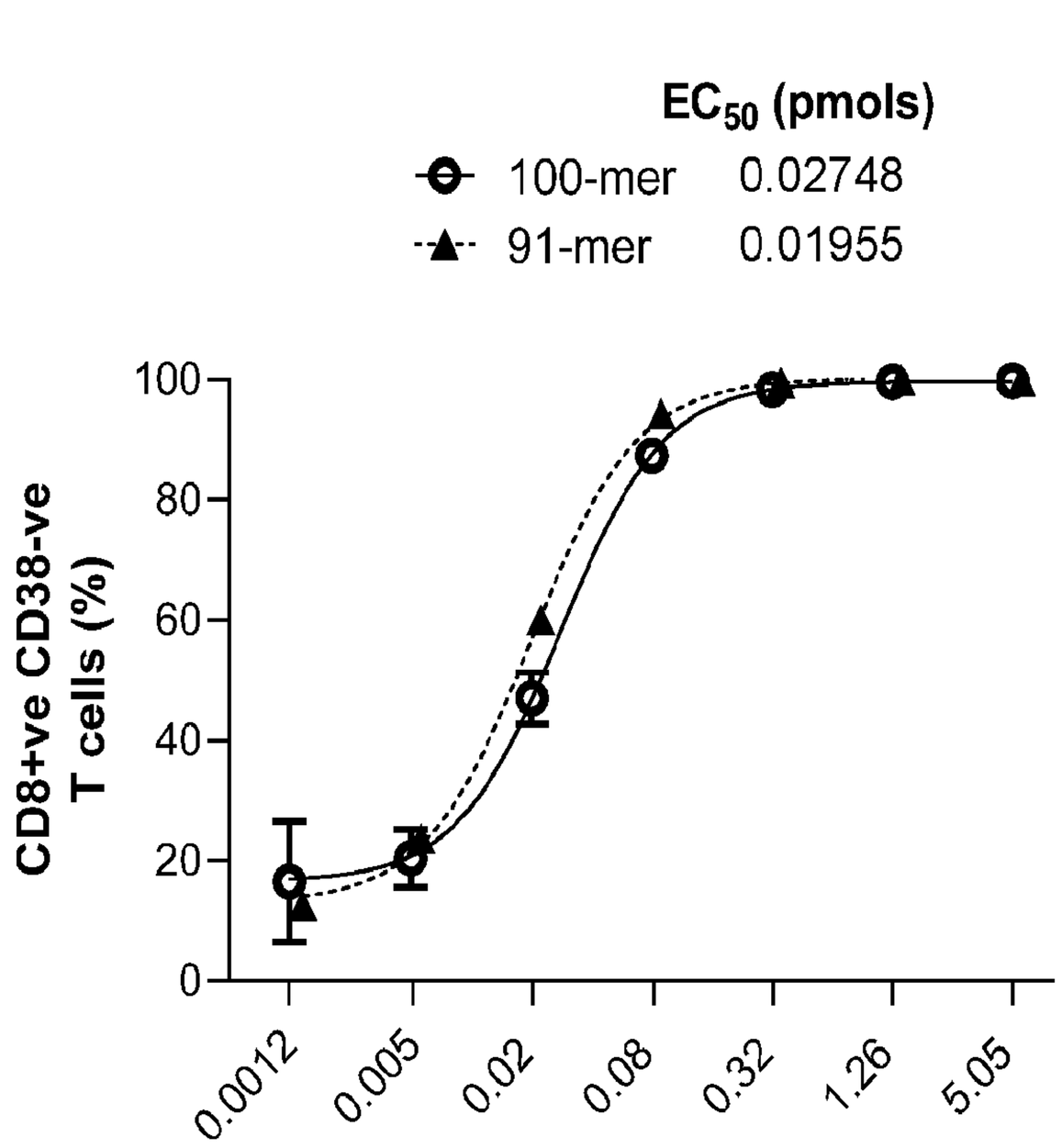


Fig. 55B

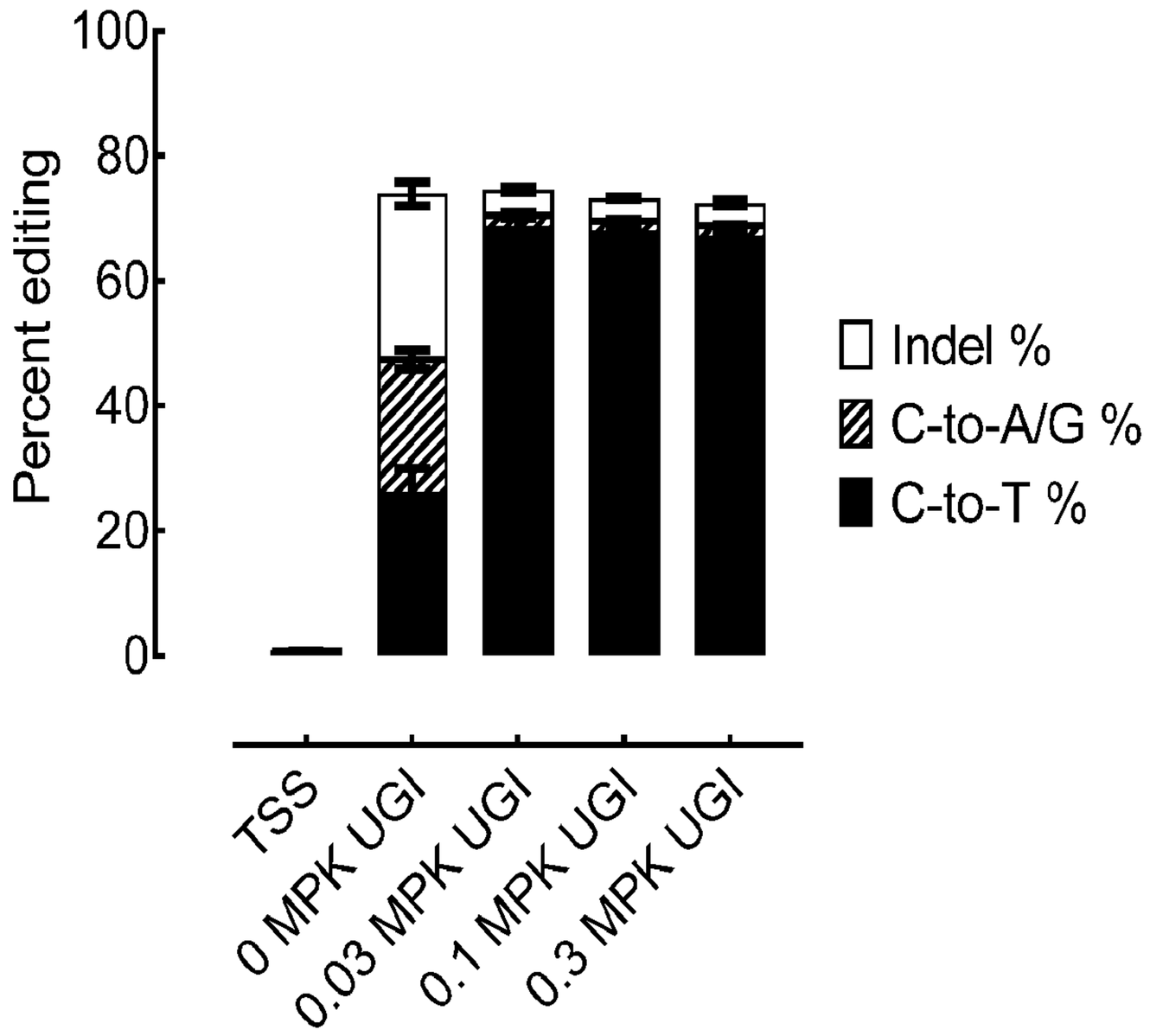

Percent editing
100
80
60
40
20
0
Indel %
C-to-A/G %
C-to-T %
TSS
0 MPK UGI
0.03 MPK UGI
0.1 MPK UGI
0.3 MPK UGI

Fig. 56

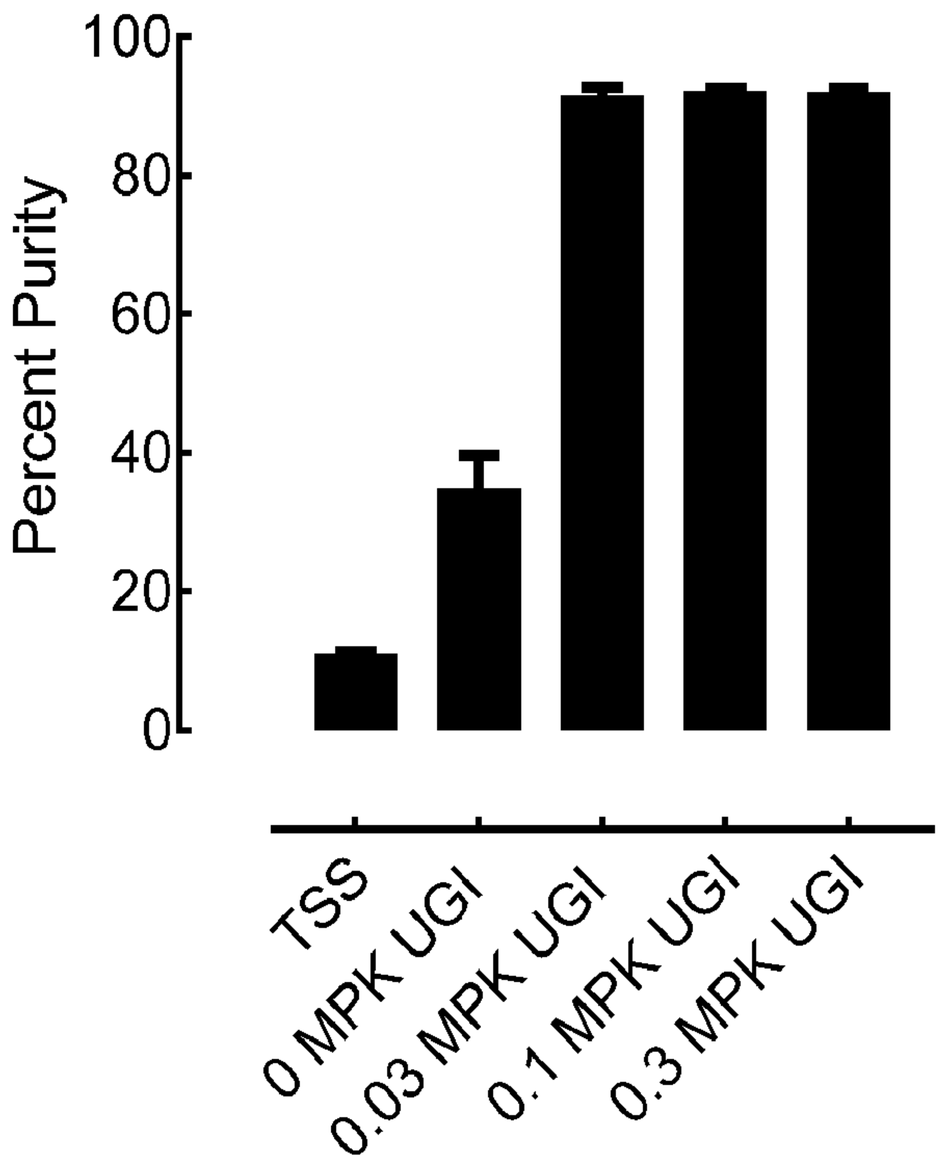

Percent Purity
100
80
60
40
20
0
TSS
0 MPK UGI
0.03 MPK UGI
0.1 MPK UGI
0.3 MPK UGI

Fig. 57

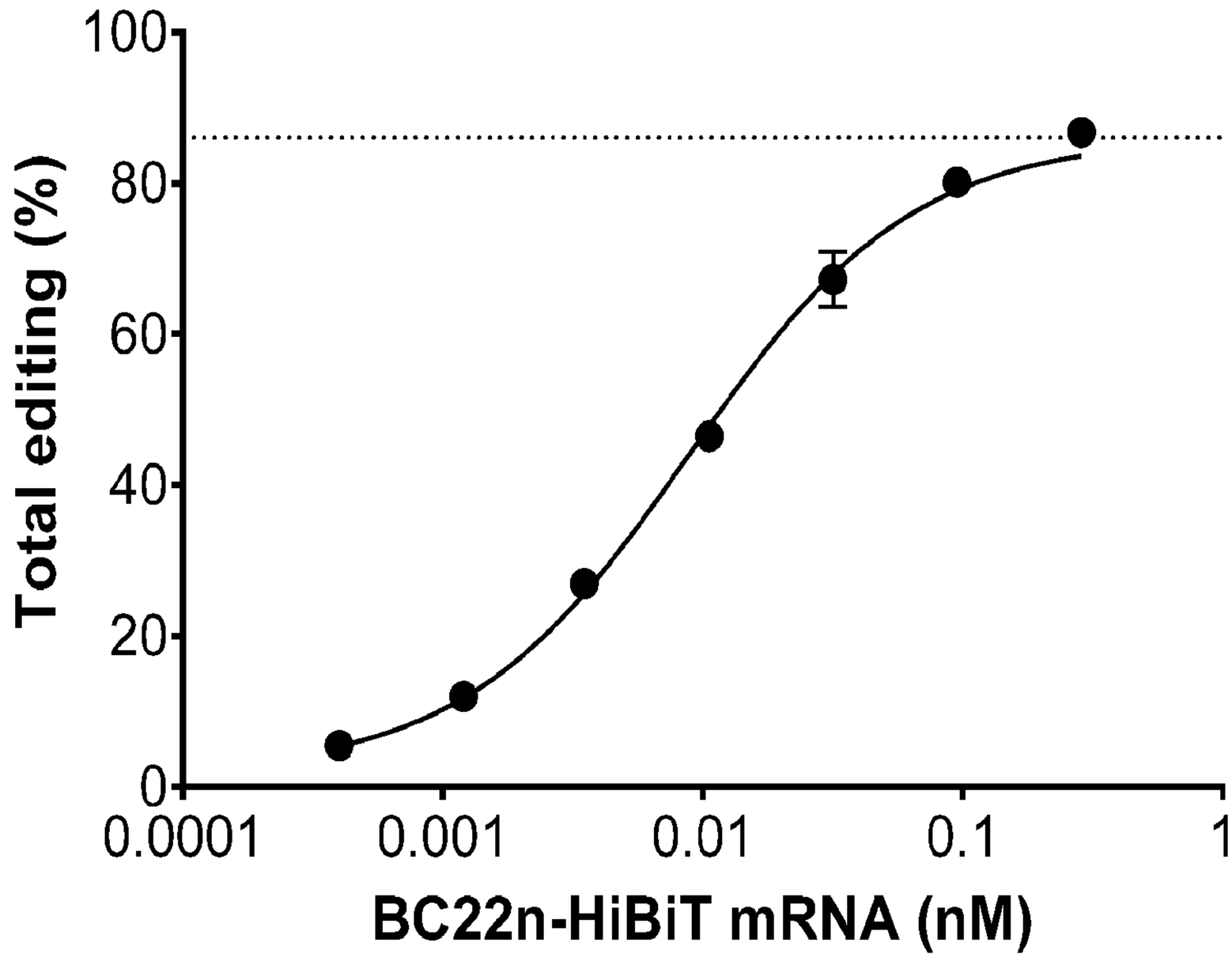


Fig. 58A

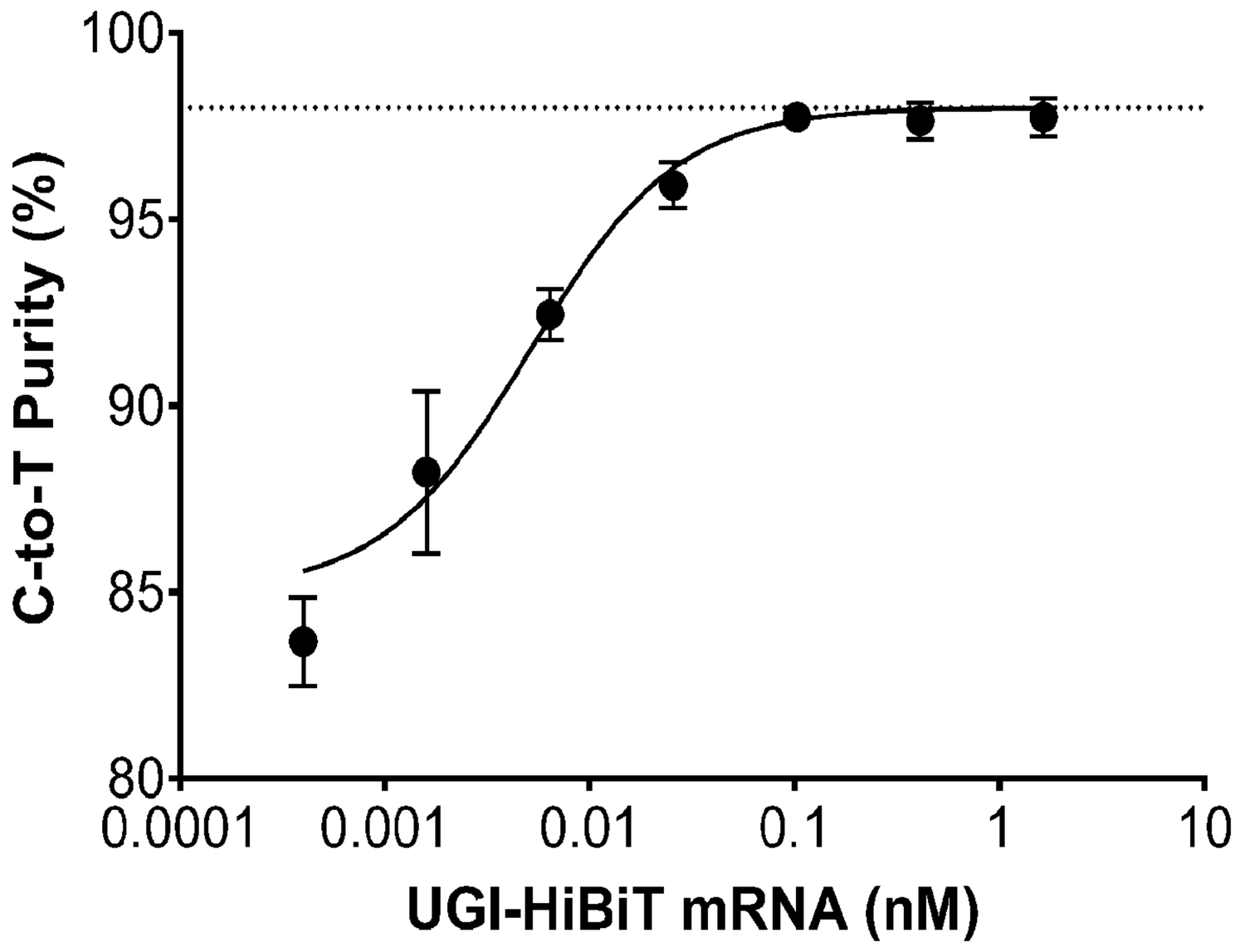


Fig. 58B

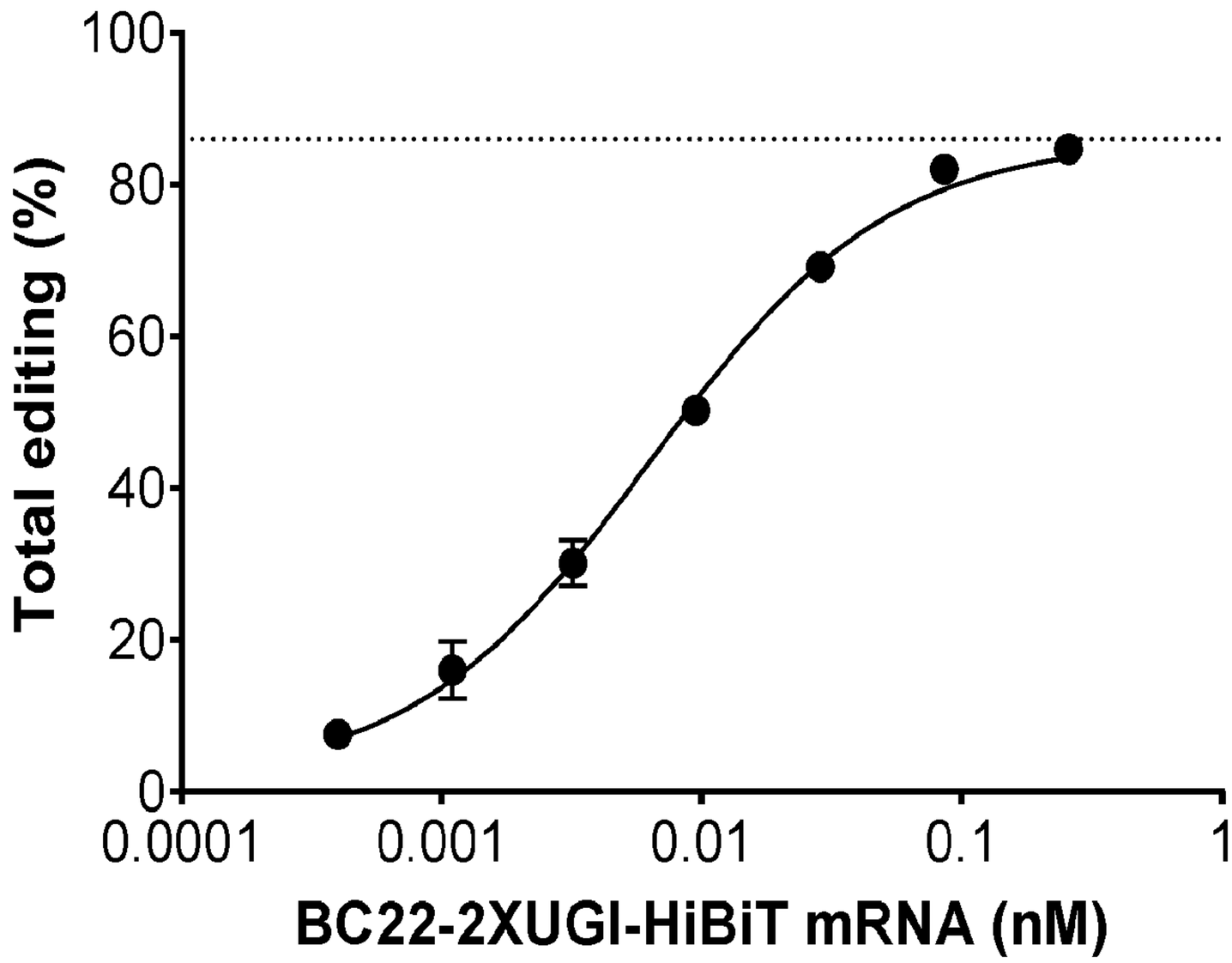


Fig. 58C

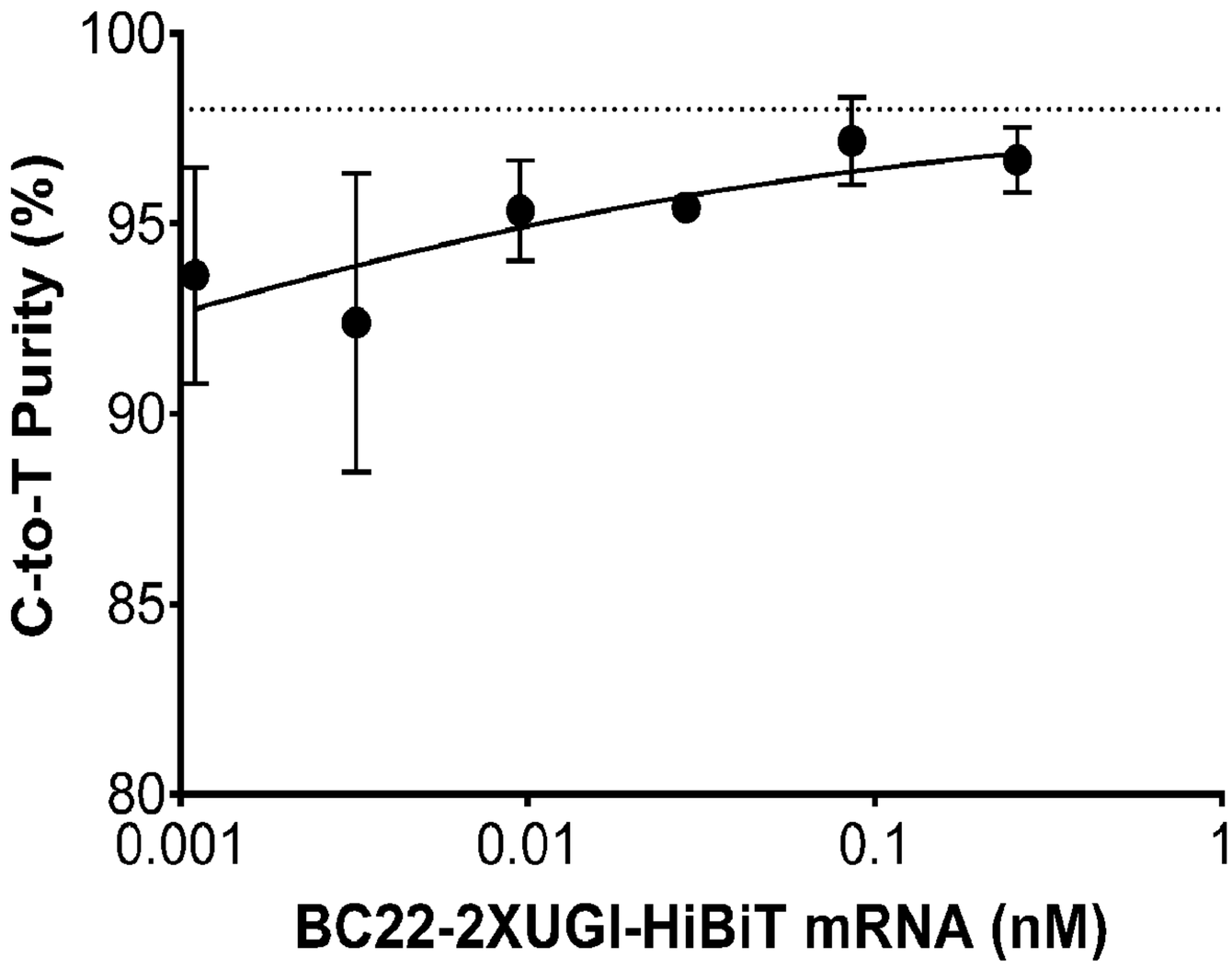


Fig. 58D

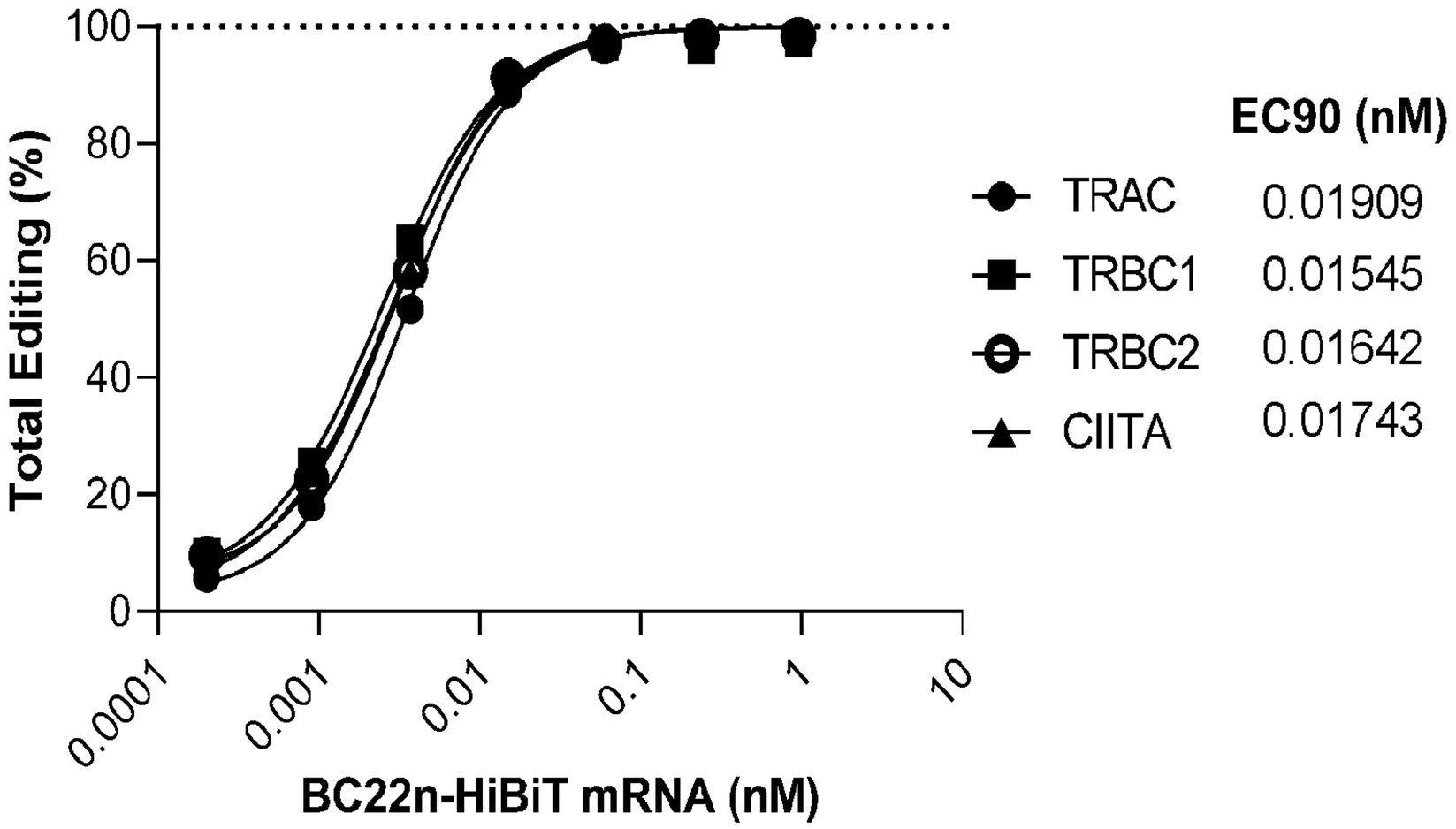


Fig. 59A

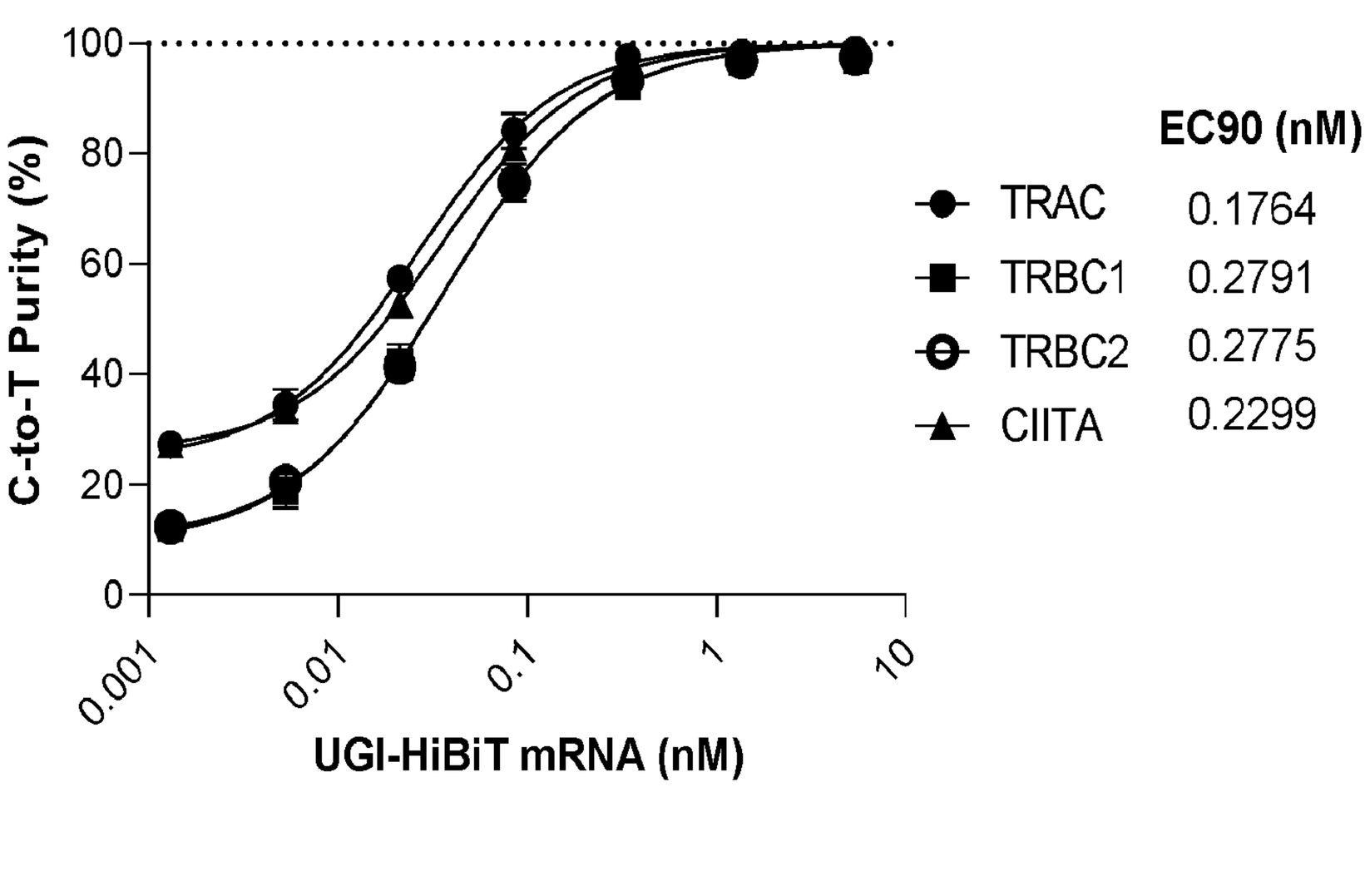


Fig. 59B

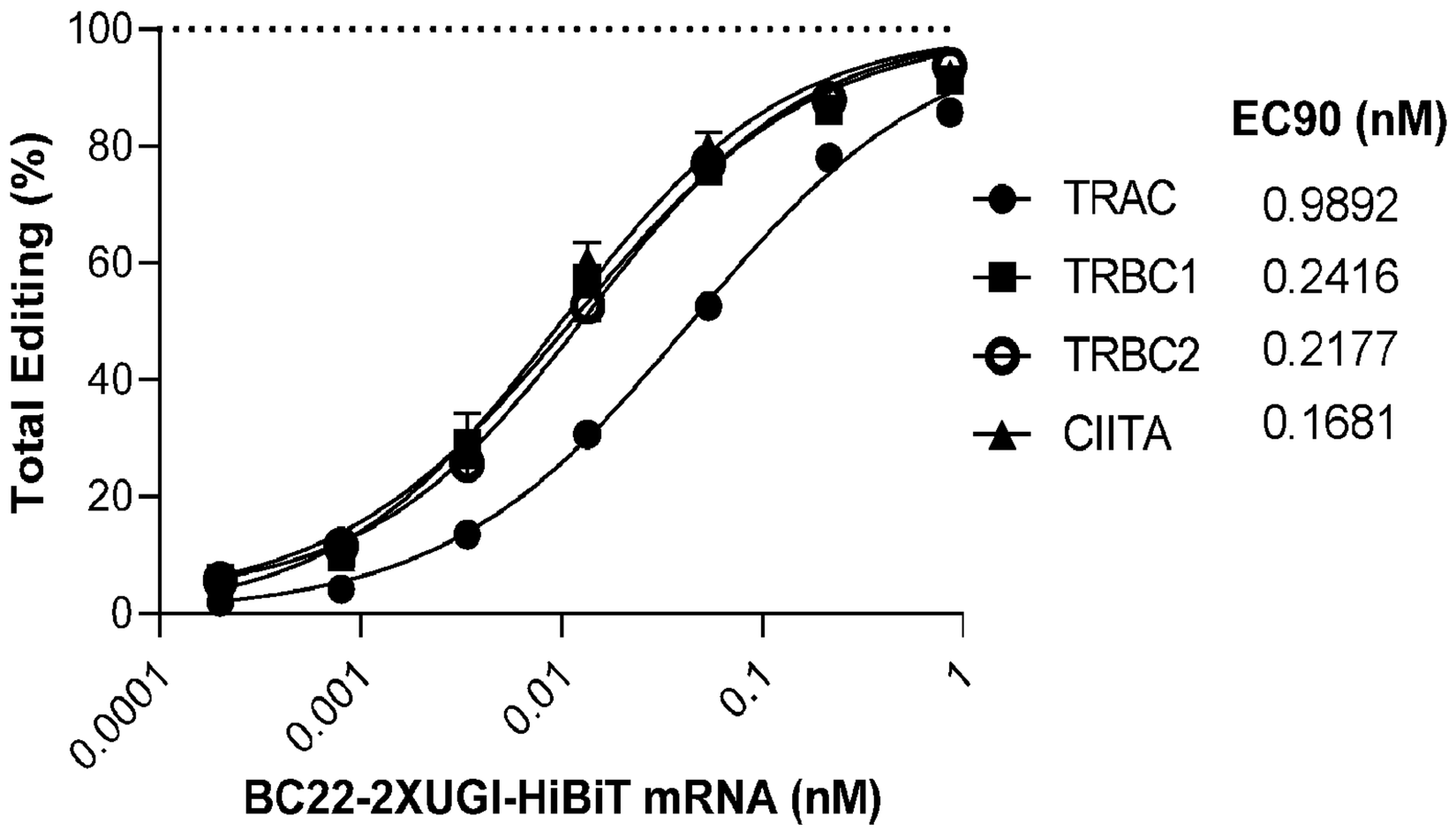


Fig. 59C

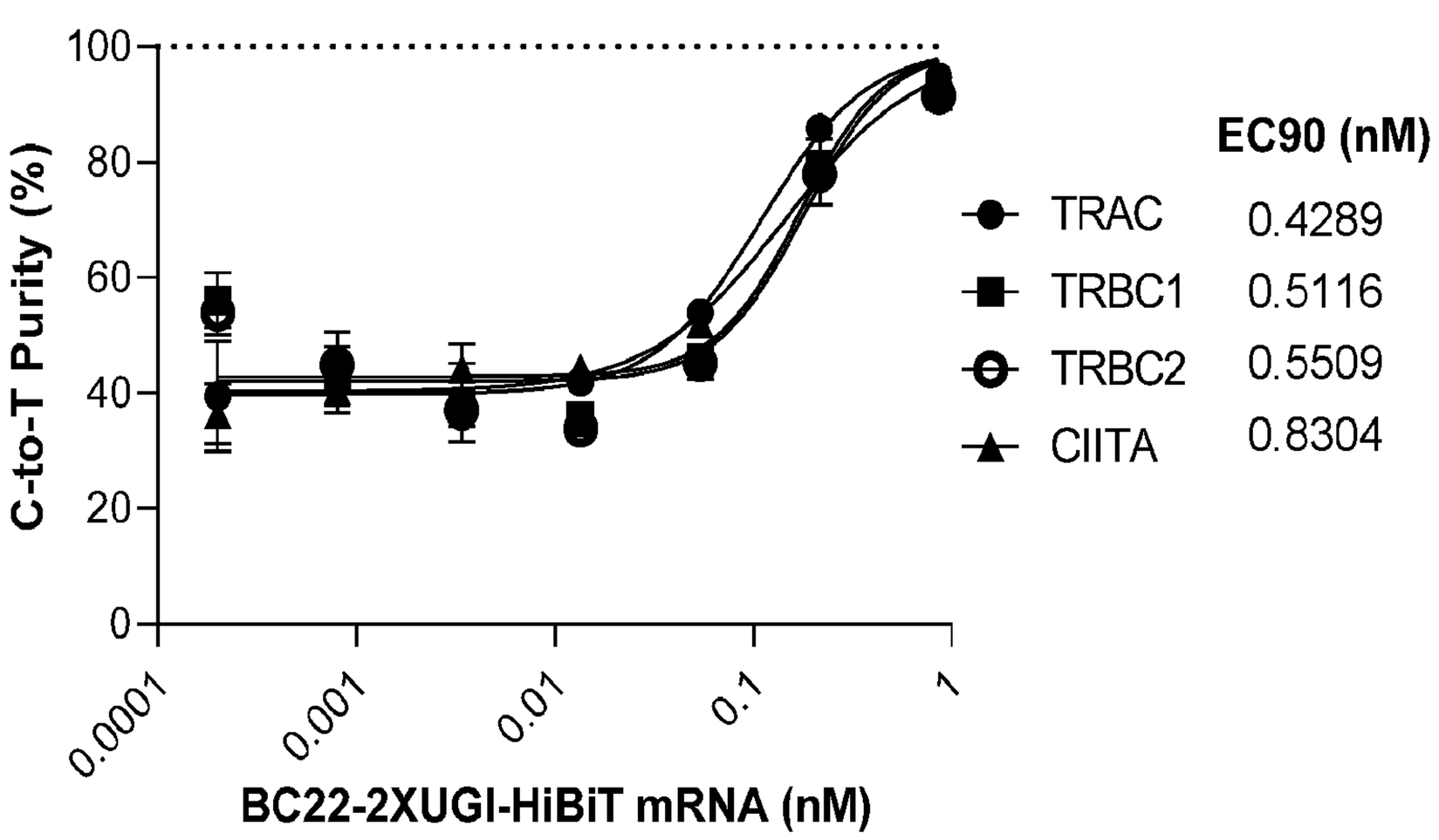


Fig. 59D

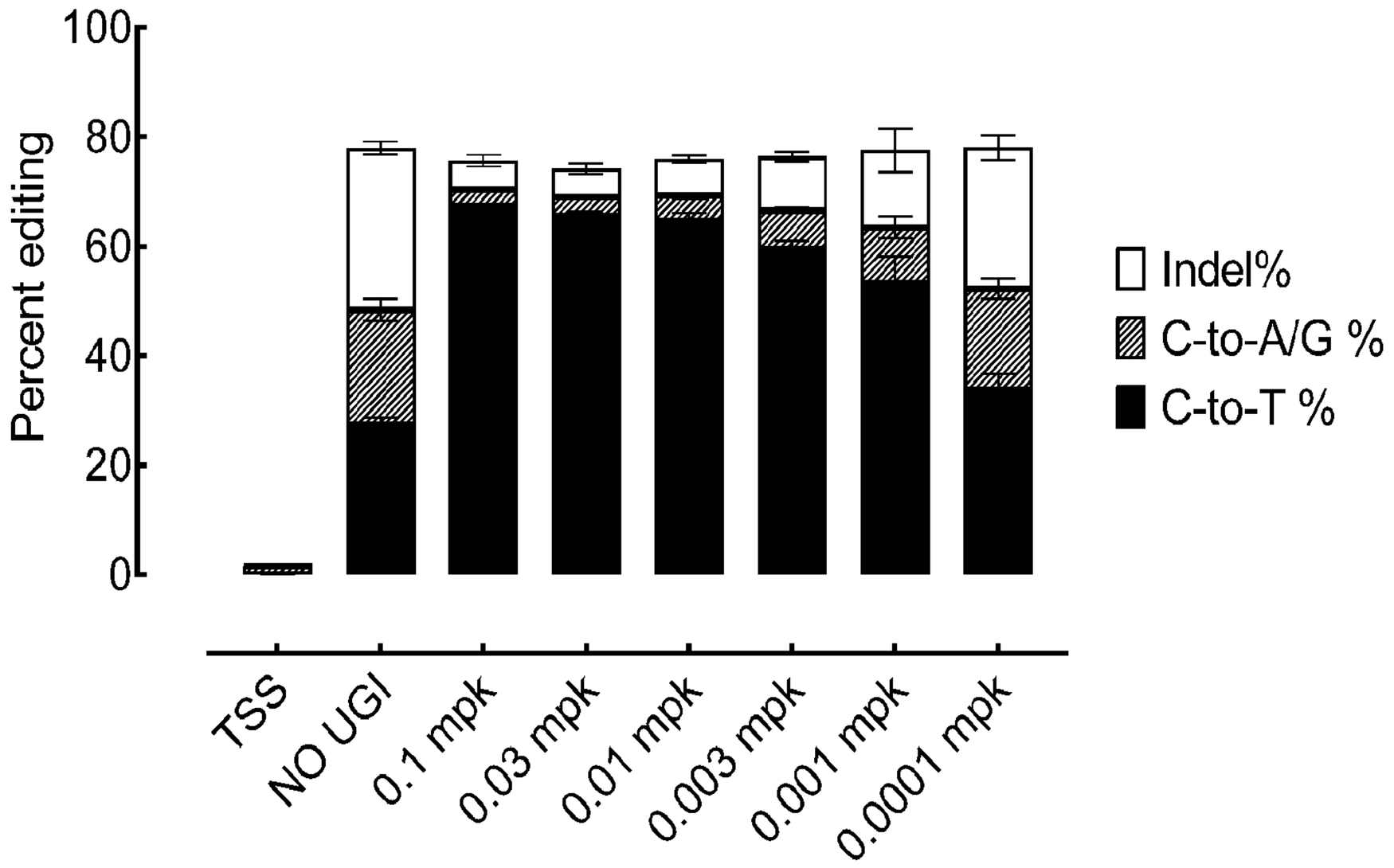

Percent editing
100
80
60
40
20
0
Indel%
C-to-A/G %
C-to-T %
TSS
NO UGI
0.1 mpk
0.03 mpk
0.01 mpk
0.003 mpk
0.001 mpk
0.0001 mpk

Fig. 60

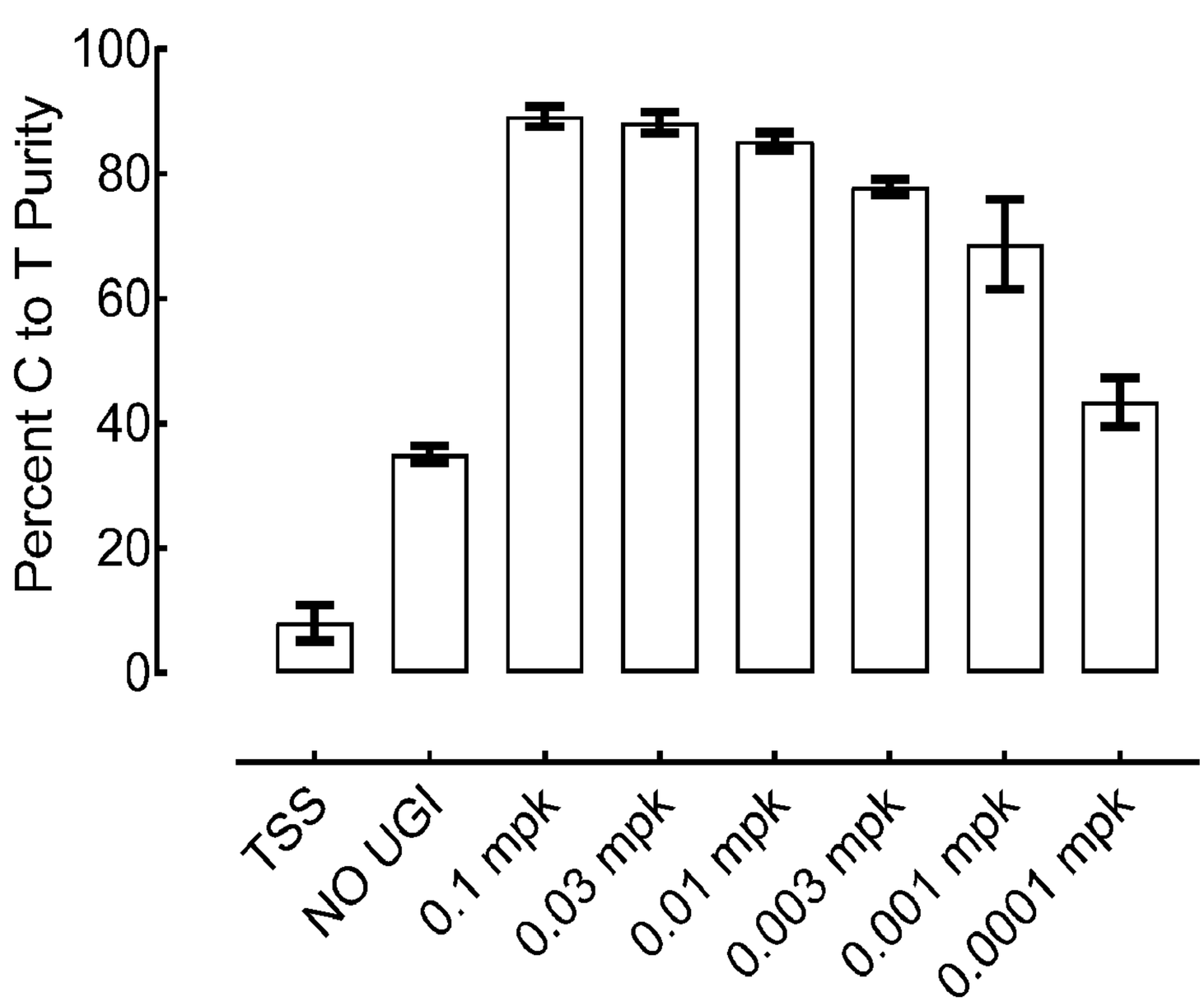

Percent C to T Purity
100
80
60
40
20
0
TSS
NO UGI
0.1 mpk
0.03 mpk
0.01 mpk
0.003 mpk
0.001 mpk
0.0001 mpk

Fig. 61

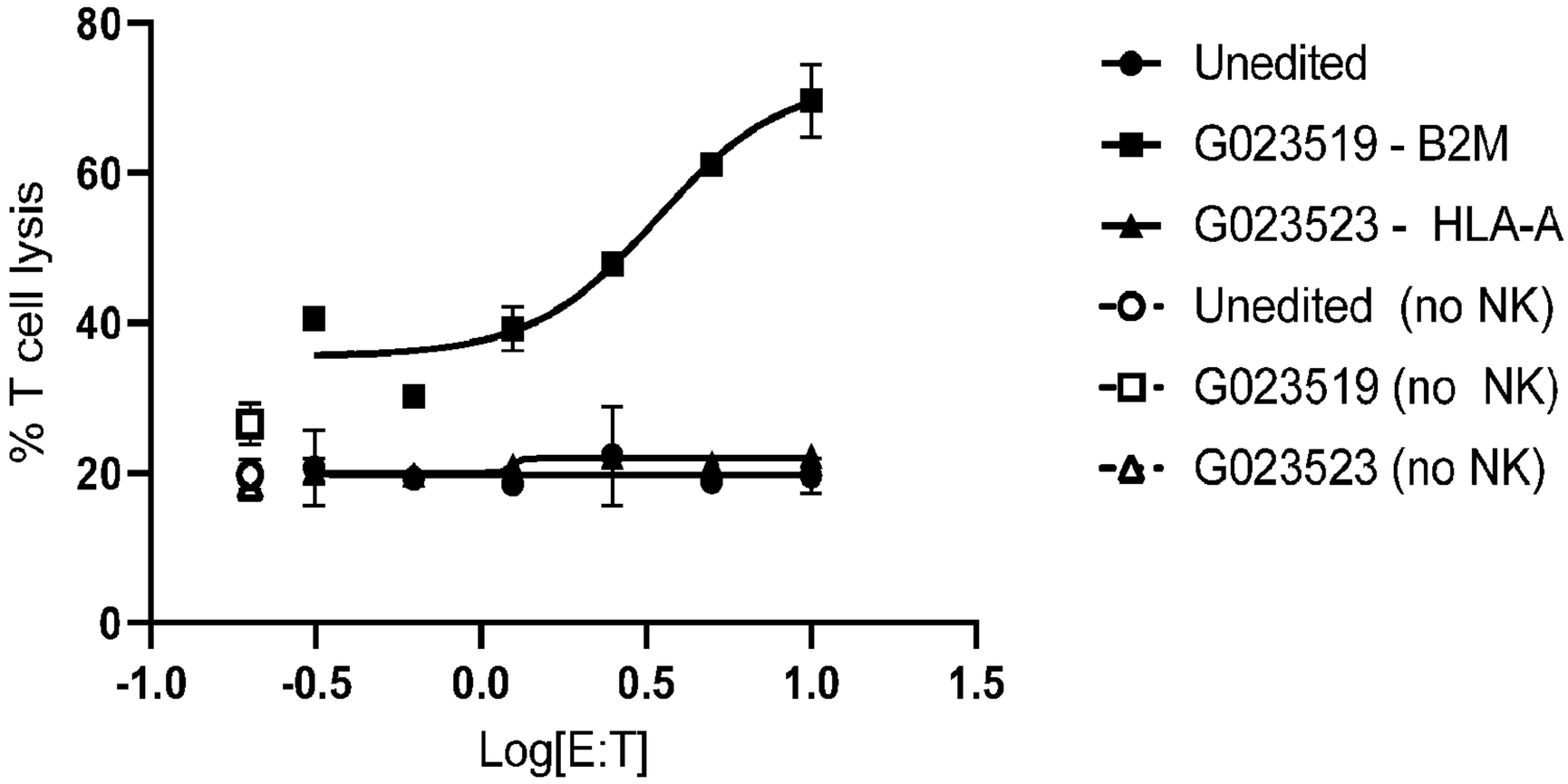

80
60
40
20
0
% T cell lysis
-1.0
-0.5
0.0
0.5
1.0
1.5
Log[E:T]
Unedited
G023519 - B2M
G023523 - HLA-A
Unedited (no NK)
G023519 (no NK)
G023523 (no NK)

Fig. 62

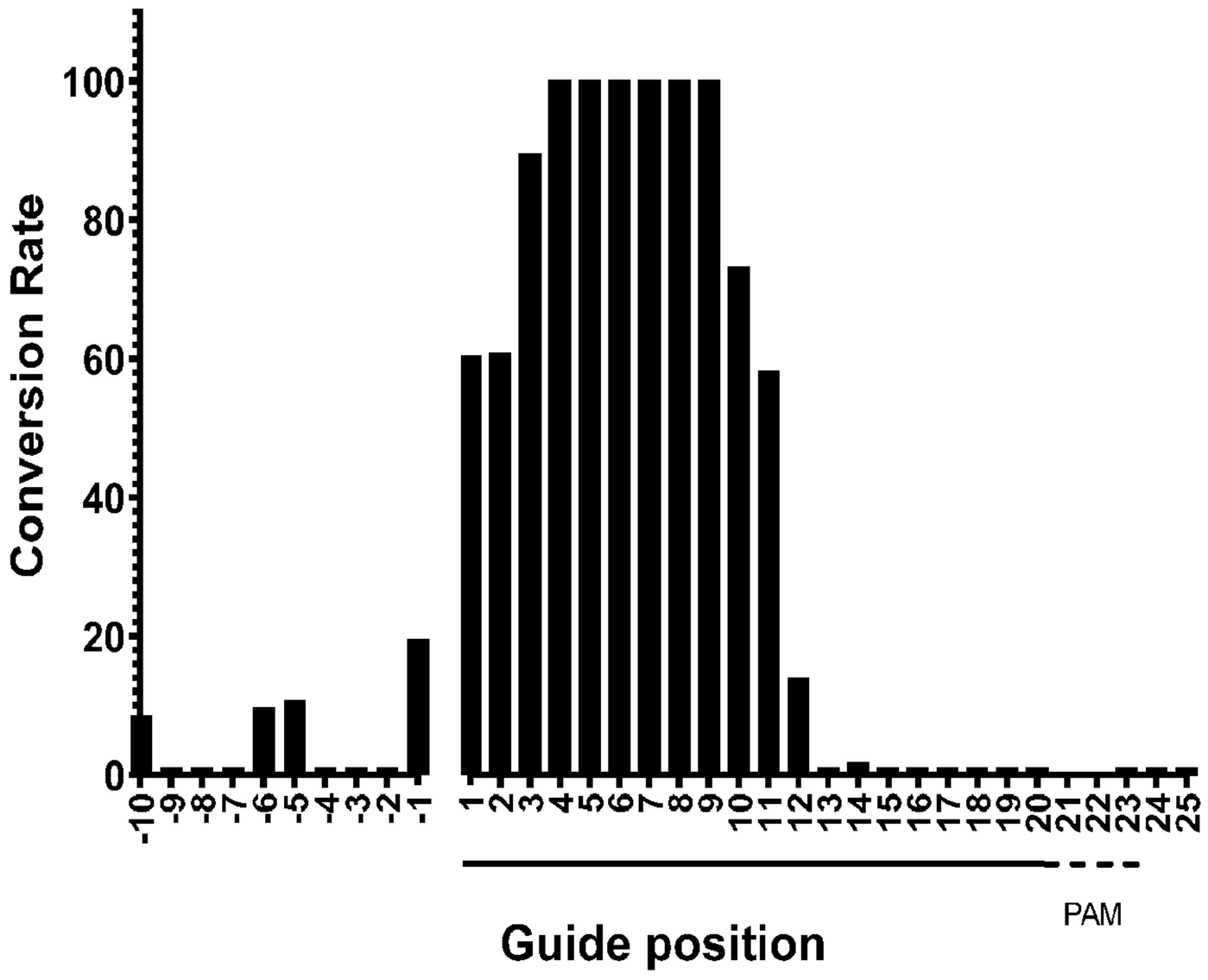

Conversion Rate
100
80
60
40
20
0
-10 -9 -8 -7 -6 -5 -4 -3 -2 -1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
PAM
Guide position

Fig. 63

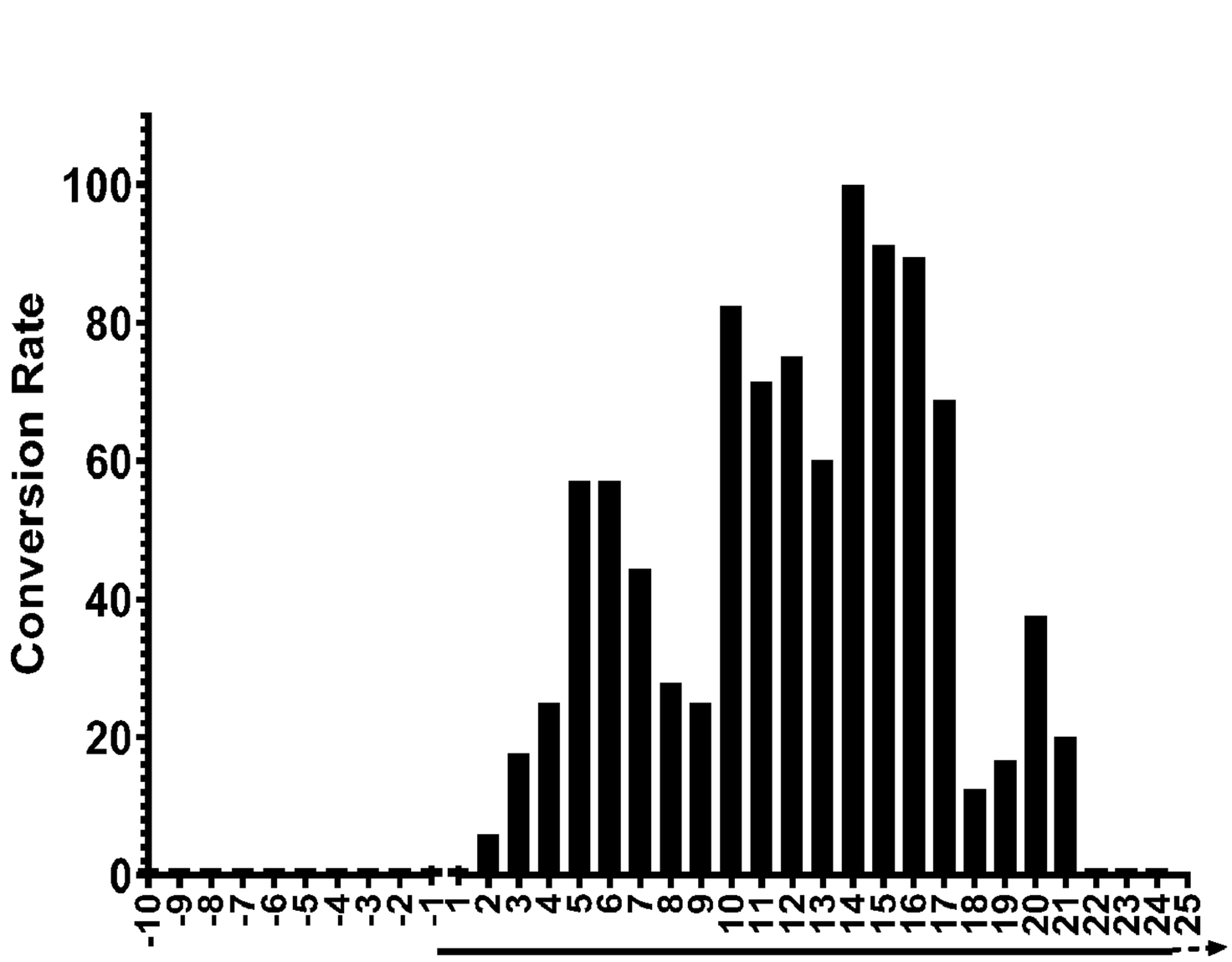

Conversion Rate
100
80
60
40
20
0
-10 -9 -8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Guide Position

Fig. 64

SEQUENCE LISTING

<110> INTELLIA THERAPEUTICS, INC.

<120> POLYNUCLEOTIDES, COMPOSITIONS, AND METHODS FOR GENOME EDITING INVOLVING DEAMINATION

<130> 01155-0016-00PCT

<150> US 63/124,060
<151> 2020-12-11

<150> US 63/130,104
<151> 2020-12-23

<150> US 63/165,636
<151> 2020-03-24

<150> US 63/275,424
<151> 2021-11-03

<160> 1163

<170> PatentIn version 3.5

<210> 1
<211> 5342
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc gagacccccg 660

gcaccuccga guccgccacc cccgaguccg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

uggugggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucguguggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac ggcggcggcu 4800

cccccaagaa gaagcggaag gugugacuag caccagccuc aagaacaccc gaauggaguc 4860

ucuaagcuac auaauaccaa cuuacacuuu acaaaauguu gucccccaaa auguagccau 4920

ucguaucugc uccuaauaaa aagaaaguuu cuucacauuc ucucgagaaa aaaaaaaaau 4980

ggaaaaaaaa aaaacggaaa aaaaaaaaag guaaaaaaaa aaaauauaaa aaaaaaaaac 5040

auaaaaaaaa aaaacgaaaa aaaaaaaacg uaaaaaaaaa aaacucaaaa aaaaaaaaga 5100

uaaaaaaaaa aaaccuaaaa aaaaaaaaug uaaaaaaaaa aaagggaaaa aaaaaaaacg 5160

caaaaaaaaa aaacacaaaa aaaaaaaaug caaaaaaaaa aaaucgaaaa aaaaaaaauc 5220

uaaaaaaaaa aaacgaaaaa aaaaaaaccc aaaaaaaaaa aagacaaaaa aaaaaaauag 5280

aaaaaaaaaa aaguuaaaaa aaaaaaacug aaaaaaaaaa aauuuaaaaa aaaaaaaucu 5340

ag 5342

<210> 2
<211> 4782
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 2
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacucc 600

ggcuccgaga cccccggcac cuccgagucc gccacccccg aguccgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggauccccu acuacguggg cccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau cacccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccuggggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacggcg gcggcucccc caagaagaag cggaaggugu ga 4782

<210> 3
<211> 1593
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 3

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr

```
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser Glu Thr Pro Gly Thr Ser
        195                 200                 205

Glu Ser Ala Thr Pro Glu Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220
```

```
Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
```

```
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575

Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620
```

```
Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
    770                 775                 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
```

```
            820                 825                 830

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
                965                 970                 975

Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020
```

```
Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
```

```
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395
```

```
Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500

Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1535                1540                1545

Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560

Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575

Leu Gly  Gly Asp Gly Gly Gly  Ser Pro Lys Lys Lys  Arg Lys Val
```

1580 1585 1590

<210> 4
<211> 5399
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 4
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc gagacccccg 660

gcaccuccga guccgccacc cccgaguccg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

uggugggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucguguggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac ggcggcggcu 4800

cccccaagaa gaagcggaag guguccgagu ccgccacccc cgaguccgug uccggcuggc 4860

ggcuguucaa gaagaucucc ugacuagcac cagccucaag aacacccgaa uggagucucu 4920

aagcuacaua auaccaacuu acacuuuaca aaauguuguc ccccaaaaug uagccauucg 4980

uaucugcucc uaauaaaaag aaaguuucuu cacauucucu cgagaaaaaa aaaaaaugga 5040

aaaaaaaaaa acggaaaaaa aaaaaaggua aaaaaaaaaa auauaaaaaa aaaaaacaua 5100

aaaaaaaaaa acgaaaaaaa aaaaacguaa aaaaaaaaaa cucaaaaaaa aaaaagauaa 5160

aaaaaaaaaa ccuaaaaaaa aaaaauguaa aaaaaaaaaa gggaaaaaaa aaaaacgcaa 5220

aaaaaaaaaa cacaaaaaaa aaaaaugcaa aaaaaaaaaa ucgaaaaaaa aaaaaucuaa 5280

aaaaaaaaaa cgaaaaaaaa aaaacccaaa aaaaaaaaag acaaaaaaaa aaaauagaaa 5340

aaaaaaaaag uuaaaaaaaa aaaacugaaa aaaaaaaaau uuaaaaaaaa aaaaucuag 5399

<210> 5
<211> 4839
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 5
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacucc 600

ggcuccgaga cccccggcac cuccgagucc gccacccccg aguccgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggauccccu acuacguggg ccccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau cacccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccuggggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacggcg gcggcucccc caagaagaag cggaaggugu ccgaguccgc caccccgag 4800

uccguguccg gcuggcggcu guucaagaag aucuccuga 4839

<210> 6
<211> 1612
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 6

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
20 25 30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
35 40 45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
50 55 60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65 70 75 80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
85 90 95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala

```
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser Glu Thr Pro Gly Thr Ser
        195                 200                 205

Glu Ser Ala Thr Pro Glu Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300
```

```
Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
```

```
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575

Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700
```

```
Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
    770                 775                 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
            820                 825                 830

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
```

```
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
                965                 970                 975

Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095
```

```
Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
```

```
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470
```

```
Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485


Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500


Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515


Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530


Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1535                1540                1545


Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560


Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575


Leu Gly  Gly Asp Gly Gly Gly  Ser Pro Lys Lys Lys  Arg Lys Val
    1580                1585                1590


Ser Glu  Ser Ala Thr Pro Glu  Ser Val Ser Gly Trp  Arg Leu Phe
    1595                1600                1605


Lys Lys  Ile Ser
    1610


<210>  7

<400>  7
000

<210>  8
<211>  4140
<212>  RNA
<213>  Artificial Sequence
```

<220>
<223> Synthetic

<400> 8
auggacaaga aguacagcau cggacuggac aucggaacaa acagcgucgg augggcaguc 60

aucacagacg aauacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120

cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180

gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc 240

uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga 300

cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360

aacaucgucg acgaagucgc auaccacgaa aaguacccga caaucuacca ccugagaaag 420

aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480

augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540

gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600

aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660

agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720

cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa 780

gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840

cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900

cugcugagcg acauccugag agucaacaca gaaaucacaa aggcaccgcu gagcgcaagc 960

augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020

cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca 1080

ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucaucaa gccgauccug 1140

gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200

aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260

gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc 1320

gaaaagaucc ugacauucag aaucccguac uacgucggac cgcuggcaag aggaaacagc 1380

agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440

gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500

aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560

uacaacgaac ugacaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620

agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680

gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc 1740

agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800

aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860

cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920

caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980

agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040

gacuuccuga agagcgacgg auucgcaaac agaaacuuca ugcagcugau ccacgacgac 2100

agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160

cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220

gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280

aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340

augaagagaa ucgaagaagg aaucaaggaa cugggaagcc agauccugaa ggaacacccg 2400

gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460

gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520

aucgucccgc agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580

gacaagaaca gaggaaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640

aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700

acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760

cuggucgaaa caagacagau cacaaagcac gucgcacaga uccuggacag cagaaugaac 2820

acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880

aagcugguca gcgacuucag aaaggacuuc caguucuaca aggucagaga aaucaacaac 2940

uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000

uacccgaagc uggaaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag 3060

augaucgcaa agagcgaaca ggaaaucgga aaggcaacag caaaguacuu cuucuacagc 3120

aacaucauga acuucuucaa gacagaaauc acacuggcaa acggagaaau cagaaagaga 3180

ccgcugaucg aaacaaacgg agaaacagga gaaaucgucu gggacaaggg aagagacuuc 3240

gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc 3300

cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360

gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420

uacagcgucc uggucgucgc aaaggucgaa aagggaaaga gcaagaagcu gaagagcguc 3480

aaggaacugc ugggaaucac aaucauggaa agaagcagcu ucgaaaagaa cccgaucgac 3540

uuccuggaag caaagggaua caaggaaguc aagaaggacc ugaucaucaa gcugccgaag 3600

uacagccugu ucgaacugga aaacggaaga aagagaaugc uggcaagcgc aggagaacug 3660

cagaagggaa acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720

cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780

cagcacaagc acuaccugga cgaaaucauc gaacagauca gcgaauucag caagagaguc 3840

auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900

ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960

ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaaag 4020

gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaauc 4080

gaccugagcc agcugggagg agacggagga ggaagcccga agaagaagag aaaggucuag 4140

<210> 9
<211> 1379
<212> PRT
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 9

```
Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1               5                  10                  15


Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
            20                  25                  30


Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
        35                  40                  45


Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
    50                  55                  60


Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65                  70                  75                  80


Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                85                  90                  95


Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110


His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125


His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140


Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160


Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175


Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190
```

```
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380
```

```
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590
```

```
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780
```

```
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990
```

```
Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005


Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020


Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035


Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050


Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065


Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080


Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
    1085                1090                1095


Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110


Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
    1115                1120                1125


Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140


Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155


Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170
```

```
Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185

Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215

Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245

Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260

His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275

Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290

Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305

Ile Ile  His Leu Phe Thr Leu  Thr Asn Leu Gly Ala  Pro Ala Ala
    1310                1315                1320

Phe Lys  Tyr Phe Asp Thr Thr  Ile Asp Arg Lys Arg  Tyr Thr Ser
    1325                1330                1335

Thr Lys  Glu Val Leu Asp Ala  Thr Leu Ile His Gln  Ser Ile Thr
    1340                1345                1350

Gly Leu  Tyr Glu Thr Arg Ile  Asp Leu Ser Gln Leu  Gly Gly Asp
    1355                1360                1365
```

Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
    1370   1375

<210> 10

<400> 10
000

<210> 11
<211> 4140
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 11
auggacaaga aguacuccau cggccuggac aucggcacca acuccguggg cugggccgug 60

aucaccgacg aguacaaggu gcccuccaag aaguucaagg ugcugggcaa caccgaccgg 120

cacuccauca agaagaaccu gaucggcgcc cugcuguucg acuccggcga gaccgccgag 180

gccacccggc ugaagcggac cgcccggcgg cgguacaccc ggcggaagaa ccggaucugc 240

uaccugcagg agaucuucuc caacgagaug gccaaggugg acgacuccuu cuuccaccgg 300

cuggaggagu ccuuccuggu ggaggaggac aagaagcacg agcggcaccc caucuucggc 360

aacaucgugg acgagguggc cuaccacgag aaguacccca ccaucuacca ccugcggaag 420

aagcuggugg acuccaccga caaggccgac cugcggcuga ucuaccuggc ccuggcccac 480

augaucaagu uccggggcca cuuccugauc gagggcgacc ugaaccccga caacuccgac 540

guggacaagc uguucaucca gcuggugcag accuacaacc agcuguucga ggagaacccc 600

aucaacgccu ccggcgugga cgccaaggcc auccuguccg cccggcuguc caagucccgg 660

cggcuggaga accugaucgc ccagcugccc ggcgagaaga agaacggccu guucggcaac 720

cugaucgccc ugucccuggg ccugaccccc aacuucaagu ccaacuucga ccuggccgag 780

gacgccaagc ugcagcuguc caaggacacc uacgacgacg accuggacaa ccugcuggcc 840

cagaucggcg accaguacgc cgaccuguuc cuggccgcca agaaccuguc cgacgccauc 900

cugcuguccg acauccugcg ggugaacacc gagaucacca aggccccccu guccgccucc 960

augaucaagc gguacgacga gcaccaccag gaccugaccc ugcugaaggc ccuggugcgg 1020

cagcagcugc ccgagaagua caaggagauc uucuucgacc aguccaagaa cggcuacgcc 1080

ggcuacaucg acggcggcgc cucccaggag gaguucuaca aguucaucaa gcccauccug 1140

gagaagaugg acggcaccga ggagcugcug gugaagcuga accgggagga ccugcugcgg 1200

aagcagcgga ccuucgacaa cggcuccauc ccccaccaga uccaccuggg cgagcugcac 1260

gccauccugc ggcggcagga ggacuucuac cccuuccuga aggacaaccg ggagaagauc 1320

gagaagaucc ugaccuuccg gaucccccuac uacgugggcc cccuggcccg gggcaacucc 1380

cgguucgccu ggaugacccg gaaguccgag gagaccauca ccccccuggaa cuucgaggag 1440

gugguggaca agggcgccuc cgcccagucc uucaucgagc ggaugaccaa cuucgacaag 1500

aaccugccca acgagaaggu gcugcccaag cacucccugc uguacgagua cuucaccgug 1560

uacaacgagc ugaccaaggu gaaguacgug accgagggca ugcggaagcc cgccuuccug 1620

uccggcgagc agaagaaggc caucguggac cugcuguuca agaccaaccg gaaggugacc 1680

gugaagcagc ugaaggagga cuacuucaag aagaucgagu gcuucgacuc cguggagauc 1740

uccggcgugg aggaccgguu caacgccucc cugggcaccu accacgaccu gcugaagauc 1800

aucaaggaca aggacuuccu ggacaacgag gagaacgagg acauccugga ggacaucgug 1860

cugacccuga cccuguucga ggaccgggag augaucgagg agcggcugaa gaccuacgcc 1920

caccuguucg acgacaaggu gaugaagcag cugaagcggc ggcgguacac cggcuggggc 1980

cggcuguccc ggaagcugau caacggcauc cgggacaagc aguccggcaa gaccauccug 2040

gacuuccuga aguccgacgg cuucgccaac cggaacuuca ugcagcugau ccacgacgac 2100

ucccugaccu ucaaggagga cauccagaag gcccaggugu ccggccaggg cgacucccug 2160

cacgagcaca ucgccaaccu ggccggcucc cccgccauca agaagggcau ccugcagacc 2220

gugaaggugg uggacgagcu ggugaaggug augggccggc acaagcccga gaacaucgug 2280

aucgagaugg cccgggagaa ccagaccacc cagaagggcc agaagaacuc ccgggagcgg 2340

augaagcgga ucgaggaggg caucaaggag cugggcuccc agauccugaa ggagcacccc 2400

guggagaaca cccagcugca gaacgagaag cuguaccugu acuaccugca gaacggccgg 2460

gacauguacg uggaccagga gcuggacauc aaccggcugu ccgacuacga cguggaccac 2520

aucgugcccc aguccuuccu gaaggacgac uccaucgaca acaaggugcu gacccggucc 2580

gacaagaacc ggggcaaguc cgacaacgug cccuccgagg agguggugaa gaagaugaag 2640

aacuacuggc ggcagcugcu gaacgccaag cugaucaccc agcggaaguu cgacaaccug 2700

accaaggccg agcggggcgg ccuguccgag cuggacaagg ccggcuucau caagcggcag 2760

cugguggaga cccggcagau caccaagcac guggcccaga uccuggacuc ccggaugaac 2820

accaaguacg acgagaacga caagcugauc cgggagguga aggugaucac ccugaagucc 2880

aagcuggugu ccgacuuccg gaaggacuuc caguucuaca aggugcggga gaucaacaac 2940

uaccaccacg cccacgacgc cuaccugaac gccguggugg gcaccgcccu gaucaagaag 3000

uaccccaagc uggaguccga guucguguac ggcgacuaca agguguacga cgugcggaag 3060

augaucgcca aguccgagca ggagaucggc aaggccaccg ccaaguacuu cuucuacucc 3120

aacaucauga acuucuucaa gaccgagauc acccuggcca acggcgagau ccggaagcgg 3180

ccccugaucg agaccaacgg cgagaccggc gagaucgugu gggacaaggg ccgggacuuc 3240

gccaccgugc ggaaggugcu guccaugccc caggugaaca ucgugaagaa gaccgaggug 3300

cagaccggcg gcuucuccaa ggaguccauc cugcccaagc ggaacuccga caagcugauc 3360

gcccggaaga aggacuggga ccccaagaag uacggcggcu ucgacucccc caccguggcc 3420

uacuccgugc uggugguggc caagguggag aagggcaagu ccaagaagcu gaaguccgug 3480

aaggagcugc ugggcaucac caucauggag cgguccuccu ucgagaagaa ccccaucgac 3540

uuccuggagg ccaagggcua caaggaggug aagaaggacc ugaucaucaa gcugcccaag 3600

uacucccugu ucgagcugga gaacggccgg aagcggaugc uggccuccgc cggcgagcug 3660

cagaagggca acgagcuggc ccugcccucc aaguacguga acuuccugua ccuggccucc 3720

cacuacgaga agcugaaggg cucccccgag gacaacgagc agaagcagcu guucguggag 3780

cagcacaagc acuaccugga cgagaucauc gagcagaucu ccgaguucuc caagcgggug 3840

auccuggccg acgccaaccu ggacaaggug cuguccgccu acaacaagca ccgggacaag 3900

cccauccggg agcaggccga gaacaucauc caccuguuca cccugaccaa ccugggcgcc 3960

cccgccgccu ucaaguacuu cgacaccacc aucgaccgga agcgguacac cuccaccaag 4020

gaggugcugg acgccacccu gauccaccag uccaucaccg gccuguacga gacccggauc 4080

gaccuguccc agcugggcgg cgacggcggc ggcuccccca agaagaagcg gaagguguga 4140

<210> 12
<211> 1379
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 12

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

```
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
```

```
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525
```

```
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
```

```
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925
```

```
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990

Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005

Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020

Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035

Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050

Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065

Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080

Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
    1085                1090                1095

Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110

Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
```

```
    1115                1120                1125

Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140

Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155

Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170

Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185

Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215

Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245

Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260

His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275

Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290

Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305
```

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
    1310   1315   1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
    1325   1330   1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
    1340   1345   1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
    1355   1360   1365

Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
    1370   1375

<210> 13
<211> 5548
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 13
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc augagcagcg 120

aaacaggacc ggucgcaguc gacccgacac ugagaagaag aaucgaaccg cacgaauucg 180

aagucuucuu cgacccgaga gaacugagaa aggaaacaug ccugcuguac gaaaucaacu 240

ggggaggaag acacagcauc uggagacaca caagccagaa cacaaacaag cacgucgaag 300

ucaacuucau cgaaaaguuc acaacagaaa gauacuucug cccgaacaca agaugcagca 360

ucacaugguu ccugagcugg agcccgugcg gagaaugcag cagagcaauc acagaauucc 420

ugagcagaua cccgcacguc acacuguuca ucuacaucgc aagacuguac caccacgcag 480

acccgagaaa cagacaggga cugagagacc ugaucagcag cggagucaca auccagauca 540

ugacagaaca ggaaagcgga uacugcugga gaaacuucgu caacuacagc ccgagcaacg 600

aagcacacug gccgagauac ccgcaccugu gggucagacu guacguccug gaacuguacu 660

gcaucauccu gggacugccg ccgugccuga acauccugag aagaaagcag ccgcagcuga 720

cauucuucac aaucgcacug cagagcugcc acuaccagag acugccgccg cacauccugu 780

gggcaacagg acugaagagc ggaagcgaaa caccgggaac aagcgaaagc gcaacaccgg 840

aaagcgacaa gaaguacagc aucggacugg ccaucggaac aaacagcguc ggaugggcag 900

ucaucacaga cgaauacaag gucccgagca agaaguucaa gguccuggga aacacagaca 960

gacacagcau caagaagaac cugaucggag cacugcuguu cgacagcgga gaaacagcag 1020

aagcaacaag acugaagaga acagcaagaa gaagauacac aagaagaaag aacagaaucu 1080

gcuaccugca ggaaaucuuc agcaacgaaa uggcaaaggu cgacgacagc uucuuccaca 1140

gacuggaaga aagcuuccug gucgaagaag acaagaagca cgaaagacac ccgaucuucg 1200

gaaacaucgu cgacgaaguc gcauaccacg aaaaguaccc gacaaucuac caccugagaa 1260

agaagcuggu cgacagcaca gacaaggcag accugagacu gaucuaccug gcacuggcac 1320

acaugaucaa guucagagga cacuuccuga ucgaaggaga ccugaacccg gacaacagcg 1380

acgucgacaa gcuguucauc cagcuggucc agacauacaa ccagcuguuc gaagaaaacc 1440

cgaucaacgc aagcggaguc gacgcaaagg caauccugag cgcaagacug agcaagagca 1500

gaagacugga aaaccugauc gcacagcugc cgggagaaaa gaagaacgga cuguucggaa 1560

accugaucgc acugagccug ggacugacac cgaacuucaa gagcaacuuc gaccuggcag 1620

aagacgcaaa gcugcagcug agcaaggaca cauacgacga cgaccuggac aaccugcugg 1680

cacagaucgg agaccaguac gcagaccugu uccuggcagc aaagaaccug agcgacgcaa 1740

uccugcugag cgacauccug agagucaaca cagaaaucac aaaggcaccg cugagcgcaa 1800

gcaugaucaa gagauacgac gaacaccacc aggaccugac acugcugaag gcacugguca 1860

gacagcagcu gccggaaaag uacaaggaaa ucuucuucga ccagagcaag aacggauacg 1920

caggauacau cgacggagga gcaagccagg aagaauucua caaguucauc aagccgaucc 1980

uggaaaagau ggacggaaca gaagaacugc uggucaagcu gaacagagaa gaccugcuga 2040

gaaagcagag aacauucgac aacggaagca ucccgcacca gauccaccug ggagaacugc 2100

acgcaauccu gagaagacag gaagacuucu acccguuccu gaaggacaac agagaaaaga 2160

ucgaaaagau ccugacauuc agaaucccgu acuacgucgg accgcuggca agaggaaaca 2220

gcagauucgc auggaugaca agaaagagcg aagaaacaau cacaccgugg aacuucgaag 2280

aagucgucga caagggagca agcgcacaga gcuucaucga aagaaugaca aacuucgaca 2340

agaaccugcc gaacgaaaag guccugccga agcacagccu gcuguacgaa uacuucacag 2400

ucuacaacga acugacaaag gucaaguacg ucacagaagg aaugagaaag ccggcauucc 2460

ugagcggaga acagaagaag gcaaucgucg accugcuguu caagacaaac agaaagguca 2520

cagucaagca gcugaaggaa gacuacuuca agaagaucga augcuucgac agcgucgaaa 2580

ucagcggagu cgaagacaga uucaacgcaa gccuggggaac auaccacgac cugcugaaga 2640

ucaucaagga caaggacuuc cuggacaacg aagaaaacga agacauccug gaagacaucg 2700

uccugacacu gacacuguuc gaagacagag aaaugaucga agaaagacug aagacauacg 2760

cacaccuguu cgacgacaag gucaugaagc agcugaagag aagaagauac acaggauggg 2820

gaagacugag cagaaagcug aucaacggaa ucagagacaa gcagagcgga aagacaaucc 2880

uggacuuccu gaagagcgac ggauucgcaa acagaaacuu caugcagcug auccacgacg 2940

acagccugac auucaaggaa gacauccaga aggcacaggu cagcggacag ggagacagcc 3000

ugcacgaaca caucgcaaac cuggcaggaa gcccggcaau caagaaggga auccugcaga 3060

cagucaaggu cgucgacgaa cuggucaagg ucaugggaag acacaagccg gaaaacaucg 3120

ucaucgaaau ggcaagagaa aaccagacaa cacagaaggg acagaagaac agcagagaaa 3180

gaaugaagag aaucgaagaa ggaaucaagg aacugggaag ccagauccug aaggaacacc 3240

cggucgaaaa cacacagcug cagaacgaaa agcuguaccu guacuaccug cagaacggaa 3300

gagacaugua cgucgaccag gaacuggaca ucaacagacu gagcgacuac gacgucgacc 3360

acaucguccc gcagagcuuc cugaaggacg acagcaucga caacaagguc cugacaagaa 3420

gcgacaagaa cagaggaaag agcgacaacg ucccgagcga agaagucguc aagaagauga 3480

agaacuacug gagacagcug cugaacgcaa agcugaucac acagagaaag uucgacaacc 3540

ugacaaaggc agagagagga ggacugagcg aacuggacaa ggcaggauuc aucaagagac 3600

agcuggucga aacaagacag aucacaaagc acgucgcaca gauccuggac agcagaauga 3660

acacaaagua cgacgaaaac gacaagcuga ucagagaagu caaggucauc acacugaaga 3720

gcaagcuggu cagcgacuuc agaaaggacu uccaguucua caaggucaga gaaaucaaca 3780

acuaccacca cgcacacgac gcauaccuga acgcagucgu cggaacagca cugaucaaga 3840

aguacccgaa gcuggaaagc gaauucgucu acggagacua caaggucuac gacgucagaa 3900

agaugaucgc aaagagcgaa caggaaaucg gaaaggcaac agcaaaguac uucuucuaca 3960

gcaacaucau gaacuucuuc aagacagaaa ucacacuggc aaacggagaa aucagaaaga 4020

gaccgcugau cgaaacaaac ggagaaacag gagaaaucgu cugggacaag ggaagagacu 4080

ucgcaacagu cagaaagguc cugagcaugc cgcaggucaa caucgucaag aagacagaag 4140

uccagacagg aggauucagc aaggaaagca uccugccgaa gagaaacagc gacaagcuga 4200

ucgcaagaaa gaaggacugg gacccgaaga aguacggagg auucgacagc ccgacagucg 4260

cauacagcgu ccuggucguc gcaaaggucg aaaagggaaa gagcaagaag cugaagagcg 4320

ucaaggaacu gcugggaauc acaaucaugg aaagaagcag cuucgaaaag aacccgaucg 4380

acuuccugga agcaaaggga uacaaggaag ucaagaagga ccugaucauc aagcugccga 4440

aguacagccu guucgaacug gaaaacggaa gaaagagaau gcuggcaagc gcaggagaac 4500

ugcagaaggg aaacgaacug gcacugccga gcaaguacgu caacuuccug uaccuggcaa 4560

gccacuacga aaagcugaag ggaagcccgg aagacaacga acagaagcag cuguucgucg 4620

aacagcacaa gcacuaccug gacgaaauca ucgaacagau cagcgaauuc agcaagagag 4680

ucauccuggc agacgcaaac cuggacaagg uccugagcgc auacaacaag cacagagaca 4740

agccgaucag agaacaggca gaaaacauca uccaccuguu cacacugaca aaccugggag 4800

caccggcagc auucaaguac uucgacacaa caaucgacag aaagagauac acaagcacaa 4860

aggaaguccu ggacgcaaca cugauccacc agagcaucac aggacuguac gaaacaagaa 4920

ucgaucugag ccagcuggga ggagacagcg gaggaagcac aaaccugagc gacaucaucg 4980

aaaaggaaac aggaaagcag cuggucaucc aggaaagcau ccugaugcug ccggaagaag 5040

ucgaagaagu caucggaaac aagccggaaa gcgacauccu gguccacaca gcauacgacg 5100

aaagcacaga cgaaaacguc augcugcuga caagcgacgc accggaauac aagccguggg 5160

cacuggucau ccaggacagc aacggagaaa acaagaucaa gaugcugagc ggaggaagcc 5220

cgaagaagaa gagaaagguc uaauagucua gacaucacau uuaaaagcau cucagccuac 5280

caugagaaua agagaaagaa aaugaagauc aauagcuuau ucaucucuuu uucuuuuucg 5340

uugguguaaa gccaacaccc ugucuaaaaa acauaaauuu cuuuaaucau uuugccucuu 5400

uucucugugc uucaauuaau aaaaaaugga aagaaccucg agaaaaaaaa aaaaaaaaaa 5460

aaaaaaaaaa aagcgaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaccgaa aaaaaaaaaa 5520

aaaaaaaaaa aaaaaaaaaa aaaaaaau 5548

<210> 14
<211> 5130
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 14
augagcagcg aaacaggacc ggucgcaguc gacccgacac ugagaagaag aaucgaaccg 60

cacgaauucg aagucuucuu cgacccgaga gaacugagaa aggaaacaug ccugcuguac 120

gaaaucaacu ggggaggaag acacagcauc uggagacaca caagccagaa cacaaacaag 180

cacgucgaag ucaacuucau cgaaaaguuc acaacagaaa gauacuucug cccgaacaca 240

agaugcagca ucacaugguu ccugagcugg agcccgugcg gagaaugcag cagagcaauc 300

acagaauucc ugagcagaua cccgcacguc acacuguuca ucuacaucgc aagacuguac 360

caccacgcag acccgagaaa cagacaggga cugagagacc ugaucagcag cggagucaca 420

auccagauca ugacagaaca ggaaagcgga uacugcugga gaaacuucgu caacuacagc 480

ccgagcaacg aagcacacug gccgagauac ccgcaccugu gggucagacu guacguccug 540

gaacuguacu gcaucauccu gggacugccg ccgugccuga acauccugag aagaaagcag 600

ccgcagcuga cauucuucac aaucgcacug cagagcugcc acuaccagag acugccgccg 660

cacauccugu gggcaacagg acugaagagc ggaagcgaaa caccgggaac aagcgaaagc 720

gcaacaccgg aaagcgacaa gaaguacagc aucggacugg ccaucggaac aaacagcguc 780

ggaugggcag ucaucacaga cgaauacaag gucccgagca agaaguucaa gguccuggga 840

aacacagaca gacacagcau caagaagaac cugaucggag cacugcuguu cgacagcgga 900

gaaacagcag aagcaacaag acugaagaga acagcaagaa gaagauacac aagaagaaag 960

aacagaaucu gcuaccugca ggaaaucuuc agcaacgaaa uggcaaaggu cgacgacagc 1020

uucuuccaca gacuggaaga aagcuuccug gucgaagaag acaagaagca cgaaagacac 1080

ccgaucuucg gaaacaucgu cgacgaaguc gcauaccacg aaaaguaccc gacaaucuac 1140

caccugagaa agaagcuggu cgacagcaca gacaaggcag accugagacu gaucuaccug 1200

gcacuggcac acaugaucaa guucagagga cacuuccuga ucgaaggaga ccugaacccg 1260

gacaacagcg acgucgacaa gcuguucauc cagcuggucc agacauacaa ccagcuguuc 1320

gaagaaaacc cgaucaacgc aagcggaguc gacgcaaagg caauccugag cgcaagacug 1380

agcaagagca gaagacugga aaaccugauc gcacagcugc cgggagaaaa gaagaacgga 1440

cuguucggaa accugaucgc acugagccug ggacugacac cgaacuucaa gagcaacuuc 1500

gaccuggcag aagacgcaaa gcugcagcug agcaaggaca cauacgacga cgaccuggac 1560

aaccugcugg cacagaucgg agaccaguac gcagaccugu uccuggcagc aaagaaccug 1620

agcgacgcaa uccugcugag cgacauccug agagucaaca cagaaaucac aaaggcaccg 1680

cugagcgcaa gcaugaucaa gagauacgac gaacaccacc aggaccugac acugcugaag 1740

gcacugguca gacagcagcu gccggaaaag uacaaggaaa ucuucuucga ccagagcaag 1800

aacggauacg caggauacau cgacggagga gcaagccagg aagaauucua caaguucauc 1860

aagccgaucc uggaaaagau ggacggaaca gaagaacugc uggucaagcu gaacagagaa 1920

gaccugcuga gaaagcagag aacauucgac aacggaagca ucccgcacca gauccaccug 1980

ggagaacugc acgcaauccu gagaagacag gaagacuucu acccguuccu gaaggacaac 2040

agagaaaaga ucgaaaagau ccugacauuc agaaucccgu acuacgucgg accgcuggca 2100

agaggaaaca gcagauucgc auggaugaca agaaagagcg aagaaacaau cacaccgugg 2160

aacuucgaag aagucgucga caagggagca agcgcacaga gcuucaucga aagaaugaca 2220

aacuucgaca agaaccugcc gaacgaaaag guccugccga agcacagccu gcuguacgaa 2280

uacuucacag ucuacaacga acugacaaag gucaaguacg ucacagaagg aaugagaaag 2340

ccggcauucc ugagcggaga acagaagaag gcaaucgucg accugcuguu caagacaaac 2400

agaaagguca cagucaagca gcugaaggaa gacuacuuca agaagaucga augcuucgac 2460

agcgucgaaa ucagcggagu cgaagacaga uucaacgcaa gccugggaac auaccacgac 2520

cugcugaaga ucaucaagga caaggacuuc cuggacaacg aagaaaacga agacauccug 2580

gaagacaucg uccugacacu gacacuguuc gaagacagag aaaugaucga agaaagacug 2640

aagacauacg cacaccuguu cgacgacaag gucaugaagc agcugaagag aagaagauac 2700

acaggauggg gaagacugag cagaaagcug aucaacggaa ucagagacaa gcagagcgga 2760

aagacaaucc uggacuuccu gaagagcgac ggauucgcaa acagaaacuu caugcagcug 2820

auccacgacg acagccugac auucaaggaa gacauccaga aggcacaggu cagcggacag 2880

ggagacagcc ugcacgaaca caucgcaaac cuggcaggaa gcccggcaau caagaaggga 2940

auccugcaga cagucaaggu cgucgacgaa cuggucaagg ucaugggaag acacaagccg 3000

gaaaacaucg ucaucgaaau ggcaagagaa aaccagacaa cacagaaggg acagaagaac 3060

agcagagaaa gaaugaagag aaucgaagaa ggaaucaagg aacugggaag ccagauccug 3120

aaggaacacc cggucgaaaa cacacagcug cagaacgaaa agcuguaccu guacuaccug 3180

cagaacggaa gagacaugua cgucgaccag gaacuggaca ucaacagacu gagcgacuac 3240

gacgucgacc acaucguccc gcagagcuuc cugaaggacg acagcaucga caacaagguc 3300

cugacaagaa gcgacaagaa cagaggaaag agcgacaacg ucccgagcga agaagucguc 3360

aagaagauga agaacuacug gagacagcug cugaacgcaa agcugaucac acagagaaag 3420

uucgacaacc ugacaaaggc agagagagga ggacugagcg aacuggacaa ggcaggauuc 3480

aucaagagac agcuggucga aacaagacag aucacaaagc acgucgcaca gauccuggac 3540

agcagaauga acacaaagua cgacgaaaac gacaagcuga ucagagaagu caaggucauc 3600

acacugaaga gcaagcuggu cagcgacuuc agaaaggacu uccaguucua caaggucaga 3660

gaaaucaaca acuaccacca cgcacacgac gcauaccuga acgcagucgu cggaacagca 3720

cugaucaaga aguacccgaa gcuggaaagc gaauucgucu acggagacua caaggucuac 3780

gacgucagaa agaugaucgc aaagagcgaa caggaaaucg gaaaggcaac agcaaaguac 3840

uucuucuaca gcaacaucau gaacuucuuc aagacagaaa ucacacuggc aaacggagaa 3900

aucagaaaga gaccgcugau cgaaacaaac ggagaaacag gagaaaucgu cugggacaag 3960

ggaagagacu ucgcaacagu cagaaagguc cugagcaugc cgcaggucaa caucgucaag 4020

aagacagaag uccagacagg aggauucagc aaggaaagca uccugccgaa gagaaacagc 4080

gacaagcuga ucgcaagaaa gaaggacugg gacccgaaga aguacggagg auucgacagc 4140

ccgacagucg cauacagcgu ccuggucguc gcaaaggucg aaaagggaaa gagcaagaag 4200

cugaagagcg ucaaggaacu gcugggaauc acaaucaugg aaagaagcag cuucgaaaag 4260

aacccgaucg acuuccugga agcaaaggga uacaaggaag ucaagaagga ccugaucauc 4320

aagcugccga aguacagccu guucgaacug gaaaacggaa gaaagagaau gcuggcaagc 4380

gcaggagaac ugcagaaggg aaacgaacug gcacugccga gcaaguacgu caacuuccug 4440

uaccuggcaa gccacuacga aaagcugaag ggaagcccgg aagacaacga acagaagcag 4500

cuguucgucg aacagcacaa gcacuaccug gacgaaauca ucgaacagau cagcgaauuc 4560

agcaagagag ucauccuggc agacgcaaac cuggacaagg uccugagcgc auacaacaag 4620

cacagagaca agccgaucag agaacaggca gaaaacauca uccaccuguu cacacugaca 4680

aaccugggag caccggcagc auucaaguac uucgacacaa caaucgacag aaagagauac 4740

acaagcacaa aggaaguccu ggacgcaaca cugauccacc agagcaucac aggacuguac 4800

gaaacaagaa ucgaucugag ccagcuggga ggagacagcg gaggaagcac aaaccugagc 4860

gacaucaucg aaaaggaaac aggaaagcag cuggucaucc aggaaagcau ccugaugcug 4920

ccggaagaag ucgaagaagu caucggaaac aagccggaaa gcgacauccu gguccacaca 4980

gcauacgacg aaagcacaga cgaaaacguc augcugcuga caagcgacgc accggaauac 5040

aagccguggg cacuggucau ccaggacagc aacggagaaa acaagaucaa gaugcugagc 5100

ggaggaagcc cgaagaagaa gagaaagguc 5130

<210> 15

<211> 1710
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 15

Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1 5 10 15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
20 25 30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
35 40 45

Ser Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val
50 55 60

Asn Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr
65 70 75 80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
85 90 95

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val Thr Leu
100 105 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg Asn Arg
115 120 125

Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met
130 135 140

Thr Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn Tyr Ser
145 150 155 160

Pro Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp Val Arg

```
                165                 170                 175

Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
            180                 185                 190

Leu Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile
        195                 200                 205

Ala Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp
    210                 215                 220

Ala Thr Gly Leu Lys Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser
225                 230                 235                 240

Ala Thr Pro Glu Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly
                245                 250                 255

Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro
            260                 265                 270

Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys
        275                 280                 285

Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu
    290                 295                 300

Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys
305                 310                 315                 320

Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys
                325                 330                 335

Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu
            340                 345                 350

Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp
        355                 360                 365
```

```
Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys
    370                 375                 380

Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu
385                 390                 395                 400

Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly
                405                 410                 415

Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu
            420                 425                 430

Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser
        435                 440                 445

Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg
    450                 455                 460

Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly
465                 470                 475                 480

Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe
                485                 490                 495

Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys
            500                 505                 510

Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp
        515                 520                 525

Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile
    530                 535                 540

Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro
545                 550                 555                 560

Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu
```

```
                565                 570                 575

Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys
            580                 585                 590

Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp
        595                 600                 605

Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu
    610                 615                 620

Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu
625                 630                 635                 640

Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His
                645                 650                 655

Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp
            660                 665                 670

Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu
        675                 680                 685

Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser
    690                 695                 700

Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp
705                 710                 715                 720

Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile
                725                 730                 735

Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu
            740                 745                 750

Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu
        755                 760                 765
```

```
Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu
    770                 775                 780

Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn
785                 790                 795                 800

Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile
                805                 810                 815

Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn
            820                 825                 830

Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys
        835                 840                 845

Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val
    850                 855                 860

Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu
865                 870                 875                 880

Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys
                885                 890                 895

Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn
            900                 905                 910

Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys
        915                 920                 925

Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp
    930                 935                 940

Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln
945                 950                 955                 960

Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala
```

```
               965                 970                 975

Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val
            980                 985                 990

Lys Val Met Gly Arg His Lys Pro  Glu Asn Ile Val Ile  Glu Met Ala
        995                 1000                1005

Arg Glu  Asn Gln Thr Thr Gln  Lys Gly Gln Lys Asn  Ser Arg Glu
    1010                1015                1020

Arg Met  Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
    1025                1030                1035

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1040                1045                1050

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1055                1060                1065

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1070                1075                1080

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1085                1090                1095

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1100                1105                1110

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1115                1120                1125

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1130                1135                1140

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1145                1150                1155
```

```
Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1160                1165                1170

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1175                1180                1185

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1190                1195                1200

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1205                1210                1215

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1220                1225                1230

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1235                1240                1245

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1250                1255                1260

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1265                1270                1275

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1280                1285                1290

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1295                1300                1305

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1310                1315                1320

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1325                1330                1335

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
```

```
    1340                1345                1350

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1355                1360                1365

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1370                1375                1380

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1385                1390                1395

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1400                1405                1410

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1415                1420                1425

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1430                1435                1440

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1445                1450                1455

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1460                1465                1470

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1475                1480                1485

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1490                1495                1500

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1505                1510                1515

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1520                1525                1530
```

```
Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1535                1540                1545

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1550                1555                1560

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1565                1570                1575

Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1580                1585                1590

Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1595                1600                1605

Leu Gly  Gly Asp Ser Gly Gly  Ser Thr Asn Leu Ser  Asp Ile Ile
    1610                1615                1620

Glu Lys  Glu Thr Gly Lys Gln  Leu Val Ile Gln Glu  Ser Ile Leu
    1625                1630                1635

Met Leu  Pro Glu Glu Val Glu  Glu Val Ile Gly Asn  Lys Pro Glu
    1640                1645                1650

Ser Asp  Ile Leu Val His Thr  Ala Tyr Asp Glu Ser  Thr Asp Glu
    1655                1660                1665

Asn Val  Met Leu Leu Thr Ser  Asp Ala Pro Glu Tyr  Lys Pro Trp
    1670                1675                1680

Ala Leu  Val Ile Gln Asp Ser  Asn Gly Glu Asn Lys  Ile Lys Met
    1685                1690                1695

Leu Ser  Gly Gly Ser Pro Lys  Lys Lys Arg Lys Val
    1700                1705                1710
```

<210> 16

<211> 5511
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 16
gggucccgca gucggcgucc agcggcucug cuuguucgug ugugugucgu ugcaggccuu 60

auucggaucc accaugagcu cagagacugg cccaguggcu guggacccca cauugagacg 120

gcggaucgag ccccaugagu uugagguauu cuucgauccg agagagcucc gcaaggagac 180

cugccugcuu uacgaaauua auuggggggg ccggcacucc auuuggcgac auacaucaca 240

gaacacuaac aagcacgucg aagucaacuu caucgagaag uucacgacag aaagauauuu 300

cuguccgaac acaaggugca gcauuaccug guuucucagc uggagcccau gcggcgaaug 360

uaguagggcc aucacugaau uccugucaag guaucccCac gucacucugu uuauuuacau 420

cgcaaggcug uaccaccacg cugacccccg caaucgacaa ggccugcggg auuugaucuc 480

uucaggugug acuauccaaa uuaugacuga gcaggaguca ggauacugcu ggagaaacuu 540

ugugaauuau agcccgagua augaagccca cuggccuagg uaucccCauc uguggguacg 600

acuguacguu cuugaacugu acugcaucau acugggccug ccuccuuguc ucaacauucu 660

gagaaggaag cagccacagc ugacauucuu uaccaucgcu cuucagucuu gucauuacca 720

gcgacugccc ccacacauuc ucugggccac cggguugaaa agcggcagcg agacuccggg 780

caccucagag uccgccacac ccgaaaguga uaagaaguac ucaaucgggc uggccaucgg 840

aacuaauucc gugggguggg cagugaucac ggaugaauac aaagugccgu ccaagaaguu 900

caagguccug gggaacaccg auagacacag caucaagaaa aaucucaucg gagcccugcu 960

guuugacucc ggcgaaaccg cagaagcgac ccggcucaaa cguaccgcga ggcgacgcua 1020

cacccggcgg aagaaucgca ucugcuaucu gcaagagauc uuuucgaacg aaauggcaaa 1080

ggucgacgac agcuucuucc accgccugga agaaucuuuc cugguggagg aggacaagaa 1140

gcaugaacgg cauccuaucu uuggaaacau cgucgacgaa guggcguacc acgaaaagua 1200

cccgaccauc uaccaucugc ggaagaaguu gguugacuca acugacaagg ccgaccucag 1260

auugaucuac uuggcccucg cccauaugau caaauuccgc ggacacuucc ugaucgaagg 1320

cgaucugaac ccugauaacu ccgacgugga uaagcuuuuc auucaacugg ugcagaccua 1380

caaccaacug uucgaagaaa acccaaucaa ugcuagcggc gucgaugcca aggccauccu 1440

guccgcccgg cugucgaagu cgcggcgccu cgaaaaccug aucgcacagc ugccgggaga 1500

gaaaaagaac ggacuuuucg gcaacuugau cgcucucuca cugggacuca cucccaauuu 1560

caaguccaau uuugaccugg ccgaggacgc gaagcugcaa cucucaaagg acaccuacga 1620

cgacgacuug gacaauuugc uggcacaaau uggcgaucag uacgcggauc uguuccuugc 1680

cgcuaagaac cuuucggacg caaucuugcu guccgauauc cugcgcguga acaccgaaau 1740

aaccaaagcg ccgcuuagcg ccucgaugau uaagcgguac gacgagcauc accaggaucu 1800

cacgcugcuc aaagcgcucg ugagacagca acugccugaa aaguacaagg agaucuucuu 1860

cgaccagucc aagaaugggu acgcagggua caucgaugga ggcgcuagcc aggaagaguu 1920

cuauaaguuc aucaagccaa uccuggaaaa gauggacgga accgaagaac ugcuggucaa 1980

gcugaacagg gaggaucugc uccggaaaca gagaaccuuu gacaacggau ccauucccca 2040

ccagauccau cugggugagc ugcacgccau cuugcggcgc caggaggacu uuuacccauu 2100

ccucaaggac aaccgggaaa agaucgagaa aauucugacg uuccgcaucc cguauuacgu 2160

gggcccacug gcgcgcggca auucgcgcuu cgcguggaug acuagaaaau cagaggaaac 2220

caucacuccu uggaauuucg aggaaguugu ggauaaggga gcuucggcac aaagcuucau 2280

cgaacgaaug accaacuucg acaagaaucu cccaaacgag aaggugcuuc cuaagcacag 2340

ccuccuuuac gaauacuuca cugucuacaa cgaacugacu aaagugaaau acguuacuga 2400

aggaaugagg aagccggccu uucuguccgg agaacagaag aaagcaauug ucgaucugcu 2460

guucaagacc aaccgcaagg ugaccgucaa gcagcuuaaa gaggacuacu ucaagaagau 2520

cgaguguuuc gacucagugg aaaucagcgg gguggaggac agauucaacg cuucgcuggg 2580

aaccuaucau gaucuccuga agaucaucaa ggacaaggac uuccuugaca acgaggagaa 2640

cgaggacauc cuggaagaua ucguccugac cuugacccuu uucgaggauc gcgagaugau 2700

cgaggagagg cuuaagaccu acgcucaucu cuucgacgau aaggucauga aacaacucaa 2760

gcgccgccgg uacacugguu ggggccgccu cucccgcaag cugaucaacg guauucgcga 2820

uaaacagagc gguaaaacua uccuggauuu ccucaaaucg gauggcuucg cuaaucguaa 2880

cuucaugcaa uugauccacg acgacagccu gaccuuuaag gaggacaucc aaaaagcaca 2940

aguguccgga cagggagacu cacuccauga acacaucgcg aaucuggccg guucgccggc 3000

gauuaagaag ggaauucugc aaacugugaa gguggucgac gagcugguga aggucauggg 3060

acggcacaaa ccggagaaua ucgugauuga aauggcccga gaaaaccaga cuacccagaa 3120

gggccagaaa aacucccgcg aaaggaugaa gcggaucgaa gaaggaauca aggagcuggg 3180

cagccagauc cugaaagagc acccggugga aaacacgcag cugcagaacg agaagcucua 3240

ccuguacuau uugcaaaaug gacgggacau guacguggac caagagcugg acaucaaucg 3300

guugucugau uacgacgugg accacaucgu uccacagucc uuucugaagg augacucgau 3360

cgauaacaag guguugacuc gcagcgacaa gaacagaggg aagucagaua augugccauc 3420

ggaggagguc gugaagaaga ugaagaauua cuggcggcag cuccugaaug cgaagcugau 3480

uacccagaga aaguuugaca aucucacuaa agccgagcgc ggcggacucu cagagcugga 3540

uaaggcugga uucaucaaac ggcagcuggu cgagacucgg cagauuacca agcacguggc 3600

gcagaucuug gacucccgca ugaacacuaa auacgacgag aacgauaagc ucauccggga 3660

agugaaggug auuacccuga aaagcaaacu ugugucggac uuucggaagg acuuucaguu 3720

uuacaaagug agagaaauca acaacuacca ucacgcgcau gacgcauacc ucaacgcugu 3780

ggucgguacc gcccugauca aaaaguaccc uaaacuugaa ucggaguuug uguacggaga 3840

cuacaagguc uacgacguga ggaagaugau agccaagucc gaacaggaaa ucgggaaagc 3900

aacugcgaaa uacuucuuuu acucaaacau caugaacuuu uucaagacug aaauuacgcu 3960

ggccaaugga gaaaucagga agaggccacu gaucgaaacu aacggagaaa cgggcgaaau 4020

cgugugggac aagggcaggg acuucgcaac uguucgcaaa gugcucucua ugccgcaagu 4080

caauauugug aagaaaaccg aagugcaaac cggcggauuu ucaaaggaau cgauccuccc 4140

aaagagaaau agcgacaagc ucauugcacg caagaaagac ugggacccga agaaguacgg 4200

aggauucgau ucgccgacug ucgcauacuc cguccucgug guggccaagg uggagaaggg 4260

aaagagcaaa aagcucaaau ccgucaaaga gcugcugggg auuaccauca uggaacgauc 4320

cucguucgag aagaacccga uugauuuccu cgaggcgaag gguuacaagg aggugaagaa 4380

ggaucugauc aucaaacucc ccaaguacuc acuguucgaa cuggaaaaug gucggaagcg 4440

caugcuggcu ucggccggag aacuccaaaa aggaaaugag cuggccuugc cuagcaagua 4500

cgucaacuuc cucuaucuug cuucgcacua cgaaaaacuc aaagggucac cggaagauaa 4560

cgaacagaag cagcuuuucg uggagcagca caagcauuau cuggaugaaa ucaucgaaca 4620

aaucuccgag uuuucaaagc gcgugauccu cgccgacgcc aaccucgaca aaguccuguc 4680

ggccuacaau aagcauagag auaagccgau cagagaacag gccgagaaca uuauccacuu 4740

guucacccug acuaaccugg gagccccagc cgccuucaag uacuucgaua cuacuaucga 4800

ucgcaaaaga uacacgucca ccaaggaagu ucuggacgcg acccugaucc accaaagcau 4860

cacuggacuc uacgaaacua ggaucgaucu gucgcagcug gguggcgauu cuggugguuc 4920

uacuaaucug ucagauauua uugaaaagga gaccgguaag caacugguua uccaggaauc 4980

cauccucaug cucccagagg agguggaaga agucauuggg aacaagccgg aaagcgauau 5040

acucgugcac accgccuacg acgagagcac cgacgagaau gucaugcuuc ugacuagcga 5100

cgccccugaa uacaagccuu gggcucuggu cauacaggau agcaacggug agaacaagau 5160

uaagaugcuc ucuggugguu cucccaagaa gaagaggaaa gucuaauagu cuagccauca 5220

cauuuaaaag caucucagcc uaccaugaga auaagagaaa gaaaaugaag aucaauagcu 5280

uauucaucuc uuuuucuuuu ucguuggugu aaagccaaca cccugucuaa aaaacauaaa 5340

uuucuuuaau cauuuugccu cuuuucucug ugcuucaauu aauaaaaaau ggaaagaacc 5400

ucgagaaaaa aaaaaaaaaa aaaaaaaaaa aaaaagcgaa aaaaaaaaaa aaaaaaaaaa 5460

aaaaaaaacc gaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa u 5511

<210> 17
<211> 5133
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 17
augagcucag agacuggccc aguggcugug gaccccacau ugagacggcg gaucgagccc 60

caugaguuug agguauucuu cgauccgaga gagcuccgca aggagaccug ccugcuuuac 120

gaaauuaauu gggggggccg gcacuccauu uggcgacaua caucacagaa cacuaacaag 180

cacgucgaag ucaacuucau cgagaaguuc acgacagaaa gauauuucug uccgaacaca 240

aggugcagca uuaccugguu ucucagcugg agcccaugcg gcgaauguag uagggccauc 300

acugaauucc ugucaaggua uccccacguc acucuguuua uuuacaucgc aaggcuguac 360

caccacgcug acccccgcaa ucgacaaggc cugcgggauu ugaucucuuc aggugugacu 420

auccaaauua ugacugagca ggagucagga uacugcugga gaaacuuugu gaauuauagc 480

ccgaguaaug aagcccacug gccuagguau ccccaucugu ggguacgacu guacguucuu 540

gaacuguacu gcaucauacu gggccugccu ccuugucuca acauucugag aaggaagcag 600

ccacagcuga cauucuuuac caucgcucuu cagucuuguc auuaccagcg acugccccca 660

cacauucucu gggccaccgg guugaaaagc ggcagcgaga cuccgggcac cucagagucc 720

gccacacccg aaagugauaa gaaguacuca aucgggcugg ccaucggaac uaauuccgug 780

gguugggcag ugaucacgga ugaauacaaa gugccgucca agaaguucaa gguccugggg 840

aacaccgaua gacacagcau caagaaaaau cucaucggag cccugcuguu ugacuccggc 900

gaaaccgcag aagcgacccg gcucaaacgu accgcgaggc gacgcuacac ccggcggaag 960

aaucgcaucu gcuaucugca agagaucuuu ucgaacgaaa uggcaaaggu cgacgacagc 1020

uucuuccacc gccuggaaga aucuuuccug guggaggagg acaagaagca ugaacggcau 1080

ccuaucuuug gaaacaucgu cgacgaagug gcguaccacg aaaaguaccc gaccaucuac 1140

caucugcgga agaaguuggu ugacucaacu gacaaggccg accucagauu gaucuacuug 1200

gcccucgccc auaugaucaa auuccgcgga cacuuccuga ucgaaggcga ucugaacccu 1260

gauaacuccg acguggauaa gcuuuucauu caacuggugc agaccuacaa ccaacuguuc 1320

gaagaaaacc caaucaaugc uagcggcguc gaugccaagg ccauccuguc cgcccggcug 1380

ucgaagucgc ggcgccucga aaaccugauc gcacagcugc cgggagagaa aaagaacgga 1440

cuuuucggca acuugaucgc ucucucacug ggacucacuc ccaauuucaa guccaauuuu 1500

gaccuggccg aggacgcgaa gcugcaacuc ucaaaggaca ccuacgacga cgacuuggac 1560

aauuugcugg cacaaauugg cgaucaguac gcggaucugu uccuugccgc uaagaaccuu 1620

ucggacgcaa ucuugcuguc cgauauccug cgcgugaaca ccgaaauaac caaagcgccg 1680

cuuagcgccu cgaugauuaa gcgguacgac gagcaucacc aggaucucac gcugcucaaa 1740

gcgcucguga gacagcaacu gccugaaaag uacaaggaga ucuucuucga ccaguccaag 1800

aauggguacg caggguacau cgauggaggc gcuagccagg aagaguucua uaaguucauc 1860

aagccaaucc uggaaaagau ggacggaacc gaagaacugc uggucaagcu gaacagggag 1920

gaucugcucc ggaaacagag aaccuuugac aacggaucca uuccccacca gauccaucug 1980

ggugagcugc acgccaucuu gcggcgccag gaggacuuuu acccauuccu caaggacaac 2040

cgggaaaaga ucgagaaaau ucugacguuc cgcaucccgu auuacguggg cccacuggcg 2100

cgcggcaauu cgcgcuucgc guggaugacu agaaaaucag aggaaaccau cacuccuugg 2160

aauuucgagg aaguugugga uaagggagcu ucggcacaaa gcuucaucga acgaaugacc 2220

aacuucgaca agaaucuccc aaacgagaag gugcuuccua agcacagccu ccuuuacgaa 2280

uacuucacug ucuacaacga acugacuaaa gugaaauacg uuacugaagg aaugaggaag 2340

ccggccuuuc uguccggaga acagaagaaa gcaauugucg aucugcuguu caagaccaac 2400

cgcaagguga ccgucaagca gcuuaaagag gacuacuuca agaagaucga guguuucgac 2460

ucaguggaaa ucagcggggu ggaggacaga uucaacgcuu cgcugggaac cuaucaugau 2520

cuccugaaga ucaucaagga caaggacuuc cuugacaacg aggagaacga ggacauccug 2580

gaagauaucg uccugaccuu gacccuuuuc gaggaucgcg agaugaucga ggagaggcuu 2640

aagaccuacg cucaucucuu cgacgauaag gucaugaaac aacucaagcg ccgccgguac 2700

acugguuggg gccgccucuc ccgcaagcug aucaacggua uucgcgauaa acagagcggu 2760

aaaacuaucc uggauuuccu caaaucggau ggcuucgcua aucguaacuu caugcaauug 2820

auccacgacg acagccugac cuuuaaggag gacauccaaa aagcacaagu guccggacag 2880

ggagacucac uccaugaaca caucgcgaau cuggccgguu cgccggcgau uaagaaggga 2940

auucugcaaa cugugaaggu ggucgacgag cuggugaagg ucaugggacg gcacaaaccg 3000

gagaauaucg ugauugaaau ggcccgagaa aaccagacua cccagaaggg ccagaaaaac 3060

ucccgcgaaa ggaugaagcg gaucgaagaa ggaaucaagg agcugggcag ccagauccug 3120

aaagagcacc cgguggaaaa cacgcagcug cagaacgaga agcucuaccu guacuauuug 3180

caaaauggac gggacaugua cguggaccaa gagcuggaca ucaaucgguu gucugauuac 3240

gacguggacc acaucguucc acaguccuuu cugaaggaug acucgaucga uaacaaggug 3300

uugacucgca gcgacaagaa cagagggaag ucagauaaug ugccaucgga ggaggucgug 3360

aagaagauga agaauuacug gcggcagcuc cugaaugcga agcugauuac ccagagaaag 3420

uuugacaauc ucacuaaagc cgagcgcggc ggacucucag agcuggauaa ggcuggauuc 3480

aucaaacggc agcuggucga gacucggcag auuaccaagc acguggcgca gaucuuggac 3540

ucccgcauga acacuaaaua cgacgagaac gauaagcuca uccgggaagu gaaggugauu 3600

acccugaaaa gcaaacuugu gucggacuuu cggaaggacu uucaguuuua caaagugaga 3660

gaaaucaaca acuaccauca cgcgcaugac gcauaccuca acgcuguggu cgguaccgcc 3720

cugaucaaaa aguacccuaa acuugaaucg gaguuugugu acggagacua caaggucuac 3780

gacgugagga agaugauagc caaguccgaa caggaaaucg ggaaagcaac ugcgaaauac 3840

uucuuuuacu caaacaucau gaacuuuuuc aagacugaaa uuacgcuggc caauggagaa 3900

aucaggaaga ggccacugau cgaaacuaac ggagaaacgg gcgaaaucgu gugggacaag 3960

ggcagggacu ucgcaacugu ucgcaaagug cucucuaugc cgcaagucaa uauugugaag 4020

aaaaccgaag ugcaaaccgg cggauuuuca aaggaaucga uccucccaaa gagaaauagc 4080

gacaagcuca uugcacgcaa gaaagacugg gacccgaaga aguacggagg auucgauucg 4140

ccgacugucg cauacuccgu ccucguggug gccaaggugg agaagggaaa gagcaaaaag 4200

cucaaauccg ucaaagagcu gcuggggauu accaucaugg aacgauccuc guucgagaag 4260

aacccgauug auuuccucga ggcgaagggu uacaaggagg ugaagaagga ucugaucauc 4320

aaacucccca aguacucacu guucgaacug gaaaaugguc ggaagcgcau gcuggcuucg 4380

gccggagaac uccaaaaagg aaaugagcug gccuugccua gcaaguacgu caacuuccuc 4440

uaucuugcuu cgcacuacga aaaacucaaa gggucaccgg aagauaacga acagaagcag 4500

cuuuucgugg agcagcacaa gcauuaucug gaugaaauca ucgaacaaau cuccgaguuu 4560

ucaaagcgcg ugauccucgc cgacgccaac cucgacaaag uccugucggc cuacaauaag 4620

cauagagaua agccgaucag agaacaggcc gagaacauua uccacuuguu cacccugacu 4680

aaccugggag ccccagccgc cuucaaguac uucgauacua cuaucgaucg caaaagauac 4740

acguccacca aggaaguucu ggacgcgacc cugauccacc aaagcaucac uggacucuac 4800

gaaacuagga ucgaucuguc gcagcugggu ggcgauucug gugguucuac uaaucuguca 4860

gauauuauug aaaaggagac cgguaagcaa cugguuaucc aggaauccau ccucaugcuc 4920

ccagaggagg uggaagaagu cauugggaac aagccggaaa gcgauauacu cgugcacacc 4980

gccuacgacg agagcaccga cgagaauguc augcuucuga cuagcgacgc cccugaauac 5040

aagccuuggg cucuggucau acaggauagc aacggugaga acaagauuaa gaugcucucu 5100

ggugguucuc ccaagaagaa gaggaaaguc uaa 5133

<210> 18
<211> 1710
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 18

Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1 5 10 15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
20 25 30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
35 40 45

Ser Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val
50 55 60

```
Asn Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr
65                  70                  75                  80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
                85                  90                  95

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val Thr Leu
            100                 105                 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg Asn Arg
        115                 120                 125

Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met
    130                 135                 140

Thr Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn Tyr Ser
145                 150                 155                 160

Pro Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp Val Arg
                165                 170                 175

Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
            180                 185                 190

Leu Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile
        195                 200                 205

Ala Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp
    210                 215                 220

Ala Thr Gly Leu Lys Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser
225                 230                 235                 240

Ala Thr Pro Glu Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly
                245                 250                 255

Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro
```

```
            260                 265                 270

Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys
        275                 280                 285

Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu
    290                 295                 300

Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys
305                 310                 315                 320

Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys
                325                 330                 335

Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu
            340                 345                 350

Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp
        355                 360                 365

Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys
    370                 375                 380

Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu
385                 390                 395                 400

Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly
                405                 410                 415

Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu
            420                 425                 430

Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser
        435                 440                 445

Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg
    450                 455                 460
```

```
Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly
465                 470                 475                 480

Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe
                485                 490                 495

Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys
            500                 505                 510

Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp
        515                 520                 525

Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile
    530                 535                 540

Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro
545                 550                 555                 560

Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu
                565                 570                 575

Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys
            580                 585                 590

Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp
        595                 600                 605

Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu
    610                 615                 620

Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu
625                 630                 635                 640

Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His
                645                 650                 655

Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp
```

```
            660                 665                 670

Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu
        675                 680                 685

Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser
    690                 695                 700

Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp
705                 710                 715                 720

Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile
                725                 730                 735

Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu
            740                 745                 750

Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu
        755                 760                 765

Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu
    770                 775                 780

Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn
785                 790                 795                 800

Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile
                805                 810                 815

Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn
            820                 825                 830

Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys
        835                 840                 845

Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val
    850                 855                 860
```

```
Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu
865                 870                 875                 880


Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys
                885                 890                 895


Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn
            900                 905                 910


Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys
        915                 920                 925


Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp
    930                 935                 940


Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln
945                 950                 955                 960


Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala
                965                 970                 975


Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val
            980                 985                 990


Lys Val Met Gly Arg His Lys Pro  Glu Asn Ile Val Ile  Glu Met Ala
        995                 1000                1005


Arg Glu  Asn Gln Thr Thr Gln  Lys Gly Gln Lys Asn  Ser Arg Glu
    1010                1015                1020


Arg Met  Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
    1025                1030                1035


Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1040                1045                1050


Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
```

```
    1055                1060                1065

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1070                1075                1080

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1085                1090                1095

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1100                1105                1110

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1115                1120                1125

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1130                1135                1140

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1145                1150                1155

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1160                1165                1170

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1175                1180                1185

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1190                1195                1200

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1205                1210                1215

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1220                1225                1230

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1235                1240                1245
```

```
Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1250                1255                1260

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1265                1270                1275

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1280                1285                1290

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1295                1300                1305

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1310                1315                1320

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1325                1330                1335

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1340                1345                1350

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1355                1360                1365

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1370                1375                1380

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1385                1390                1395

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1400                1405                1410

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1415                1420                1425

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
```

```
    1430                1435                1440

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1445                1450                1455

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1460                1465                1470

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1475                1480                1485

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1490                1495                1500

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1505                1510                1515

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1520                1525                1530

Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1535                1540                1545

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1550                1555                1560

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1565                1570                1575

Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1580                1585                1590

Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1595                1600                1605

Leu Gly  Gly Asp Ser Gly Gly  Ser Thr Asn Leu Ser  Asp Ile Ile
    1610                1615                1620
```

Glu Lys Glu Thr Gly Lys Gln Leu Val Ile Gln Glu Ser Ile Leu
1625 1630 1635

Met Leu Pro Glu Glu Val Glu Glu Val Ile Gly Asn Lys Pro Glu
1640 1645 1650

Ser Asp Ile Leu Val His Thr Ala Tyr Asp Glu Ser Thr Asp Glu
1655 1660 1665

Asn Val Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr Lys Pro Trp
1670 1675 1680

Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile Lys Met
1685 1690 1695

Leu Ser Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
1700 1705 1710

<210> 19
<211> 5458
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 19
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgugggа cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacagc ggaagcgaaa 720

caccgggaac aagcgaaagc gcaacaccgg aaagcgacaa gaaguacagc aucggacugg 780

ccaucggaac aaacagcguc ggaugggcag ucaucacaga cgaauacaag gucccgagca 840

agaaguucaa gguccuggga aacacagaca gacacagcau caagaagaac cugaucggag 900

cacugcuguu cgacagcgga gaaacagcag aagcaacaag acugaagaga acagcaagaa 960

gaagauacac aagaagaaag aacagaaucu gcuaccugca ggaaaucuuc agcaacgaaa 1020

uggcaaaggu cgacgacagc uucuuccaca gacuggaaga aagcuuccug gucgaagaag 1080

acaagaagca cgaaagacac ccgaucuucg gaaacaucgu cgacgaaguc gcauaccacg 1140

aaaaguaccc gacaaucuac caccugagaa agaagcuggu cgacagcaca gacaaggcag 1200

accugagacu gaucuaccug gcacuggcac acaugaucaa guucagagga cacuuccuga 1260

ucgaaggaga ccugaacccg gacaacagcg acgucgacaa gcuguucauc cagcuggucc 1320

agacauacaa ccagcuguuc gaagaaaacc cgaucaacgc aagcggaguc gacgcaaagg 1380

caauccugag cgcaagacug agcaagagca gaagacugga aaaccugauc gcacagcugc 1440

cgggagaaaa gaagaacgga cuguucggaa accugaucgc acugagccug ggacugacac 1500

cgaacuucaa gagcaacuuc gaccuggcag aagacgcaaa gcugcagcug agcaaggaca 1560

cauacgacga cgaccuggac aaccugcugg cacagaucgg agaccaguac gcagaccugu 1620

uccuggcagc aaagaaccug agcgacgcaa uccugcugag cgacauccug agagucaaca 1680

cagaaaucac aaaggcaccg cugagcgcaa gcaugaucaa gagauacgac gaacaccacc 1740

aggaccugac acugcugaag gcacugguca gacagcagcu gccggaaaag uacaaggaaa 1800

ucuucuucga ccagagcaag aacggauacg caggauacau cgacggagga gcaagccagg 1860

aagaauucua caaguucauc aagccgaucc uggaaaagau ggacggaaca gaagaacugc 1920

uggucaagcu gaacagagaa gaccugcuga gaaagcagag aacauucgac aacggaagca 1980

ucccgcacca gauccaccug ggagaacugc acgcaauccu gagaagacag gaagacuucu 2040

acccguuccu gaaggacaac agagaaaaga ucgaaaagau ccugacauuc agaaucccgu 2100

acuacgucgg accgcuggca agaggaaaca gcagauucgc auggaugaca agaaagagcg 2160

aagaaacaau cacaccgugg aacuucgaag aagucgucga caagggagca agcgcacaga 2220

gcuucaucga aagaaugaca aacuucgaca agaaccugcc gaacgaaaag guccugccga 2280

agcacagccu gcuguacgaa uacuucacag ucuacaacga acugacaaag gucaaguacg 2340

ucacagaagg aaugagaaag ccggcauucc ugagcggaga acagaagaag gcaaucgucg 2400

accugcuguu caagacaaac agaaagguca cagucaagca gcugaaggaa gacuacuuca 2460

agaagaucga augcuucgac agcgucgaaa ucagcggagu cgaagacaga uucaacgcaa 2520

gccugggaac auaccacgac cugcugaaga ucaucaagga caaggacuuc cuggacaacg 2580

aagaaaacga agacauccug gaagacaucg uccugacacu gacacuguuc gaagacagag 2640

aaaugaucga agaaagacug aagacauacg cacaccuguu cgacgacaag gucaugaagc 2700

agcugaagag aagaagauac acaggauggg gaagacugag cagaaagcug aucaacggaa 2760

ucagagacaa gcagagcgga aagacaaucc uggacuuccu gaagagcgac ggauucgcaa 2820

acagaaacuu caugcagcug auccacgacg acagccugac auucaaggaa gacauccaga 2880

aggcacaggu cagcggacag ggagacagcc ugcacgaaca caucgcaaac cuggcaggaa 2940

gcccggcaau caagaaggga auccugcaga cagucaaggu cgucgacgaa cuggucaagg 3000

ucaugggaag acacaagccg gaaaacaucg ucaucgaaau ggcaagagaa aaccagacaa 3060

cacagaaggg acagaagaac agcagagaaa gaaugaagag aaucgaagaa ggaaucaagg 3120

aacugggaag ccagauccug aaggaacacc cggucgaaaa cacacagcug cagaacgaaa 3180

agcuguaccu guacuaccug cagaacggaa gagacaugua cgucgaccag gaacuggaca 3240

ucaacagacu gagcgacuac gacgucgacc acaucguccc gcagagcuuc cugaaggacg 3300

acagcaucga caacaagguc cugacaagaa gcgacaagaa cagaggaaag agcgacaacg 3360

ucccgagcga agaagucguc aagaagauga agaacuacug gagacagcug cugaacgcaa 3420

agcugaucac acagagaaag uucgacaacc ugacaaaggc agagagagga ggacugagcg 3480

aacuggacaa ggcaggauuc aucaagagac agcuggucga aacaagacag aucacaaagc 3540

acgucgcaca gauccuggac agcagaauga acacaaagua cgacgaaaac gacaagcuga 3600

ucagagaagu caaggucauc acacugaaga gcaagcuggu cagcgacuuc agaaaggacu 3660

uccaguucua caaggucaga gaaaucaaca acuaccacca cgcacacgac gcauaccuga 3720

acgcagucgu cggaacagca cugaucaaga aguacccgaa gcuggaaagc gaauucgucu 3780

acggagacua caaggucuac gacgucagaa agaugaucgc aaagagcgaa caggaaaucg 3840

gaaaggcaac agcaaaguac uucuucuaca gcaacaucau gaacuucuuc aagacagaaa 3900

ucacacuggc aaacggagaa aucagaaaga gaccgcugau cgaaacaaac ggagaaacag 3960

gagaaaucgu cugggacaag ggaagagacu ucgcaacagu cagaaagguc cugagcaugc 4020

cgcaggucaa caucgucaag aagacagaag uccagacagg aggauucagc aaggaaagca 4080

uccugccgaa gagaaacagc gacaagcuga ucgcaagaaa gaaggacugg gacccgaaga 4140

aguacggagg auucgacagc ccgacagucg cauacagcgu ccuggucguc gcaaaggucg 4200

aaaagggaaa gagcaagaag cugaagagcg ucaaggaacu gcugggaauc acaaucaugg 4260

aaagaagcag cuucgaaaag aacccgaucg acuuccugga agcaaaggga uacaaggaag 4320

ucaagaagga ccugaucauc aagcugccga aguacagccu guucgaacug gaaaacggaa 4380

gaaagagaau gcuggcaagc gcaggagaac ugcagaaggg aaacgaacug gcacugccga 4440

gcaaguacgu caacuuccug uaccuggcaa gccacuacga aaagcugaag ggaagcccgg 4500

aagacaacga acagaagcag cuguucgucg aacagcacaa gcacuaccug gacgaaauca 4560

ucgaacagau cagcgaauuc agcaagagag ucauccuggc agacgcaaac cuggacaagg 4620

uccugagcgc auacaacaag cacagagaca agccgaucag agaacaggca gaaaacauca 4680

uccaccuguu cacacugaca aaccugggag caccggcagc auucaaguac uucgacacaa 4740

caaucgacag aaagagauac acaagcacaa aggaaguccu ggacgcaaca cugauccacc 4800

agagcaucac aggacuguac gaaacaagaa ucgaucugag ccagcuggga ggagacagcg 4860

gaggaagcac aaaccugagc gacaucaucg aaaaggaaac aggaaagcag cuggucaucc 4920

aggaaagcau ccugaugcug ccggaagaag ucgaagaagu caucggaaac aagccggaaa 4980

gcgacauccu gguccacaca gcauacgacg aaagcacaga cgaaaacguc augcugcuga 5040

caagcgacgc accggaauac aagccguggg cacuggucau ccaggacagc aacggagaaa 5100

acaagaucaa gaugcugagc ggaggaagcc cgaagaagaa gagaaagguc uaauagucua 5160

gacaucacau uuaaaagcau cucagccuac caugagaaua agagaaagaa aaugaagauc 5220

aauagcuuau ucaucucuuu uucuuuuucg uugguguaaa gccaacaccc ugucuaaaaa 5280

acauaaauuu cuuuaaucau uuugccucuu uucucugugc uucaauuaau aaaaaaugga 5340

aagaaccucg agaaaaaaaa aaaaaaaaaa aaaaaaaaaa aagcgaaaaa aaaaaaaaaa 5400

aaaaaaaaaa aaaaaccgaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaau 5458

<210> 20
<211> 5043
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 20
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacagc 600

ggaagcgaaa caccgggaac aagcgaaagc gcaacaccgg aaagcgacaa gaaguacagc 660

aucggacugg ccaucggaac aaacagcguc ggaugggcag ucaucacaga cgaauacaag 720

gucccgagca agaaguucaa gguccuggga aacacagaca gacacagcau caagaagaac 780

cugaucggag cacugcuguu cgacagcgga gaaacagcag aagcaacaag acugaagaga 840

acagcaagaa gaagauacac aagaagaaag aacagaaucu gcuaccugca ggaaaucuuc 900

agcaacgaaa uggcaaaggu cgacgacagc uucuuccaca gacuggaaga aagcuuccug 960

gucgaagaag acaagaagca cgaaagacac ccgaucuucg gaaacaucgu cgacgaaguc 1020

gcauaccacg aaaaguaccc gacaaucuac caccugagaa agaagcuggu cgacagcaca 1080

gacaaggcag accugagacu gaucuaccug gcacuggcac acaugaucaa guucagagga 1140

cacuuccuga ucgaaggaga ccugaacccg gacaacagcg acgucgacaa gcuguucauc 1200

cagcuggucc agacauacaa ccagcuguuc gaagaaaacc cgaucaacgc aagcggaguc 1260

gacgcaaagg caauccugag cgcaagacug agcaagagca gaagacugga aaaccugauc 1320

gcacagcugc cgggagaaaa gaagaacgga cuguucggaa accugaucgc acugagccug 1380

ggacugacac cgaacuucaa gagcaacuuc gaccuggcag aagacgcaaa gcugcagcug 1440

agcaaggaca cauacgacga cgaccuggac aaccugcugg cacagaucgg agaccaguac 1500

gcagaccugu uccuggcagc aaagaaccug agcgacgcaa uccugcugag cgacauccug 1560

agagucaaca cagaaaucac aaaggcaccg cugagcgcaa gcaugaucaa gagauacgac 1620

gaacaccacc aggaccugac acugcugaag gcacugguca gacagcagcu gccggaaaag 1680

uacaaggaaa ucuucuucga ccagagcaag aacggauacg caggauacau cgacggagga 1740

gcaagccagg aagaauucua caaguucauc aagccgaucc uggaaaagau ggacggaaca 1800

gaagaacugc uggucaagcu gaacagagaa gaccugcuga gaaagcagag aacauucgac 1860

aacggaagca ucccgcacca gauccaccug ggagaacugc acgcaauccu gagaagacag 1920

gaagacuucu acccguuccu gaaggacaac agagaaaaga ucgaaaagau ccugacauuc 1980

agaaucccgu acuacgucgg accgcuggca agaggaaaca gcagauucgc auggaugaca 2040

agaaagagcg aagaaacaau cacaccgugg aacuucgaag aagucgucga caagggagca 2100

agcgcacaga gcuucaucga aagaaugaca aacuucgaca agaaccugcc gaacgaaaag 2160

guccugccga agcacagccu gcuguacgaa uacuucacag ucuacaacga acugacaaag 2220

gucaaguacg ucacagaagg aaugagaaag ccggcauucc ugagcggaga acagaagaag 2280

gcaaucgucg accugcuguu caagacaaac agaaagguca cagucaagca gcugaaggaa 2340

gacuacuuca agaagaucga augcuucgac agcgucgaaa ucagcggagu cgaagacaga 2400

uucaacgcaa gccugggaac auaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aagaaaacga agacauccug gaagacaucg uccugacacu gacacuguuc 2520

gaagacagag aaaugaucga agaaagacug aagacauacg cacaccuguu cgacgacaag 2580

gucaugaagc agcugaagag aagaagauac acaggauggg gaagacugag cagaaagcug 2640

aucaacggaa ucagagacaa gcagagcgga aagacaaucc uggacuuccu gaagagcgac 2700

ggauucgcaa acagaaacuu caugcagcug auccacgacg acagccugac auucaaggaa 2760

gacauccaga aggcacaggu cagcggacag ggagacagcc ugcacgaaca caucgcaaac 2820

cuggcaggaa gcccggcaau caagaaggga auccugcaga cagucaaggu cgucgacgaa 2880

cuggucaagg ucaugggaag acacaagccg gaaaacaucg ucaucgaaau ggcaagagaa 2940

aaccagacaa cacagaaggg acagaagaac agcagagaaa gaaugaagag aaucgaagaa 3000

ggaaucaagg aacugggaag ccagauccug aaggaacacc cggucgaaaa cacacagcug 3060

cagaacgaaa agcuguaccu guacuaccug cagaacggaa gagacaugua cgucgaccag 3120

gaacuggaca ucaacagacu gagcgacuac gacgucgacc acaucguccc gcagagcuuc 3180

cugaaggacg acagcaucga caacaagguc cugacaagaa gcgacaagaa cagaggaaag 3240

agcgacaacg ucccgagcga agaagucguc aagaagauga agaacuacug gagacagcug 3300

cugaacgcaa agcugaucac acagagaaag uucgacaacc ugacaaaggc agagagagga 3360

ggacugagcg aacuggacaa ggcaggauuc aucaagagac agcuggucga aacaagacag 3420

aucacaaagc acgucgcaca gauccuggac agcagaauga acacaaagua cgacgaaaac 3480

gacaagcuga ucagagaagu caaggucauc acacugaaga gcaagcuggu cagcgacuuc 3540

agaaaggacu uccaguucua caaggucaga gaaaucaaca acuaccacca cgcacacgac 3600

gcauaccuga acgcagucgu cggaacagca cugaucaaga aguacccgaa gcuggaaagc 3660

gaauucgucu acggagacua caaggucuac gacgucagaa agaugaucgc aaagagcgaa 3720

caggaaaucg gaaaggcaac agcaaaguac uucuucuaca gcaacaucau gaacuucuuc 3780

aagacagaaa ucacacuggc aaacggagaa aucagaaaga gaccgcugau cgaaacaaac 3840

ggagaaacag gagaaaucgu cugggacaag ggaagagacu ucgcaacagu cagaaagguc 3900

cugagcaugc cgcaggucaa caucgucaag aagacagaag uccagacagg aggauucagc 3960

aaggaaagca uccugccgaa gagaaacagc gacaagcuga ucgcaagaaa gaaggacugg 4020

gacccgaaga aguacggagg auucgacagc ccgacagucg cauacagcgu ccuggucguc 4080

gcaaaggucg aaaagggaaa gagcaagaag cugaagagcg ucaaggaacu gcugggaauc 4140

acaaucaugg aaagaagcag cuucgaaaag aacccgaucg acuuccugga agcaaaggga 4200

uacaaggaag ucaagaagga ccugaucauc aagcugccga aguacagccu guucgaacug 4260

gaaaacggaa gaaagagaau gcuggcaagc gcaggagaac ugcagaaggg aaacgaacug 4320

gcacugccga gcaaguacgu caacuuccug uaccuggcaa gccacuacga aaagcugaag 4380

ggaagcccgg aagacaacga acagaagcag cuguucgucg aacagcacaa gcacuaccug 4440

gacgaaauca ucgaacagau cagcgaauuc agcaagagag ucauccuggc agacgcaaac 4500

cuggacaagg uccugagcgc auacaacaag cacagagaca agccgaucag agaacaggca 4560

gaaaacauca uccaccuguu cacacugaca aaccugggag caccggcagc auucaaguac 4620

uucgacacaa caaucgacag aaagagauac acaagcacaa aggaaguccu ggacgcaaca 4680

cugauccacc agagcaucac aggacuguac gaaacaagaa ucgaucugag ccagcuggga 4740

ggagacagcg gaggaagcac aaaccugagc gacaucaucg aaaaggaaac aggaaagcag 4800

cuggucaucc aggaaagcau ccugaugcug ccggaagaag ucgaagaagu caucggaaac 4860

aagccggaaa gcgacauccu gguccacaca gcauacgacg aaagcacaga cgaaaacguc 4920

augcugcuga caagcgacgc accggaauac aagccguggg cacuggucau ccaggacagc 4980

aacggagaaa acaagaucaa gaugcugagc ggaggaagcc cgaagaagaa gagaaagguc 5040

uaa 5043

<210> 21
<211> 1680
<212> PRT

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 21

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15


Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30


Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45


Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60


Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80


Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95


Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110


Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125


Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140


Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160


Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175
```

```
Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser Glu Thr Pro Gly Thr Ser
        195                 200                 205

Glu Ser Ala Thr Pro Glu Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
```

```
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575
```

```
Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
```

```
    770                 775                 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
            820                 825                 830

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
                965                 970                 975
```

```
Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
```

```
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350
```

```
Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500

Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
```

```
    1535                1540                1545

Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560

Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575

Leu Gly  Gly Asp Ser Gly Gly  Ser Thr Asn Leu Ser  Asp Ile Ile
    1580                1585                1590

Glu Lys  Glu Thr Gly Lys Gln  Leu Val Ile Gln Glu  Ser Ile Leu
    1595                1600                1605

Met Leu  Pro Glu Glu Val Glu  Glu Val Ile Gly Asn  Lys Pro Glu
    1610                1615                1620

Ser Asp  Ile Leu Val His Thr  Ala Tyr Asp Glu Ser  Thr Asp Glu
    1625                1630                1635

Asn Val  Met Leu Leu Thr Ser  Asp Ala Pro Glu Tyr  Lys Pro Trp
    1640                1645                1650

Ala Leu  Val Ile Gln Asp Ser  Asn Gly Glu Asn Lys  Ile Lys Met
    1655                1660                1665

Leu Ser  Gly Gly Ser Pro Lys  Lys Lys Arg Lys Val
    1670                1675                1680
```

<210> 22

<400> 22
000

<210> 23
<211> 4197
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 23

auggacaaga aguacuccau cggccuggac aucggcacca acuccguggg cugggccgug 60

aucaccgacg aguacaaggu gcccuccaag aaguucaagg ugcugggcaa caccgaccgg 120

cacuccauca agaagaaccu gaucggcgcc cugcuguucg acuccggcga gaccgccgag 180

gccacccggc ugaagcggac cgcccggcgg cgguacaccc ggcggaagaa ccggaucugc 240

uaccugcagg agaucuucuc caacgagaug gccaaggugg acgacuccuu cuuccaccgg 300

cuggaggagu ccuuccuggu ggaggaggac aagaagcacg agcggcaccc caucuucggc 360

aacaucgugg acgagguggc cuaccacgag aaguacccca ccaucuacca ccugcggaag 420

aagcuggugg acuccaccga caaggccgac cugcggcuga ucuaccuggc ccuggcccac 480

augaucaagu uccggggcca cuuccugauc gagggcgacc ugaaccccga caacuccgac 540

guggacaagc uguucaucca gcuggugcag accuacaacc agcuguucga ggagaacccc 600

aucaacgccu ccggcgugga cgccaaggcc auccuguccg cccggcuguc caagucccgg 660

cggcuggaga accugaucgc ccagcugccc ggcgagaaga agaacggccu guucggcaac 720

cugaucgccc ugucccuggg ccugaccccc aacuucaagu ccaacuucga ccuggccgag 780

gacgccaagc ugcagcuguc caaggacacc uacgacgacg accuggacaa ccugcuggcc 840

cagaucggcg accaguacgc cgaccuguuc cuggccgcca agaaccuguc cgacgccauc 900

cugcuguccg acauccugcg ggugaacacc gagaucacca aggccccccu guccgccucc 960

augaucaagc gguacgacga gcaccaccag gaccugaccc ugcugaaggc ccuggugcgg 1020

cagcagcugc ccgagaagua caaggagauc uucuucgacc aguccaagaa cggcuacgcc 1080

ggcuacaucg acggcggcgc cucccaggag gaguucuaca aguucaucaa gcccauccug 1140

gagaagaugg acggcaccga ggagcugcug gugaagcuga accgggagga ccugcugcgg 1200

aagcagcgga ccuucgacaa cggcuccauc ccccaccaga uccaccuggg cgagcugcac 1260

gccauccugc ggcggcagga ggacuucuac cccuuccuga aggacaaccg ggagaagauc 1320

gagaagaucc ugaccuuccg gaucccccuac uacgugggcc cccuggcccg gggcaacucc 1380

cgguucgccu ggaugacccg gaaguccgag gagaccauca cccccuggaa cuucgaggag 1440

gugguggaca agggcgccuc cgcccagucc uucaucgagc ggaugaccaa cuucgacaag 1500

aaccugccca acgagaaggu gcugcccaag cacucccugc uguacgagua cuucaccgug 1560

uacaacgagc ugaccaaggu gaaguacgug accgagggca ugcggaagcc cgccuuccug 1620

uccggcgagc agaagaaggc caucguggac cugcuguuca agaccaaccg gaaggugacc 1680

gugaagcagc ugaaggagga cuacuucaag aagaucgagu gcuucgacuc cguggagauc 1740

uccggcgugg aggaccgguu caacgccucc cugggcaccu accacgaccu gcugaagauc 1800

aucaaggaca aggacuuccu ggacaacgag gagaacgagg acauccugga ggacaucgug 1860

cugacccuga cccuguucga ggaccgggag augaucgagg agcggcugaa gaccuacgcc 1920

caccuguucg acgacaaggu gaugaagcag cugaagcggc ggcgguacac cggcuggggc 1980

cggcuguccc ggaagcugau caacggcauc cgggacaagc aguccggcaa gaccauccug 2040

gacuuccuga aguccgacgg cuucgccaac cggaacuuca ugcagcugau ccacgacgac 2100

ucccugaccu ucaaggagga cauccagaag gcccaggugu ccggccaggg cgacucccug 2160

cacgagcaca ucgccaaccu ggccggcucc cccgccauca agaagggcau ccugcagacc 2220

gugaaggugg uggacgagcu ggugaaggug augggccggc acaagcccga gaacaucgug 2280

aucgagaugg cccgggagaa ccagaccacc cagaagggcc agaagaacuc ccgggagcgg 2340

augaagcgga ucgaggaggg caucaaggag cugggcuccc agauccugaa ggagcacccc 2400

guggagaaca cccagcugca gaacgagaag cuguaccugu acuaccugca gaacggccgg 2460

gacauguacg uggaccagga gcuggacauc aaccggcugu ccgacuacga cguggaccac 2520

aucgugcccc aguccuuccu gaaggacgac uccaucgaca acaaggugcu gacccggucc 2580

gacaagaacc ggggcaaguc cgacaacgug cccuccgagg agguggugaa gaagaugaag 2640

aacuacuggc ggcagcugcu gaacgccaag cugaucaccc agcggaaguu cgacaaccug 2700

accaaggccg agcggggcgg ccuguccgag cuggacaagg ccggcuucau caagcggcag 2760

cugguggaga cccggcagau caccaagcac guggcccaga uccuggacuc ccggaugaac 2820

accaaguacg acgagaacga caagcugauc cgggagguga aggugaucac ccugaagucc 2880

aagcuggugu ccgacuuccg gaaggacuuc caguucuaca aggugcggga gaucaacaac 2940

uaccaccacg cccacgacgc cuaccugaac gccguggugg gcaccgcccu gaucaagaag 3000

uaccccaagc uggaguccga guucguguac ggcgacuaca agguguacga cgugcggaag 3060

augaucgcca aguccgagca ggagaucggc aaggccaccg ccaaguacuu cuucuacucc 3120

aacaucauga acuucuucaa gaccgagauc acccuggcca acggcgagau ccggaagcgg 3180

ccccugaucg agaccaacgg cgagaccggc gagaucgugu gggacaaggg ccgggacuuc 3240

gccaccgugc ggaaggugcu guccaugccc caggugaaca ucgugaagaa gaccgaggug 3300

cagaccggcg gcuucuccaa ggaguccauc cugcccaagc ggaacuccga caagcugauc 3360

gcccggaaga aggacuggga ccccaagaag uacggcggcu ucgacucccc caccguggcc 3420

uacuccgugc uggugguggc caagguggag aagggcaagu ccaagaagcu gaaguccgug 3480

aaggagcugc ugggcaucac caucauggag cgguccuccu ucgagaagaa ccccaucgac 3540

uuccuggagg ccaagggcua caaggaggug aagaaggacc ugaucaucaa gcugcccaag 3600

uacucccugu ucgagcugga gaacggccgg aagcggaugc uggccuccgc cggcgagcug 3660

cagaagggca acgagcuggc ccugcccucc aaguacguga acuuccugua ccuggccucc 3720

cacuacgaga agcugaaggg cucccccgag gacaacgagc agaagcagcu guucguggag 3780

cagcacaagc acuaccugga cgagaucauc gagcagaucu ccgaguucuc caagcgggug 3840

auccuggccg acgccaaccu ggacaaggug cuguccgccu acaacaagca ccgggacaag 3900

cccauccggg agcaggccga gaacaucauc caccuguuca cccugaccaa ccugggcgcc 3960

cccgccgccu ucaaguacuu cgacaccacc aucgaccgga agcgguacac cuccaccaag 4020

gaggugcugg acgccacccu gauccaccag uccaucaccg gccuguacga gacccggauc 4080

gaccuguccc agcugggcgg cgacggcggc ggcuccccca agaagaagcg gaaggugucc 4140

gaguccgcca cccccgaguc cguguccggc uggcggcugu ucaagaagau cuccuga 4197

<210> 24
<211> 1398
<212> PRT
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 24

```
Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1               5                  10                  15


Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
            20                  25                  30


Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
        35                  40                  45


Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
    50                  55                  60


Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65                  70                  75                  80


Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                85                  90                  95


Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110


His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125


His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140


Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160


Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175


Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190
```

```
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380
```

```
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590
```

```
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780
```

```
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990
```

```
Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005


Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020


Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035


Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050


Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065


Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080


Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
    1085                1090                1095


Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110


Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
    1115                1120                1125


Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140


Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155


Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170
```

```
Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185

Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215

Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245

Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260

His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275

Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290

Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305

Ile Ile  His Leu Phe Thr Leu  Thr Asn Leu Gly Ala  Pro Ala Ala
    1310                1315                1320

Phe Lys  Tyr Phe Asp Thr Thr  Ile Asp Arg Lys Arg  Tyr Thr Ser
    1325                1330                1335

Thr Lys  Glu Val Leu Asp Ala  Thr Leu Ile His Gln  Ser Ile Thr
    1340                1345                1350

Gly Leu  Tyr Glu Thr Arg Ile  Asp Leu Ser Gln Leu  Gly Gly Asp
    1355                1360                1365
```

Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val Ser Glu Ser Ala
    1370   1375   1380

Thr Pro Glu Ser Val Ser Gly Trp Arg Leu Phe Lys Lys Ile Ser
    1385   1390   1395

<210> 25
<211> 589
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 25
gggagaccca agcuggcuag cucccgcagu cggcguccag cggcucugcu uguucgugug 60

ugugucguug caggccuuau ucggauccgc caccauggga ccgaagaaga agagaaaggu 120

cggaggagga agcacaaacc ugucggacau caucgaaaag gaaacaggaa agcagcuggu 180

cauccaggaa ucgauccuga ugcugccgga agaagucgaa gaagucaucg gaaacaagcc 240

ggaaucggac auccuggucc acacagcaua cgacgaaucg acagacgaaa acgucaugcu 300

gcugacaucg gacgcaccgg aauacaagcc gugggcacug gucauccagg acucgaacgg 360

agaaaacaag aucaagaugc ugugauaguc uagacaucac auuuaaaagc aucucagccu 420

accaugagaa uaagagaaag aaaaugaaga ucaauagcuu auucaucucu uuuucuuuuu 480

cguuggugua aagccaacac ccugucuaaa aaacauaaau uucuuuaauc auuuugccuc 540

uuuucucugu gcuucaauua auaaaaaaug gaaagaaccu cgagucuag   589

<210> 26
<211> 291
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 26
augggaccga agaagaagag aaaggucgga ggaggaagca caaaccuguc ggacaucauc 60

gaaaaggaaa caggaaagca gcuggucauc caggaaucga uccugaugcu gccggaagaa 120

gucgaagaag ucaucggaaa caagccggaa ucggacaucc ugguccacac agcauacgac 180

gaaucgacag acgaaaacgu caugcugcug acaucggacg caccggaaua caagccgugg 240

gcacugguca uccaggacuc gaacggagaa aacaagauca agaugcugug a 291

<210> 27
<211> 107
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 27

Met Thr Asn Leu Ser Asp Ile Ile Glu Lys Glu Thr Gly Lys Gln Leu
1 5 10 15

Val Ile Gln Glu Ser Ile Leu Met Leu Pro Glu Glu Val Glu Glu Val
20 25 30

Ile Gly Asn Lys Pro Glu Ser Asp Ile Leu Val His Thr Ala Tyr Asp
35 40 45

Glu Ser Thr Asp Glu Asn Val Met Leu Leu Thr Ser Asp Ala Pro Glu
50 55 60

Tyr Lys Pro Trp Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn Lys
65 70 75 80

Ile Lys Met Leu Ser Gly Gly Ser Lys Arg Thr Ala Asp Gly Ser Glu
85 90 95

Phe Glu Ser Pro Lys Lys Lys Arg Lys Val Glu
100 105

<210> 28
<211> 5933

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 28
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc gagacccccg 660

gcaccuccga guccgccacc cccgaguccg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

```
agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg     2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg     2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga     3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg     3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu     3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc     3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca     3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu     3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga     3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg     3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg     3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg     3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu     3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg     3660

uggugggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg     3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg     3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc     3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga     3900

ucgugugggа caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg     3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc     4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg     4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg     4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu     4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga     4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc     4320
```

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac uccggcggcu 4800

ccggcggcuc cggcggcucc accaaccugu ccgacaucau cgagaaggag accggcaagc 4860

agcuggugau ccaggagucc auccugaugc ugcccgagga gguggaggag gugaucggca 4920

acaagcccga guccgacauc cuggugcaca ccgccuacga cgaguccacc gacgagaacg 4980

ugaugcugcu gaccuccgac gcccccgagu acaagcccug ggcccuggug auccaggacu 5040

ccaacggcga gaacaagauc aagaugcugu ccggcggcuc cggcggcucc ggcggcucca 5100

ccaaccuguc cgacaucauc gagaaggaga ccggcaagca gcuggugauc caggagucca 5160

uccugaugcu gcccgaggag guggaggagg ugaucggcaa caagcccgag uccgacaucc 5220

uggugcacac cgccuacgac gaguccaccg acgagaacgu gaugcugcug accuccgacg 5280

cccccgagua caagcccugg gcccugguga uccaggacuc caacggcgag aacaagauca 5340

agaugcuguc cggcggcucc aagcggaccg ccgacggcuc cgaguucgag cccaagaaga 5400

agcggaaggu gugauagcua gcaccagccu caagaacacc cgaauggagu cucuaagcua 5460

cauaauacca acuuacacuu uacaaaaugu ugucccccaa aauguagcca uucguaucug 5520

cuccuaauaa aaagaaaguu ucuucacauu cucucgagaa aaaaaaaaaa uggaaaaaaa 5580

aaaaacggaa aaaaaaaaaa gguaaaaaaa aaaaauauaa aaaaaaaaaa cauaaaaaaa 5640

aaaaacgaaa aaaaaaaaac guaaaaaaaa aaaacucaaa aaaaaaaaag auaaaaaaaa 5700

aaaaccuaaa aaaaaaaaau guaaaaaaaa aaaagggaaa aaaaaaaaac gcaaaaaaaa 5760

aaaacacaaa aaaaaaaaau gcaaaaaaaa aaaaucgaaa aaaaaaaaau cuaaaaaaaa 5820

aaaacgaaaa aaaaaaaacc caaaaaaaaa aaagacaaaa aaaaaaaaua gaaaaaaaaa 5880

aaaguuaaaa aaaaaaaacu gaaaaaaaaa aaauuuaaaa aaaaaaaauc uag 5933

<210> 29
<211> 5373
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 29
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacucc 600

ggcuccgaga cccccggcac cuccgagucc gccacccccg aguccgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggauccccu acuacguggg ccccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau cacccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccuggggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacuccg gcggcuccgg cggcuccggc ggcuccacca accuguccga caucaucgag 4800

aaggagaccg gcaagcagcu ggugauccag gaguccaucc ugaugcugcc cgaggaggug 4860

gaggagguga ucggcaacaa gcccgagucc gacauccugg ugcacaccgc cuacgacgag 4920

uccaccgacg agaacgugau gcugcugacc uccgacgccc ccgaguacaa gcccugggcc 4980

cuggugaucc aggacuccaa cggcgagaac aagaucaaga ugcuguccgg cggcuccggc 5040

ggcuccggcg gcuccaccaa ccuguccgac aucaucgaga aggagaccgg caagcagcug 5100

gugauccagg aguccauccu gaugcugccc gaggaggugg aggaggugau cggcaacaag 5160

cccgaguccg acauccuggu gcacaccgcc uacgacgagu ccaccgacga gaacgugaug 5220

cugcugaccu ccgacgcccc cgaguacaag cccugggccc uggugaucca ggacuccaac 5280

ggcgagaaca agaucaagau gcuguccggc ggcuccaagc ggaccgccga cggcuccgag 5340

uucgagccca agaagaagcg gaagguguga uag 5373

<210> 30
<211> 1789
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 30

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1   5   10   15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20   25   30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35   40   45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50   55   60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65   70   75   80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85   90   95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100   105   110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115   120   125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130   135   140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145   150   155   160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165   170   175

```
Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser Glu Thr Pro Gly Thr Ser
        195                 200                 205

Glu Ser Ala Thr Pro Glu Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
    370                 375                 380
```

```
Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575
```

```
Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
    770                 775                 780
```

```
Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800


Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815


Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
            820                 825                 830


Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845


Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860


Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880


Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895


Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
            900                 905                 910


Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925


Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940


Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960


Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
                965                 970                 975
```

```
Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1160                1165                1170
```

```
Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350
```

```
Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500

Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1535                1540                1545
```

```
Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560

Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575

Leu Gly  Gly Asp Ser Gly Gly  Ser Gly Gly Ser Gly  Gly Ser Thr
    1580                1585                1590

Asn Leu  Ser Asp Ile Ile Glu  Lys Glu Thr Gly Lys  Gln Leu Val
    1595                1600                1605

Ile Gln  Glu Ser Ile Leu Met  Leu Pro Glu Glu Val  Glu Glu Val
    1610                1615                1620

Ile Gly  Asn Lys Pro Glu Ser  Asp Ile Leu Val His  Thr Ala Tyr
    1625                1630                1635

Asp Glu  Ser Thr Asp Glu Asn  Val Met Leu Leu Thr  Ser Asp Ala
    1640                1645                1650

Pro Glu  Tyr Lys Pro Trp Ala  Leu Val Ile Gln Asp  Ser Asn Gly
    1655                1660                1665

Glu Asn  Lys Ile Lys Met Leu  Ser Gly Gly Ser Gly  Gly Ser Gly
    1670                1675                1680

Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys
    1685                1690                1695

Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val
    1700                1705                1710

Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser Asp Ile  Leu Val His
    1715                1720                1725
```

Thr Ala Tyr Asp Glu Ser Thr Asp Glu Asn Val Met Leu Leu Thr
    1730   1735   1740

Ser Asp Ala Pro Glu Tyr Lys Pro Trp Ala Leu Val Ile Gln Asp
    1745   1750   1755

Ser Asn Gly Glu Asn Lys Ile Lys Met Leu Ser Gly Gly Ser Lys
    1760   1765   1770

Arg Thr Ala Asp Gly Ser Glu Phe Glu Pro Lys Lys Lys Arg Lys
    1775   1780   1785

Val

<210> 31
<211> 6125
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 31
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccaugaag cggaccgccg   60

acggcuccga guucgagucc cccaagaaga agcggaaggu guccuccgag accggccccg   120

uggccgugga ccccacccug cggcggcgga ucgagcccca cgaguucgag guguucuucg   180

acccccggga gcugcggaag gagaccugcc ugcuguacga gaucaacugg ggcggccggc   240

acuccaucug gcggcacacc ucccagaaca ccaacaagca cguggaggug aacuucaucg   300

agaaguucac caccgagcgg uacuucugcc ccaacacccg gugcuccauc accugguucc   360

uguccugguc ccccugcggc gagugcuccc gggccaucac cgaguuccug ucccgguacc   420

cccacgugac ccuguucauc uacaucgccc ggcuguacca ccacgccgac ccccggaacc   480

ggcagggccu gcgggaccug aucuccuccg gcgugaccau ccagaucaug accgagcagg   540

aguccggcua cugcuggcgg aacuucguga acuacucccc cuccaacgag gcccacuggc   600

cccgguaccc ccaccugugg gugcggcugu acgugcugga gcuguacugc aucauccugg   660

gccugccccc cugccugaac auccugcggc ggaagcagcc ccagcugacc uucuucacca 720

ucgcccugca guccugccac uaccagcggc ugccccccca cauccugugg gccaccggcc 780

ugaaguccgg cggcuccucc ggcggcuccu ccggcuccga gacccccggc accuccgagu 840

ccgccacccc cgaguccucc ggcggcuccu ccggcggcuc cgacaagaag uacuccaucg 900

gccuggccau cggcaccaac uccgugggcu gggccgugau caccgacgag uacaaggugc 960

ccuccaagaa guucaaggug cugggcaaca ccgaccggca cuccaucaag aagaaccuga 1020

ucggcgcccu gcuguucgac uccggcgaga ccgccgaggc cacccggcug aagcggaccg 1080

cccggcggcg guacacccgg cggaagaacc ggaucugcua ccugcaggag aucuucucca 1140

acgagauggc caagguggac gacuccuucu uccaccggcu ggaggagucc uuccuggugg 1200

aggaggacaa gaagcacgag cggcacccca ucuucggcaa caucguggac gagguggccu 1260

accacgagaa guaccccacc aucuaccacc ugcggaagaa gcugguggac uccaccgaca 1320

aggccgaccu gcggcugauc uaccuggccc uggcccacau gaucaaguuc cggggccacu 1380

uccugaucga gggcgaccug aaccccgaca acuccgacgu ggacaagcug uucauccagc 1440

uggugcagac cuacaaccag cuguucgagg agaaccccau caacgccucc ggcguggacg 1500

ccaaggccau ccuguccgcc cggcugucca agucccggcg gcuggagaac cugaucgccc 1560

agcugcccgg cgagaagaag aacggccugu ucggcaaccu gaucgcccug ucccugggcc 1620

ugacccccaa cuucaagucc aacuucgacc uggccgagga cgccaagcug cagcugucca 1680

aggacaccua cgacgacgac cuggacaacc ugcuggccca gaucggcgac caguacgccg 1740

accuguuccu ggccgccaag aaccuguccg acgccauccu gcuguccgac auccugcggg 1800

ugaacaccga gaucaccaag gccccccugu ccgccuccau gaucaagcgg uacgacgagc 1860

accaccagga ccugacccug cugaaggccc uggugcggca gcagcugccc gagaaguaca 1920

aggagaucuu cuucgaccag uccaagaacg gcuacgccgg cuacaucgac ggcggcgccu 1980

cccaggagga guucuacaag uucaucaagc ccauccugga gaagauggac ggcaccgagg 2040

agcugcuggu gaagcugaac cgggaggacc ugcugcggaa gcagcggacc uucgacaacg 2100

gcuccauccc ccaccagauc caccugggcg agcugcacgc cauccugcgg cggcaggagg 2160

acuucuaccc cuuccugaag gacaaccggg agaagaucga gaagauccug accuuccgga 2220

uccccuacua cgugggcccc cuggcccggg gcaacucccg guucgccugg augacccgga 2280

aguccgagga gaccaucacc cccuggaacu ucgaggaggu gguggacaag ggcgccuccg 2340

cccaguccuu caucgagcgg augaccaacu ucgacaagaa ccugcccaac gagaaggugc 2400

ugcccaagca cucccugcug uacgaguacu ucaccgugua caacgagcug accaagguga 2460

aguacgugac cgagggcaug cggaagcccg ccuuccuguc cggcgagcag aagaaggcca 2520

ucguggaccu gcuguucaag accaaccgga aggugaccgu gaagcagcug aaggaggacu 2580

acuucaagaa gaucgagugc uucgacuccg uggagaucuc cggcguggag gaccgguuca 2640

acgccucccu gggcaccuac cacgaccugc ugaagaucau caaggacaag gacuuccugg 2700

acaacgagga gaacgaggac auccuggagg acaucgugcu gacccugacc cuguucgagg 2760

accgggagau gaucgaggag cggcugaaga ccuacgccca ccuguucgac gacaagguga 2820

ugaagcagcu gaagcggcgg cgguacaccg gcuggggccg gcugucccgg aagcugauca 2880

acggcauccg ggacaagcag uccggcaaga ccauccugga cuuccugaag uccgacggcu 2940

ucgccaaccg gaacuucaug cagcugaucc acgacgacuc ccugaccuuc aaggaggaca 3000

uccagaaggc ccaggugucc ggccagggcg acucccugca cgagcacauc gccaaccugg 3060

ccggcucccc cgccaucaag aagggcaucc ugcagaccgu gaagguggug gacgagcugg 3120

ugaaggugau gggccggcac aagcccgaga acaucgugau cgagauggcc cgggagaacc 3180

agaccaccca gaagggccag aagaacuccc gggagcggau gaagcggauc gaggagggca 3240

ucaaggagcu gggcucccag auccugaagg agcaccccgu ggagaacacc cagcugcaga 3300

acgagaagcu guaccuguac uaccugcaga acggccggga cauguacgug gaccaggagc 3360

uggacaucaa ccggcugucc gacuacgacg uggaccacau cgugccccag uccuuccuga 3420

aggacgacuc caucgacaac aaggugcuga cccgguccga caagaaccgg ggcaaguccg 3480

acaacgugcc cuccgaggag guggugaaga agaugaagaa cuacuggcgg cagcugcuga 3540

acgccaagcu gaucacccag cggaaguucg acaaccugac caaggccgag cggggcggcc 3600

uguccgagcu ggacaaggcc ggcuucauca agcggcagcu gguggagacc cggcagauca 3660

ccaagcacgu ggcccagauc cuggacuccc ggaugaacac caaguacgac gagaacgaca 3720

agcugauccg ggaggugaag gugaucaccc ugaaguccaa gcuggugucc gacuuccgga 3780

aggacuucca guucuacaag gugcgggaga ucaacaacua ccaccacgcc cacgacgccu 3840

accugaacgc cgugggugggc accgcccuga ucaagaagua ccccaagcug gaguccgagu 3900

ucguguacgg cgacuacaag guguacgacg ugcggaagau gaucgccaag uccgagcagg 3960

agaucggcaa ggccaccgcc aaguacuucu ucuacuccaa caucaugaac uucuucaaga 4020

ccgagaucac ccuggccaac ggcgagaucc ggaagcggcc ccugaucgag accaacggcg 4080

agaccggcga gaucgugugg gacaagggcc gggacuucgc caccgugcgg aaggugcugu 4140

ccaugcccca ggugaacauc gugaagaaga ccgaggugca gaccggcggc uucuccaagg 4200

aguccauccu gcccaagcgg aacuccgaca agcugaucgc ccggaagaag gacugggacc 4260

ccaagaagua cggcggcuuc gacuccccca ccguggccua cuccgugcug gugguggcca 4320

agguggagaa gggcaagucc aagaagcuga aguccgugaa ggagcugcug ggcaucacca 4380

ucauggagcg guccuccuuc gagaagaacc ccaucgacuu ccuggaggcc aagggcuaca 4440

aggaggugaa gaaggaccug aucaucaagc ugcccaagua cucccuguuc gagcuggaga 4500

acggccggaa gcggaugcug gccuccgccg gcgagcugca gaagggcaac gagcuggccc 4560

ugcccuccaa guacgugaac uuccuguacc uggccuccca cuacgagaag cugaagggcu 4620

cccccgagga caacgagcag aagcagcugu ucguggagca gcacaagcac uaccuggacg 4680

agaucaucga gcagaucucc gaguucucca agcgggugau ccuggccgac gccaaccugg 4740

acaaggugcu guccgccuac aacaagcacc gggacaagcc cauccgggag caggccgaga 4800

acaucaucca ccuguucacc cugaccaacc ugggcgcccc cgccgccuuc aaguacuucg 4860

acaccaccau cgaccggaag cgguacaccu ccaccaagga ggugcuggac gccacccuga 4920

uccaccaguc caucaccggc cuguacgaga cccggaucga ccugucccag cuggggcggcg 4980

acuccggcgg cuccggcggc uccggcggcu ccaccaaccu guccgacauc aucgagaagg 5040

agaccggcaa gcagcuggug auccaggagu ccauccugau gcugcccgag gagguggagg 5100

aggugaucgg caacaagccc gaguccgaca uccuggugca caccgccuac gacgagucca 5160

ccgacgagaa cgugaugcug cugaccuccg acgcccccga guacaagccc ugggcccugg 5220

ugauccagga cuccaacggc gagaacaaga ucaagaugcu guccggcggc uccggcggcu 5280

ccggcggcuc caccaaccug uccgacauca ucgagaagga gaccggcaag cagcugguga 5340

uccaggaguc cauccugaug cugcccgagg agguggagga ggugaucggc aacaagcccg 5400

aguccgacau ccuggugcac accgccuacg acgaguccac cgacgagaac gugaugcugc 5460

ugaccuccga cgcccccgag uacaagcccu gggcccuggu gauccaggac uccaacggcg 5520

agaacaagau caagaugcug uccggcggcu ccaagcggac cgccgacggc uccgaguucg 5580

agcccaagaa gaagcggaag gugugauagc uagcaccagc cucaagaaca cccgaaugga 5640

gucucuaagc uacauaauac caacuuacac uuuacaaaau guugucccc aaaauguagc 5700

cauucguauc ugcuccuaau aaaaagaaag uuucuucaca uucucucgag aaaaaaaaaa 5760

aauggaaaaa aaaaaaacgg aaaaaaaaaa aagguaaaaa aaaaaaauau aaaaaaaaaa 5820

aacauaaaaa aaaaaaacga aaaaaaaaaa acguaaaaaa aaaaaacuca aaaaaaaaaa 5880

agauaaaaaa aaaaaaccua aaaaaaaaaa auguaaaaaa aaaaaaggga aaaaaaaaaa 5940

acgcaaaaaa aaaaaacaca aaaaaaaaaa augcaaaaaa aaaaaaucga aaaaaaaaaa 6000

aucuaaaaaa aaaaaacgaa aaaaaaaaaa cccaaaaaaa aaaaagacaa aaaaaaaaaa 6060

uagaaaaaaa aaaaaguuaa aaaaaaaaaa cugaaaaaaa aaaaauuuaa aaaaaaaaaa 6120

ucuag 6125

<210> 32
<211> 5565
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 32
augaagcgga ccgccgacgg cuccgaguuc gaguccccca agaagaagcg gaaggugucc 60

uccgagaccg gccccguggc cguggacccc acccugcggc ggcggaucga gccccacgag 120

uucgaggugu ucuucgaccc ccgggagcug cggaaggaga ccugccugcu guacgagauc 180

aacuggggcg gccggcacuc caucuggcgg cacaccuccc agaacaccaa caagcacgug 240

gaggugaacu ucaucgagaa guucaccacc gagcgguacu ucugccccaa cacccggugc 300

uccaucaccu gguuccuguc cuggucccc ugcggcgagu gcucccgggc caucaccgag 360

uuccuguccc gguaccccca cgugacccug uucaucuaca ucgcccggcu guaccaccac 420

gccgaccccc ggaaccggca gggccugcgg gaccugaucu ccuccggcgu gaccauccag 480

aucaugaccg agcaggaguc cggcuacugc uggcggaacu ucgugaacua cucccccucc 540

aacgaggccc acuggccccg guacccccac cuguggggugc ggcuguacgu gcuggagcug 600

uacugcauca uccugggccu gccccccugc cugaacaucc ugcggcggaa gcagccccag 660

cugaccuucu ucaccaucgc ccugcagucc ugccacuacc agcggcugcc cccccacauc 720

cuguggggcca ccggccugaa guccggcggc uccuccggcg gcuccuccgg cuccgagacc 780

cccggcaccu ccgaguccgc caccccccgag uccuccggcg gcuccuccgg cggcuccgac 840

aagaaguacu ccaucggccu ggccaucggc accaacuccg ugggcuggggc cgugaucacc 900

gacgaguaca aggugcccuc caagaaguuc aaggugcugg gcaacaccga ccggcacucc 960

aucaagaaga accugaucgg cgcccugcug uucgacuccg gcgagaccgc cgaggccacc 1020

cggcugaagc ggaccgcccg gcggcgguac acccggcgga agaaccggau cugcuaccug 1080

caggagaucu ucuccaacga gauggccaag guggacgacu ccuucuucca ccggcuggag 1140

gaguccuucc ugguggagga ggacaagaag cacgagcggc accccaucuu cggcaacauc 1200

guggacgagg uggccuacca cgagaaguac cccaccaucu accaccugcg gaagaagcug 1260

guggacucca ccgacaaggc cgaccugcgg cugaucuacc uggcccuggc ccacaugauc 1320

aaguuccggg gccacuuccu gaucgagggc gaccugaacc ccgacaacuc cgacguggac 1380

aagcuguuca uccagcuggu gcagaccuac aaccagcugu ucgaggagaa ccccaucaac 1440

gccuccggcg uggacgccaa ggccauccug uccgcccggc uguccaaguc ccggcggcug 1500

gagaaccuga ucgcccagcu gcccggcgag aagaagaacg gccuguucgg caaccugauc 1560

gcccuguccc ugggccugac ccccaacuuc aaguccaacu ucgaccuggc cgaggacgcc 1620

aagcugcagc uguccaagga caccuacgac gacgaccugg acaaccugcu ggcccagauc 1680

ggcgaccagu acgccgaccu guuccuggcc gccaagaacc uguccgacgc cauccugcug 1740

uccgacaucc ugcgggugaa caccgagauc accaaggccc cccuguccgc cuccaugauc 1800

aagcgguacg acgagcacca ccaggaccug acccugcuga aggcccuggu gcggcagcag 1860

cugcccgaga aguacaagga gaucuucuuc gaccagucca agaacggcua cgccggcuac 1920

aucgacggcg gcgccuccca ggaggaguuc uacaaguuca ucaagcccau ccuggagaag 1980

auggacggca ccgaggagcu gcuggugaag cugaaccggg aggaccugcu gcggaagcag 2040

cggaccuucg acaacggcuc caucccccac cagauccacc ugggcgagcu gcacgccauc 2100

cugcggcggc aggaggacuu cuaccccuuc cugaaggaca accgggagaa gaucgagaag 2160

auccugaccu uccggauccc cuacuacgug ggcccccugg cccggggcaa cucccgguuc 2220

gccuggauga cccggaaguc cgaggagacc aucacccccu ggaacuucga ggagguggug 2280

gacaagggcg ccuccgccca guccuucauc gagcggauga ccaacuucga caagaaccug 2340

cccaacgaga aggugcugcc caagcacucc cugcuguacg aguacuucac cguguacaac 2400

gagcugacca aggugaagua cgugaccgag ggcaugcgga agcccgccuu ccuguccggc 2460

gagcagaaga aggccaucgu ggaccugcug uucaagacca accggaaggu gaccgugaag 2520

cagcugaagg aggacuacuu caagaagauc gagugcuucg acuccgugga gaucuccggc 2580

guggaggacc gguucaacgc cucccugggc accuaccacg accugcugaa gaucaucaag 2640

gacaaggacu uccuggacaa cgaggagaac gaggacaucc uggaggacau cgugcugacc 2700

cugacccugu ucgaggaccg ggagaugauc gaggagcggc ugaagaccua cgcccaccug 2760

uucgacgaca aggugaugaa gcagcugaag cggcggcggu acaccggcug gggccggcug 2820

ucccggaagc ugaucaacgg cauccgggac aagcaguccg gcaagaccau ccuggacuuc 2880

cugaaguccg acggcuucgc caaccggaac uucaugcagc ugauccacga cgacucccug 2940

accuucaagg aggacaucca gaaggcccag guguccggcc agggcgacuc ccugcacgag 3000

cacaucgcca accuggccgg cucccccgcc aucaagaagg gcauccugca gaccgugaag 3060

gugguggacg agcuggugaa ggugaugggc cggcacaagc ccgagaacau cgugaucgag 3120

auggcccggg agaaccagac cacccagaag ggccagaaga acucccggga gcggaugaag 3180

cggaucgagg agggcaucaa ggagcugggc ucccagaucc ugaaggagca ccccguggag 3240

aacacccagc ugcagaacga gaagcuguac cuguacuacc ugcagaacgg ccgggacaug 3300

uacguggacc aggagcugga caucaaccgg cuguccgacu acgacgugga ccacaucgug 3360

ccccaguccu uccugaagga cgacuccauc gacaacaagg ugcugacccg guccgacaag 3420

aaccggggca aguccgacaa cgugcccucc gaggaggugg ugaagaagau gaagaacuac 3480

uggcggcagc ugcugaacgc caagcugauc acccagcgga aguucgacaa ccugaccaag 3540

gccgagcggg gcggccuguc cgagcuggac aaggccggcu ucaucaagcg gcagcuggug 3600

gagacccggc agaucaccaa gcacguggcc cagauccugg acucccggau gaacaccaag 3660

uacgacgaga acgacaagcu gauccgggag gugaagguga ucacccugaa guccaagcug 3720

guguccgacu uccggaagga cuuccaguuc uacaaggugc gggagaucaa caacuaccac 3780

cacgcccacg acgccuaccu gaacgccgug gugggcaccg cccugaucaa gaaguacccc 3840

aagcuggagu ccgaguucgu guacggcgac uacaaggugu acgacgugcg gaagaugauc 3900

gccaaguccg agcaggagau cggcaaggcc accgccaagu acuucuucua cuccaacauc 3960

augaacuucu ucaagaccga gaucacccug gccaacggcg agauccggaa gcggccccug 4020

aucgagacca acggcgagac cggcgagauc gugugggaca agggccggga cuucgccacc 4080

gugcggaagg ugcuguccau gccccaggug aacaucguga agaagaccga ggugcagacc 4140

ggcggcuucu ccaaggaguc cauccugccc aagcggaacu ccgacaagcu gaucgcccgg 4200

aagaaggacu gggaccccaa gaaguacggc ggcuucgacu cccccaccgu ggccuacucc 4260

gugcuggugg uggccaaggu ggagaagggc aaguccaaga agcugaaguc cgugaaggag 4320

cugcugggca ucaccaucau ggagcggucc uccuucgaga agaaccccau cgacuuccug 4380

gaggccaagg gcuacaagga ggugaagaag gaccugauca ucaagcugcc caaguacucc 4440

cuguucgagc uggagaacgg ccggaagcgg augcuggccu ccgccggcga gcugcagaag 4500

ggcaacgagc uggcccugcc cuccaaguac gugaacuucc uguaccuggc cucccacuac 4560

gagaagcuga agggcucccc cgaggacaac gagcagaagc agcuguucgu ggagcagcac 4620

aagcacuacc uggacgagau caucgagcag aucuccgagu ucuccaagcg ggugauccug 4680

gccgacgcca accuggacaa ggugcugucc gccuacaaca agcaccggga caagcccauc 4740

cgggagcagg ccgagaacau cauccaccug uucacccuga ccaaccuggg cgcccccgcc 4800

gccuucaagu acuucgacac caccaucgac cggaagcggu acaccuccac caaggaggug 4860

cuggacgcca cccugaucca ccaguccauc accggccugu acgagacccg gaucgaccug 4920

ucccagcugg gcggcgacuc cggcggcucc ggcggcuccg gcggcuccac caaccugucc 4980

gacaucaucg agaaggagac cggcaagcag cuggugaucc aggaguccau ccugaugcug 5040

cccgaggagg uggaggaggu gaucggcaac aagcccgagu ccgacauccu ggugcacacc 5100

gccuacgacg aguccaccga cgagaacgug augcugcuga ccuccgacgc ccccgaguac 5160

aagcccuggg cccuggugau ccaggacucc aacggcgaga acaagaucaa gaugcugucc 5220

ggcggcuccg gcggcuccgg cggcuccacc aaccuguccg acaucaucga gaaggagacc 5280

ggcaagcagc uggugaucca ggaguccauc cugaugcugc ccgaggaggu ggaggaggug 5340

aucggcaaca agcccgaguc cgacauccug gugcacaccg ccuacgacga guccaccgac 5400

gagaacguga ugcugcugac cuccgacgcc cccgaguaca agcccuggggc ccuggugauc 5460

caggacucca acggcgagaa caagaucaag augcuguccg gcggcuccaa gcggaccgcc 5520

gacggcuccg aguucgagcc caagaagaag cggaaggugu gauag 5565

<210> 33
<211> 1853
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 33

Met Lys Arg Thr Ala Asp Gly Ser Glu Phe Glu Ser Pro Lys Lys Lys
1 5 10 15

Arg Lys Val Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu
20 25 30

```
Arg Arg Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg
        35                  40                  45

Glu Leu Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly
    50                  55                  60

Arg His Ser Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val
65                  70                  75                  80

Glu Val Asn Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro
                85                  90                  95

Asn Thr Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly
            100                 105                 110

Glu Cys Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val
        115                 120                 125

Thr Leu Phe Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg
    130                 135                 140

Asn Arg Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln
145                 150                 155                 160

Ile Met Thr Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn
                165                 170                 175

Tyr Ser Pro Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp
            180                 185                 190

Val Arg Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro
        195                 200                 205

Pro Cys Leu Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe
    210                 215                 220

Thr Ile Ala Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile
225                 230                 235                 240
```

```
Leu Trp Ala Thr Gly Leu Lys Ser Gly Gly Ser Ser Gly Gly Ser Ser
               245                 250                 255

Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Ser
            260                 265                 270

Gly Gly Ser Ser Gly Gly Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala
        275                 280                 285

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
    290                 295                 300

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
305                 310                 315                 320

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
               325                 330                 335

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
            340                 345                 350

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
        355                 360                 365

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
    370                 375                 380

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
385                 390                 395                 400

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
               405                 410                 415

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
            420                 425                 430
```

```
Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
        435                 440                 445

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
    450                 455                 460

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
465                 470                 475                 480

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
                485                 490                 495

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
            500                 505                 510

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
        515                 520                 525

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
    530                 535                 540

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
545                 550                 555                 560

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
                565                 570                 575

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
            580                 585                 590

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
        595                 600                 605

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
    610                 615                 620

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
625                 630                 635                 640
```

```
Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
                645                 650                 655

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
            660                 665                 670

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
        675                 680                 685

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
    690                 695                 700

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
705                 710                 715                 720

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
                725                 730                 735

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
            740                 745                 750

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
        755                 760                 765

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
    770                 775                 780

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
785                 790                 795                 800

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
                805                 810                 815

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
            820                 825                 830
```

```
Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
        835                 840                 845

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
    850                 855                 860

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
865                 870                 875                 880

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
                885                 890                 895

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
            900                 905                 910

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
        915                 920                 925

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
    930                 935                 940

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
945                 950                 955                 960

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
                965                 970                 975

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
            980                 985                 990

Gly Gln Gly Asp Ser Leu His Glu  His Ile Ala Asn Leu  Ala Gly Ser
        995                 1000                1005

Pro Ala  Ile Lys Lys Gly Ile  Leu Gln Thr Val Lys  Val Val Asp
    1010                1015                1020

Glu Leu  Val Lys Val Met Gly  Arg His Lys Pro Glu  Asn Ile Val
    1025                1030                1035
```

```
Ile Glu  Met Ala Arg Glu Asn  Gln Thr Thr Gln Lys  Gly Gln Lys
    1040                1045                1050

Asn Ser  Arg Glu Arg Met Lys  Arg Ile Glu Glu Gly  Ile Lys Glu
    1055                1060                1065

Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro Val Glu  Asn Thr Gln
    1070                1075                1080

Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg
    1085                1090                1095

Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp
    1100                1105                1110

Tyr Asp  Val Asp His Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp
    1115                1120                1125

Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly
    1130                1135                1140

Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val Val Lys  Lys Met Lys
    1145                1150                1155

Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg
    1160                1165                1170

Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu
    1175                1180                1185

Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg
    1190                1195                1200

Gln Ile  Thr Lys His Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn
    1205                1210                1215
```

```
Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val
    1220                 1225                 1230

Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe
    1235                 1240                 1245

Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn Tyr His  His Ala His
    1250                 1255                 1260

Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys
    1265                 1270                 1275

Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val
    1280                 1285                 1290

Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly
    1295                 1300                 1305

Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe
    1310                 1315                 1320

Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg
    1325                 1330                 1335

Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp
    1340                 1345                 1350

Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys Val Leu  Ser Met Pro
    1355                 1360                 1365

Gln Val  Asn Ile Val Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe
    1370                 1375                 1380

Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile
    1385                 1390                 1395

Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp
    1400                 1405                 1410
```

```
Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val Val Ala  Lys Val Glu
    1415                1420                1425

Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly
    1430                1435                1440

Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp
    1445                1450                1455

Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile
    1460                1465                1470

Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg
    1475                1480                1485

Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu
    1490                1495                1500

Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser
    1505                1510                1515

His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys
    1520                1525                1530

Gln Leu  Phe Val Glu Gln His  Lys His Tyr Leu Asp  Glu Ile Ile
    1535                1540                1545

Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala
    1550                1555                1560

Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys
    1565                1570                1575

Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu
    1580                1585                1590
```

```
Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr
    1595                1600                1605

Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala
    1610                1615                1620

Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile
    1625                1630                1635

Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly Gly Ser  Gly Gly Ser
    1640                1645                1650

Gly Gly  Ser Thr Asn Leu Ser  Asp Ile Ile Glu Lys  Glu Thr Gly
    1655                1660                1665

Lys Gln  Leu Val Ile Gln Glu  Ser Ile Leu Met Leu  Pro Glu Glu
    1670                1675                1680

Val Glu  Glu Val Ile Gly Asn  Lys Pro Glu Ser Asp  Ile Leu Val
    1685                1690                1695

His Thr  Ala Tyr Asp Glu Ser  Thr Asp Glu Asn Val  Met Leu Leu
    1700                1705                1710

Thr Ser  Asp Ala Pro Glu Tyr  Lys Pro Trp Ala Leu  Val Ile Gln
    1715                1720                1725

Asp Ser  Asn Gly Glu Asn Lys  Ile Lys Met Leu Ser  Gly Gly Ser
    1730                1735                1740

Gly Gly  Ser Gly Gly Ser Thr  Asn Leu Ser Asp Ile  Ile Glu Lys
    1745                1750                1755

Glu Thr  Gly Lys Gln Leu Val  Ile Gln Glu Ser Ile  Leu Met Leu
    1760                1765                1770

Pro Glu  Glu Val Glu Glu Val  Ile Gly Asn Lys Pro  Glu Ser Asp
    1775                1780                1785
```

Ile Leu Val His Thr Ala Tyr Asp Glu Ser Thr Asp Glu Asn Val
    1790   1795   1800

Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr Lys Pro Trp Ala Leu
    1805   1810   1815

Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile Lys Met Leu Ser
    1820   1825   1830

Gly Gly Ser Lys Arg Thr Ala Asp Gly Ser Glu Phe Glu Pro Lys
    1835   1840   1845

Lys Lys Arg Lys Val
    1850

<210> 34
<211> 887
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 34
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccaugacc aaccuguccg   60

acaucaucga gaaggagacc ggcaagcagc uggugaucca ggaguccauc cugaugcugc   120

ccgaggaggu ggaggaggug aucggcaaca agcccgaguc cgacauccug gugcacaccg   180

ccuacgacga guccaccgac gagaacguga ugcugcugac cuccgacgcc cccgaguaca   240

agcccugggc ccuggugauc caggacucca acggcgagaa caagaucaag augcuguccg   300

gcggcuccaa gcggaccgcc gacggcuccg aguucgaguc ccccaagaag aagcggaagg   360

uggagugaua gcuagcacca gccucaagaa cacccgaaug gagucucuaa gcuacauaau   420

accaacuuac acuuuacaaa auguuguccc ccaaaaugua gccauucgua ucugcuccua   480

auaaaaagaa aguuucuuca cauucucucg agaaaaaaaa aaaauggaaa aaaaaaaaac   540

ggaaaaaaaa aaaagguaaa aaaaaaaaau auaaaaaaaa aaaacauaaa aaaaaaaaac   600

gaaaaaaaaa aaacguaaaa aaaaaaaacu caaaaaaaaa aaagauaaaa aaaaaaaacc 660

uaaaaaaaaa aaauguaaaa aaaaaaaagg gaaaaaaaaa aaacgcaaaa aaaaaaaaca 720

caaaaaaaaa aaaugcaaaa aaaaaaaauc gaaaaaaaaa aaaucuaaaa aaaaaaaacg 780

aaaaaaaaaa aacccaaaaa aaaaaaagac aaaaaaaaaa aauagaaaaa aaaaaaaguu 840

aaaaaaaaaa aacugaaaaa aaaaaaauuu aaaaaaaaaa aaucuag 887

<210> 35
<211> 324
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 35
augaccaacc uguccgacau caucgagaag gagaccggca agcagcuggu gauccaggag 60

uccauccuga ugcugcccga ggagguggag gaggugaucg gcaacaagcc cgaguccgac 120

auccuggugc acaccgccua cgacgagucc accgacgaga acgugaugcu gcugaccucc 180

gacgcccccg aguacaagcc cugggcccug gugauccagg acuccaacgg cgagaacaag 240

aucaagaugc uguccggcgg cuccaagcgg accgccgacg gcuccgaguu cgagucccccc 300

aagaagaagc ggaaggugga guga 324

<210> 36
<211> 1379
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 36

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

```
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
        35                  40                  45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
    50                  55                  60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65                  70                  75                  80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                85                  90                  95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220
```

```
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430
```

```
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620
```

```
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830
```

```
Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990

Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005

Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020
```

```
Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035

Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050

Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065

Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080

Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
    1085                1090                1095

Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110

Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
    1115                1120                1125

Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140

Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155

Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170

Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185

Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215
```

```
Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245

Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260

His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275

Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290

Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305

Ile Ile  His Leu Phe Thr Leu  Thr Asn Leu Gly Ala  Pro Ala Ala
    1310                1315                1320

Phe Lys  Tyr Phe Asp Thr Thr  Ile Asp Arg Lys Arg  Tyr Thr Ser
    1325                1330                1335

Thr Lys  Glu Val Leu Asp Ala  Thr Leu Ile His Gln  Ser Ile Thr
    1340                1345                1350

Gly Leu  Tyr Glu Thr Arg Ile  Asp Leu Ser Gln Leu  Gly Gly Asp
    1355                1360                1365

Gly Gly  Gly Ser Pro Lys Lys  Lys Arg Lys Val
    1370                1375
```

<210> 37

<400> 37
000

<210> 38

<400> 38
000

<210> 39

<400> 39
000

<210> 40
<211> 199
<212> PRT
<213> Homo sapiens

<400> 40

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1   5   10   15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20   25   30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35   40   45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50   55   60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65   70   75   80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85   90   95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100   105   110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115   120   125

```
Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn
        195
```

<210> 41
<211> 229
<212> PRT
<213> R. norvegicus

<400> 41

```
Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1               5                   10                  15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
            20                  25                  30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
        35                  40                  45

Ser Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val
    50                  55                  60

Asn Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr
65                  70                  75                  80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
                85                  90                  95
```

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val Thr Leu
100 105 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg Asn Arg
115 120 125

Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met
130 135 140

Thr Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn Tyr Ser
145 150 155 160

Pro Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp Val Arg
165 170 175

Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
180 185 190

Leu Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile
195 200 205

Ala Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp
210 215 220

Ala Thr Gly Leu Lys
225

<210> 42
<211> 249
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 42
actaatctgt cagatattat tgaaaaggag accggtaagc aactggttat ccaggaatcc 60

atcctcatgc tcccagagga ggtggaagaa gtcattggga acaagccgga aagcgatata 120

ctcgtgcaca ccgcctacga cgagagcacc gacgagaatg tcatgcttct gactagcgac 180

gcccctgaat acaagccttg ggctctggtc atacaggata gcaacggtga gaacaagatt 240

aagatgctc 249

<210> 43
<211> 83
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 43

Thr Asn Leu Ser Asp Ile Ile Glu Lys Glu Thr Gly Lys Gln Leu Val
1 5 10 15

Ile Gln Glu Ser Ile Leu Met Leu Pro Glu Glu Val Glu Glu Val Ile
20 25 30

Gly Asn Lys Pro Glu Ser Asp Ile Leu Val His Thr Ala Tyr Asp Glu
35 40 45

Ser Thr Asp Glu Asn Val Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr
50 55 60

Lys Pro Trp Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile
65 70 75 80

Lys Met Leu

<210> 44
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 44

agcggcagcg agactccggg cacctcagag tccgccacac ccgaaagt 48

<210> 45
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 45
agcggaagcg aaacaccggg aacaagcgaa agcgcaacac cggaaagc 48

<210> 46
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 46

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser
1 5 10 15

<210> 47
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 47

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala
1 5 10

<210> 48
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 48

Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Gly
1 5 10 15

Gly Ser Gly Gly Ser
20

<210> 49
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 49

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
1 5 10 15

<210> 50
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 50

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> 51
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 51

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

1 5 10 15

<210> 52
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 52

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser
20

<210> 53
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 53

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu
1 5 10 15

Ala Ala Ala Lys
20

<210> 54
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 54

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly

1 5 10 15

Gly Gly Gly Ser
20

<210> 55
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 55

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25

<210> 56
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 56

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
1 5 10 15

Ala Ala Ala Lys
20

<210> 57
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 57

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly
1 5 10 15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
20 25

<210> 58
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 58

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25

<210> 59
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 59

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
20 25

<210> 60
<211> 12
<212> DNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 60
tctggtggtt ct 12

<210> 61
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 61

Ser Gly Gly Ser
1

<210> 62
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 62
cccaagaaga agaggaaagt c 21

<210> 63
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 63

Pro Lys Lys Lys Arg Lys Val
1 5

<210> 64

<211> 5472
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 64
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60

ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120

aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg 180

gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240

gcctggctga ccgcccaacg acccccgccc attgacgtca ataatgacgt atgttcccat 300

agtaacgcca tagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc 360

ccacttggca gtacatcaag tgtatcatat gccaagtccg ccccctattg acgtcaatga 420

cggtaaatgg cccgcctggc attatgccca gtacatgacc ttacgggact ttcctacttg 480

gcagtacatc tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacac 540

caatgggcgt ggatagcggt ttgactcacg gggatttcca agtctccacc ccattgacgt 600

caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactg 660

cgatcgcccg ccccgttgac gcaaatgggc ggtaggcgtg tacggtggga ggtctatata 720

agcagagctc gtttagtgaa ccgtcagatc actagaagct ttattgcggt agtttatcac 780

agttaaattg ctaacgcagt cagtgcttct gacacaacag tctcgaactt aagctgcagt 840

gactctctta aggtagcctt gcagaagttg gtcgtgaggc actgggcagg taagtatcaa 900

ggttacaaga caggtttaag gagaccaata gaaactgggc ttgtcgagac agagaagact 960

cttgcgtttc tgataggcac ctattggtct tactgacatc cactttgcct ttctctccac 1020

aggtgtccac tcccagttca attacagctc ttaaggctag agtacttaat acgactcact 1080

ataggctagc ctcgagaatt cacgcgtggt acctctagag tcgacccggg cggccgcttc 1140

cctttagtga gggttaatgc ttcgagcaga catgataaga tacattgatg agtttggaca 1200

aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc 1260

tttatttgta accattataa gctgcaataa acaagttaac aacaacaatt gcattcattt 1320

tatgtttcag gttcaggggg agatgtggga ggtttttaa agcaagtaaa acctctacaa 1380

atgtggtaaa atccgataag gatcgatccg ggctggcgta atagcgaaga ggcccgcacc 1440

gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggacgcgccc tgtagcggcg 1500

cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc 1560

tagcgcccgc tcctttcgct ttcttcccttcctttctcgc cacgttcgcc ggctttcccc 1620

gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta cggcacctcg 1680

accccaaaaa acttgattag ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg 1740

tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg 1800

gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt ttgccgattt 1860

cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat tttaacaaaa 1920

tattaacgct tacaatttcc tgatgcggta ttttctcctt acgcatctgt gcggtatttc 1980

acaccgcata cgcggatctg cgcagcacca tggcctgaaa taacctctga aagaggaact 2040

tggttaggta ccttctgagg cggaaagaac cagctgtgga atgtgtgtca gttagggtgt 2100

ggaaagtccc caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca 2160

gcaaccaggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 2220

ctcaattagt cagcaaccat agtcccgccc ctaactccgc ccatcccgcc cctaactccg 2280

cccagttccg cccattctcc gccccatggc tgactaattt tttttattta tgcagaggcc 2340

gaggccgcct cggcctctga gctattccag aagtagtgag gaggcttttt tggaggccta 2400

ggcttttgca aaaagcttga ttcttctgac acaacagtct cgaacttaag gctagagcca 2460

ccatgattga acaagatgga ttgcacgcag gttctccggc cgcttgggtg gagaggctat 2520

tcggctatga ctgggcacaa cagacaatcg gctgctctga tgccgccgtg ttccggctgt 2580

cagcgcaggg gcgcccggtt ctttttgtca agaccgacct gtccggtgcc ctgaatgaac 2640

tgcaggacga ggcagcgcgg ctatcgtggc tggccacgac gggcgttcct tgcgcagctg 2700

tgctcgacgt tgtcactgaa gcgggaaggg actggctgct attgggcgaa gtgccggggc 2760

aggatctcct gtcatctcac cttgctcctg ccgagaaagt atccatcatg gctgatgcaa 2820

tgcggcggct gcatacgctt gatccggcta cctgcccatt cgaccaccaa gcgaaacatc 2880

gcatcgagcg agcacgtact cggatggaag ccggtcttgt cgatcaggat gatctggacg 2940

aagagcatca ggggctcgcg ccagccgaac tgttcgccag gctcaaggcg cgcatgcccg 3000

acggcgagga tctcgtcgtg acccatggcg atgcctgctt gccgaatatc atggtggaaa 3060

atggccgctt ttctggattc atcgactgtg gccggctggg tgtggcggac cgctatcagg 3120

acatagcgtt ggctacccgt gatattgctg aagagcttgg cggcgaatgg gctgaccgct 3180

tcctcgtgct ttacggtatc gccgctcccg attcgcagcg catcgccttc tatcgccttc 3240

ttgacgagtt cttctgagcg ggactctggg gttcgaaatg accgaccaag cgacgcccaa 3300

cctgccatca cgatggccgc aataaaatat ctttattttc attacatctg tgtgttggtt 3360

ttttgtgtga atcgatagcg ataaggatcc gcgtatggtg cactctcagt acaatctgct 3420

ctgatgccgc atagttaagc cagccccgac acccgccaac acccgctgac gcgccctgac 3480

gggcttgtct gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca 3540

tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctcgtgatac 3600

gcctatttt ataggttaat gtcatgataa taatggtttc ttagacgtca ggtggcactt 3660

ttcggggaaa tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaatatgt 3720

atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta 3780

tgagtattca acatttccgt gtcgccctta ttccctttt tgcggcattt tgccttcctg 3840

tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac 3900

gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg 3960

aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc 4020

gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 4080

ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat 4140

gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg 4200

gaggaccgaa ggagctaacc gctttttttgc acaacatggg ggatcatgta actcgccttg 4260

atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc 4320

ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt 4380

cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct 4440

cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc 4500

gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca 4560

cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct 4620

cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt 4680

taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga 4740

ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca 4800

aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac 4860

caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg 4920

taactggctt cagcagagcg cagataccaa atactgttct tctagtgtag ccgtagttag 4980

gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac 5040

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt 5100

taccggataa ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg 5160

agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc 5220

ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc 5280

gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc 5340

acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa 5400

acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgg 5460

ctcgacagat ct 5472

<210> 65
<211> 4401
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 65
gataagaagt actcaatcgg gctggatatc ggaactaatt ccgtgggttg ggcagtgatc 60

acggatgaat acaaagtgcc gtccaagaag ttcaaggtcc tggggaacac cgatagacac 120

agcatcaaga aaaatctcat cggagccctg ctgtttgact ccggcgaaac cgcagaagcg 180

acccggctca aacgtaccgc gaggcgacgc tacacccggc ggaagaatcg catctgctat 240

ctgcaagaga tcttttcgaa cgaaatggca aaggtcgacg acagcttctt ccaccgcctg 300

gaagaatctt tcctggtgga ggaggacaag aagcatgaac ggcatcctat ctttggaaac 360

atcgtcgacg aagtggcgta ccacgaaaag tacccgacca tctaccatct gcggaagaag 420

ttggttgact caactgacaa ggccgacctc agattgatct acttggccct cgcccatatg 480

atcaaattcc gcggacactt cctgatcgaa ggcgatctga accctgataa ctccgacgtg 540

gataagcttt tcattcaact ggtgcagacc tacaaccaac tgttcgaaga aaacccaatc 600

aatgctagcg gcgtcgatgc caaggccatc ctgtccgccc ggctgtcgaa gtcgcggcgc 660

ctcgaaaacc tgatcgcaca gctgccggga gagaaaaaga acggactttt cggcaacttg 720

atcgctctct cactgggact cactcccaat ttcaagtcca attttgacct ggccgaggac 780

gcgaagctgc aactctcaaa ggacacctac gacgacgact tggacaattt gctggcacaa 840

attggcgatc agtacgcgga tctgttcctt gccgctaaga acctttcgga cgcaatcttg 900

ctgtccgata tcctgcgcgt gaacaccgaa ataaccaaag cgccgcttag cgcctcgatg 960

attaagcggt acgacgagca tcaccaggat ctcacgctgc tcaaagcgct cgtgagacag 1020

caactgcctg aaaagtacaa ggagatcttc ttcgaccagt ccaagaatgg gtacgcaggg 1080

tacatcgatg gaggcgctag ccaggaagag ttctataagt tcatcaagcc aatcctggaa 1140

aagatggacg gaaccgaaga actgctggtc aagctgaaca gggaggatct gctccggaaa 1200

cagagaacct ttgacaacgg atccattccc caccagatcc atctgggtga gctgcacgcc 1260

atcttgcggc gccaggagga cttttaccca ttcctcaagg acaaccggga aaagatcgag 1320

aaaattctga cgttccgcat cccgtattac gtgggcccac tggcgcgcgg caattcgcgc 1380

ttcgcgtgga tgactagaaa atcagaggaa accatcactc cttggaattt cgaggaagtt 1440

gtggataagg gagcttcggc acaaagcttc atcgaacgaa tgaccaactt cgacaagaat 1500

ctcccaaacg agaaggtgct tcctaagcac agcctccttt acgaatactt cactgtctac 1560

aacgaactga ctaaagtgaa atacgttact gaaggaatga ggaagccggc ctttctgtcc 1620

ggagaacaga agaaagcaat tgtcgatctg ctgttcaaga ccaaccgcaa ggtgaccgtc 1680

aagcagctta aagaggacta cttcaagaag atcgagtgtt tcgactcagt ggaaatcagc 1740

ggggtggagg acagattcaa cgcttcgctg ggaacctatc atgatctcct gaagatcatc 1800

aaggacaagg acttccttga caacgaggag aacgaggaca tcctggaaga tatcgtcctg 1860

accttgaccc ttttcgagga tcgcgagatg atcgaggaga ggcttaagac ctacgctcat 1920

ctcttcgacg ataaggtcat gaaacaactc aagcgccgcc ggtacactgg ttggggccgc 1980

ctctcccgca agctgatcaa cggtattcgc gataaacaga gcggtaaaac tatcctggat 2040

ttcctcaaat cggatggctt cgctaatcgt aacttcatgc aattgatcca cgacgacagc 2100

ctgaccttta aggaggacat ccaaaaagca caagtgtccg gacagggaga ctcactccat 2160

gaacacatcg cgaatctggc cggttcgccg gcgattaaga agggaattct gcaaactgtg 2220

aaggtggtcg acgagctggt gaaggtcatg ggacggcaca aaccggagaa tatcgtgatt 2280

gaaatggccc gagaaaacca gactacccag aagggccaga aaaactcccg cgaaaggatg 2340

aagcggatcg aagaaggaat caaggagctg ggcagccaga tcctgaaaga gcacccggtg 2400

gaaaacacgc agctgcagaa cgagaagctc tacctgtact atttgcaaaa tggacgggac 2460

atgtacgtgg accaagagct ggacatcaat cggttgtctg attacgacgt ggaccacatc 2520

gttccacagt cctttctgaa ggatgactcg atcgataaca aggtgttgac tcgcagcgac 2580

aagaacagag ggaagtcaga taatgtgcca tcggaggagg tcgtgaagaa gatgaagaat 2640

tactggcggc agctcctgaa tgcgaagctg attacccaga gaaagtttga caatctcact 2700

aaagccgagc gcggcggact ctcagagctg gataaggctg gattcatcaa acggcagctg 2760

gtcgagactc ggcagattac caagcacgtg gcgcagatct tggactcccg catgaacact 2820

aaatacgacg agaacgataa gctcatccgg gaagtgaagg tgattaccct gaaaagcaaa 2880

cttgtgtcgg actttcggaa ggactttcag ttttacaaag tgagagaaat caacaactac 2940

catcacgcgc atgacgcata cctcaacgct gtggtcggta ccgccctgat caaaaagtac 3000

cctaaacttg aatcggagtt tgtgtacgga gactacaagg tctacgacgt gaggaagatg 3060

atagccaagt ccgaacagga aatcgggaaa gcaactgcga aatacttctt ttactcaaac 3120

atcatgaact ttttcaagac tgaaattacg ctggccaatg gagaaatcag gaagaggcca 3180

ctgatcgaaa ctaacggaga aacgggcgaa atcgtgtggg acaagggcag ggacttcgca 3240

actgttcgca aagtgctctc tatgccgcaa gtcaatattg tgaagaaaac cgaagtgcaa 3300

accggcggat tttcaaagga atcgatcctc ccaaagagaa atagcgacaa gctcattgca 3360

cgcaagaaag actgggaccc gaagaagtac ggaggattcg attcgccgac tgtcgcatac 3420

tccgtcctcg tggtggccaa ggtggagaag ggaaagagca aaaagctcaa atccgtcaaa 3480

gagctgctgg ggattaccat catggaacga tcctcgttcg agaagaaccc gattgatttc 3540

ctcgaggcga agggttacaa ggaggtgaag aaggatctga tcatcaaact ccccaagtac 3600

tcactgttcg aactggaaaa tggtcggaag cgcatgctgg cttcggccgg agaactccaa 3660

aaaggaaatg agctggcctt gcctagcaag tacgtcaact tcctctatct tgcttcgcac 3720

tacgaaaaac tcaaagggtc accggaagat aacgaacaga agcagctttt cgtggagcag 3780

cacaagcatt atctggatga aatcatcgaa caaatctccg agttttcaaa gcgcgtgatc 3840

ctcgccgacg ccaacctcga caaagtcctg tcggcctaca ataagcatag agataagccg 3900

atcagagaac aggccgagaa cattatccac ttgttcaccc tgactaacct gggagcccca 3960

gccgccttca agtacttcga tactactatc gatcgcaaaa gatacacgtc caccaaggaa 4020

gttctggacg cgaccctgat ccaccaaagc atcactggac tctacgaaac taggatcgat 4080

ctgtcgcagc tgggtggcga ttctggtggt tctactaatc tgtcagatat tattgaaaag 4140

gagaccggta agcaactggt tatccaggaa tccatcctca tgctcccaga ggaggtggaa 4200

gaagtcattg ggaacaagcc ggaaagcgat atactcgtgc acaccgccta cgacgagagc 4260

accgacgaga atgtcatgct tctgactagc gacgcccctg aatacaagcc ttgggctctg 4320

gtcatacagg atagcaacgg tgagaacaag attaagatgc tctctggtgg ttctcccaag 4380

aagaagagga aagtctaata g 4401

<210> 66
<211> 2686
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 66
gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt 60

gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag 120

ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct 180

attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt 240

gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc 300

tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt 360

agctccttcg gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg 420

gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg 480

actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct 540

tgcccggcgt caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc 600

attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt 660

tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt 720

tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg 780

aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta tcagggttat 840

tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 900

cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta 960

acctataaaa ataggcgtat cacgaggccc tttcgtctcg cgcgtttcgg tgatgacggt 1020

gaaaacctct gacacatgca gctcccggag acggtcacag cttgtctgta agcggatgcc 1080

gggagcagac aagcccgtca gggcgcgtca gcgggtgttg gcgggtgtcg gggctggctt 1140

aactatgcgg catcagagca gattgtactg agagtgcacc atatgcggtg tgaaataccg 1200

cacagatgcg taaggagaaa ataccgcatc aggcgccatt cgccattcag gctgcgcaac 1260

tgttgggaag ggcgatcggt gcgggcctct tcgctattac gccagctggc gaaaggggga 1320

tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt cccagtcacg acgttgtaaa 1380

acgacggcca gtgaattcga gctcggtacc cggggatcct ctagagtcga cctgcaggca 1440

tgcaagcttg gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt atccgctcac 1500

aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt 1560

gagctaactc acattaattg cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc 1620

gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggcg 1680

ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt 1740

atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa 1800

gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 1860

gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag 1920

gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt 1980

gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg 2040

aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg 2100

ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 2160

taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac 2220

tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg 2280

gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt 2340

taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 2400

tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 2460

tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt 2520

ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt 2580

taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 2640

tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagtt 2686

<210> 67
<211> 233
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 67
tttcccatga ttccttcata tttgcatata cgatacaagg ctgttagaga gataattgga 60

attaatttga ctgtaaacac aaagatatta gtacaaaata cgtgacgtag aaagtaataa 120

tttcttgggt agtttgcagt tttaaaatta tgttttaaaa tggactatca tatgcttacc 180

gtaacttgaa agtatttcga tttcttggct ttatatatct tgtggaaagg acg 233

<210> 68
<211> 588
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 68
atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata 60

acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat 120

aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 180

gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtccgcc 240

ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 300

acgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 360

gcggttttgg cagtacacca atgggcgtgg atagcggttt gactcacggg gatttccaag 420

tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 480

aaaatgtcgt aacaactgcg atcgcccgcc ccgttgacgc aaatgggcgg taggcgtgta 540

cggtgggagg tctatataag cagagctcgt ttagtgaacc gtcagatc 588

<210> 69

<211> 183
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 69
caucacauuu aaaagcaucu cagccuacca ugagaauaag agaaagaaaa ugaagaucaa 60

uagcuuauuc aucucuuuuu cuuuuucguu gguguaaagc caacacccug ucuaaaaaac 120

auaaauuucu uuaaucauuu ugccucuuuu cucugugcuu caauuaauaa aaaauggaaa 180

gaa 183

<210> 70
<211> 1379
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 70

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

```
Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
```

```
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495
```

```
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
```

```
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895
```

```
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990

Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005

Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020

Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035

Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050

Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065

Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080

Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
```

```
    1085                1090                1095

Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110

Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
    1115                1120                1125

Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140

Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155

Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170

Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185

Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215

Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245

Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260

His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275
```

```
Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290

Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305

Ile Ile  His Leu Phe Thr Leu  Thr Asn Leu Gly Ala  Pro Ala Ala
    1310                1315                1320

Phe Lys  Tyr Phe Asp Thr Thr  Ile Asp Arg Lys Arg  Tyr Thr Ser
    1325                1330                1335

Thr Lys  Glu Val Leu Asp Ala  Thr Leu Ile His Gln  Ser Ile Thr
    1340                1345                1350

Gly Leu  Tyr Glu Thr Arg Ile  Asp Leu Ser Gln Leu  Gly Gly Asp
    1355                1360                1365

Gly Gly  Gly Ser Pro Lys Lys  Lys Arg Lys Val
    1370                1375
```

<210> 71
<211> 4140
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 71

```
auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc       60

aucacagacg aauacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga      120

cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa      180

gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc      240

uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga      300

cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga      360
```

aacaucgucg acgaagucgc auaccacgaa aaguacccga caaucuacca ccugagaaag 420

aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480

augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540

gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600

aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660

agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720

cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa 780

gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840

cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900

cugcugagcg acauccugag agucaacaca gaaaucacaa aggcaccgcu gagcgcaagc 960

augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020

cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca 1080

ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucaucaa gccgauccug 1140

gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200

aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260

gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc 1320

gaaaagaucc ugacauucag aaucccguac uacgucggac cgcuggcaag aggaaacagc 1380

agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440

gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500

aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560

uacaacgaac ugacaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620

agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680

gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc 1740

agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800

aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860

cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920

caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980

agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040

gacuuccuga agagcgacgg auucgcaaac agaaacuuca ugcagcugau ccacgacgac 2100

agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160

cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220

gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280

aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340

augaagagaa ucgaagaagg aaucaaggaa cugggaagcc agauccugaa ggaacacccg 2400

gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460

gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520

aucgucccgc agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580

gacaagaaca gaggaaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640

aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700

acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760

cuggucgaaa caagacagau cacaaagcac gucgcacaga uccuggacag cagaaugaac 2820

acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880

aagcugguca gcgacuucag aaaggacuuc caguucuaca aggucagaga aaucaacaac 2940

uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000

uacccgaagc uggaaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag 3060

augaucgcaa agagcgaaca ggaaaucgga aaggcaacag caaaguacuu cuucuacagc 3120

aacaucauga acuucuucaa gacagaaauc acacuggcaa acggagaaau cagaaagaga 3180

ccgcugaucg aaacaaacgg agaaacagga gaaaucgucu gggacaaggg aagagacuuc 3240

gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc 3300

cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360

gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420

uacagcgucc uggucgucgc aaaggucgaa aagggaaaga gcaagaagcu gaagagcguc 3480

aaggaacugc ugggaaucac aaucauggaa agaagcagcu ucgaaaagaa cccgaucgac 3540

uuccuggaag caaagggaua caaggaaguc aagaaggacc ugaucaucaa gcugccgaag 3600

uacagccugu ucgaacugga aaacggaaga aagagaaugc uggcaagcgc aggagaacug 3660

cagaagggaa acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720

cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780

cagcacaagc acuaccugga cgaaaucauc gaacagauca gcgaauucag caagagaguc 3840

auccuggcag acgcaaaccu ggacaaggUc cugagcgcau acaacaagca cagagacaag 3900

ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960

ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaaag 4020

gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaauc 4080

gaccugagcc agcugggagg agacggagga ggaagcccga agaagaagag aaaggucuag 4140

<210> 72
<211> 4134
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 72
gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60

acagacgaau acaaggUccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120

agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180

acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240

cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug 300

gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac 360

aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420

cuggucgaca gcacagacaa ggcagaccug agacugaucu accuggcacu ggcacacaug 480

aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccggacaa cagcgacguc 540

gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc 600

aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660

cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaaccug 720

aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780

gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840

aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900

cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960

aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020

cagcugccgg aaaaguacaa ggaaaucuuc uucgaccaga gcaagaacgg auacgcagga 1080

uacaucgacg gaggagcaag ccaggaagaa uucuacaagu ucaucaagcc gauccuggaa 1140

aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200

cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca 1260

auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320

aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380

uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacuu cgaagaaguc 1440

gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500

cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560

aacgaacuga caaaggucaa guacgucaca gaaggaauga gaaagccggc auuccugagc 1620

ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680

aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740

ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucauc 1800

aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860

acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920

cuguucgacg acaaggucau gaagcagcug aagagaagaa gauacacagg auggggaaga 1980

cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaaagac aauccuggac 2040

uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc 2100

cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160

gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220

aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280

gaaauggcaa gagaaaacca gacaacacag aagggacaga agaacagcag agaaagaaug 2340

aagagaaucg aagaaggaau caaggaacug ggaagccaga uccugaagga acacccgguc 2400

gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460

auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacauc 2520

gucccgcaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2580

aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640

uacuggagac agcugcugaa cgcaaagcug aucacacaga gaaaguucga caaccugaca 2700

aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug 2760

gucgaaacaa gacagaucac aaagcacguc gcacagaucc uggacagcag aaugaacaca 2820

aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880

cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaau caacaacuac 2940

caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000

ccgaagcugg aaagcgaauu cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060

aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac 3120

aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180

cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240

acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300

acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360

agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcauac 3420

agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcgucaag 3480

gaacugcugg gaaucacaau cauggaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540

cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600

agccuguucg aacuggaaaa cggaagaaag agaaugcugg caagcgcagg agaacugcag 3660

aagggaaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720

uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780

cacaagcacu accuggacga aaucaucgaa cagaucagcg aauucagcaa gagagucauc 3840

cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900

aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960

gcagcauuca aguacuucga cacaacaauc gacagaaaga gauacacaag cacaaaggaa 4020

guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080

cugagccagc ugggaggaga cggaggagga agcccgaaga agaagagaaa gguc 4134

<210> 73
<211> 1368
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 73

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

```
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65                  70                  75                  80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                85                  90                  95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
```

```
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460
```

```
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
```

```
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860
```

```
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880


Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895


Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910


Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925


Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940


Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960


Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975


Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990


Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005


Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020


Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035


Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050


Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
```

```
    1055                1060                1065

Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080

Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
    1085                1090                1095

Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110

Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
    1115                1120                1125

Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140

Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155

Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170

Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185

Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215

Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245
```

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
1355 1360 1365

<210> 74
<211> 4107
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 74
auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc 60

aucacagacg aauacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120

cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180

gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc 240

uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga 300

cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360

aacaucgucg acgaagucgc auaccacgaa aaguacccga caaucuacca ccugagaaag 420

aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480

augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540

gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600

aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660

agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720

cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa 780

gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840

cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900

cugcugagcg acauccugag agucaacaca gaaaucacaa aggcaccgcu gagcgcaagc 960

augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020

cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca 1080

ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucaucaa gccgauccug 1140

gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200

aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260

gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc 1320

gaaaagaucc ugacauucag aaucccguac uacgucggac cgcuggcaag aggaaacagc 1380

agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440

gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500

aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560

uacaacgaac ugacaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620

agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680

gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc 1740

agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800

aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860

cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920

caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980

agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040

gacuuccuga agagcgacgg auucgcaaac agaaacuuca ugcagcugau ccacgacgac 2100

agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160

cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220

gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280

aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340

augaagagaa ucgaagaagg aaucaaggaa cugggaagcc agauccugaa ggaacacccg 2400

gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460

gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520

aucgucccgc agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580

gacaagaaca gaggaaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640

aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700

acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760

cuggucgaaa caagacagau cacaaagcac gucgcacaga uccuggacag cagaaugaac 2820

acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880

aagcugguca gcgacuucag aaaggacuuc caguucuaca aggucagaga aaucaacaac 2940

uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000

uacccgaagc uggaaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag 3060

augaucgcaa agagcgaaca ggaaaucgga aaggcaacag caaaguacuu cuucuacagc 3120

aacaucauga acuucuucaa gacagaaauc acacuggcaa acggagaaau cagaaagaga 3180

ccgcugaucg aaacaaacgg agaaacagga gaaaucgucu gggacaaggg aagagacuuc 3240

gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc 3300

cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360

gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420

uacagcgucc uggucgucgc aaaggucgaa aagggaaaga gcaagaagcu gaagagcguc 3480

aaggaacugc ugggaaucac aaucauggaa agaagcagcu ucgaaaagaa cccgaucgac 3540

uuccuggaag caaagggaua caaggaaguc aagaaggacc ugaucaucaa gcugccgaag 3600

uacagccugu ucgaacugga aaacggaaga aagagaaugc uggcaagcgc aggagaacug 3660

cagaagggaa acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720

cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780

cagcacaagc acuaccugga cgaaaucauc gaacagauca gcgaauucag caagagaguc 3840

auccuggcag acgcaaaccu ggacaagguc cugagcgcau acaacaagca cagagacaag 3900

ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960

ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaaag 4020

gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaauc 4080

gaccugagcc agcugggagg agacuag 4107

<210> 75
<211> 4101
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 75
gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60

acagacgaau acaaggcccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120

agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180

acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240

cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug 300

gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac 360

aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420

cuggucgaca gcacagacaa ggcagaccug agacugaucu accuggcacu ggcacacaug 480

aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccggacaa cagcgacguc 540

gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc 600

aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660

cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaaccug 720

aucgcacuga gccuggggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780

gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840

aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900

cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960

aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020

cagcugccgg aaaaguacaa ggaaaucuuc uucgaccaga gcaagaacgg auacgcagga 1080

uacaucgacg gaggagcaag ccaggaagaa uucuacaagu ucaucaagcc gauccuggaa 1140

aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200

cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca 1260

auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320

aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380

uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacuu cgaagaaguc 1440

gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500

cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560

aacgaacuga caaaggucaa guacgucaca gaaggaauga gaaagccggc auuccugagc 1620

ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680

aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740

ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucauc 1800

aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860

acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920

cuguucgacg acaaggucau gaagcagcug aagagaagaa gauacacagg auggggaaga 1980

cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaaagac aauccuggac 2040

uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc 2100

cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160

gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220

aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280

gaaauggcaa gagaaaacca gacaacacag aagggacaga agaacagcag agaaagaaug 2340

aagagaaucg aagaaggaau caaggaacug ggaagccaga uccugaagga acacccgguc 2400

gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460

auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacauc 2520

gucccgcaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2580

aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640

uacuggagac agcugcugaa cgcaaagcug aucacacaga gaaaguucga caaccugaca 2700

aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug 2760

gucgaaacaa gacagaucac aaagcacguc gcacagaucc uggacagcag aaugaacaca 2820

aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880

cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaau caacaacuac 2940

caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000

ccgaagcugg aaagcgaauu cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060

aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac 3120

aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180

cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240

acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300

acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360

agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcauac 3420

agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcgucaag 3480

gaacugcugg gaaucacaau cauggaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540

cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600

agccuguucg aacuggaaaa cggaagaaag agaaugcugg caagcgcagg agaacugcag 3660

aagggaaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720

uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780

cacaagcacu accuggacga aaucaucgaa cagaucagcg aauucagcaa gagagucauc 3840

cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900

aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960

gcagcauuca aguacuucga cacaacaauc gacagaaaga gauacacaag cacaaaggaa 4020

guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080

cugagccagc ugggaggaga c 4101

<210> 76
<211> 1391
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 76

Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly
1 5 10 15

Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys
20 25 30

Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly
35 40 45

```
Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys
    50                  55                  60

Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr
65                  70                  75                  80

Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe
                85                  90                  95

Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His
            100                 105                 110

Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His
        115                 120                 125

Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser
    130                 135                 140

Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met
145                 150                 155                 160

Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp
                165                 170                 175

Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn
            180                 185                 190

Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys
        195                 200                 205

Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu
    210                 215                 220

Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu
225                 230                 235                 240

Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp
```

```
                245                 250                 255

Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp
            260                 265                 270

Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu
        275                 280                 285

Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile
    290                 295                 300

Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
305                 310                 315                 320

Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala
                325                 330                 335

Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
            340                 345                 350

Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln
        355                 360                 365

Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly
    370                 375                 380

Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys
385                 390                 395                 400

Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly
                405                 410                 415

Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu
            420                 425                 430

Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro
        435                 440                 445
```

```
Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met
    450                 455                 460

Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val
465                 470                 475                 480

Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn
                485                 490                 495

Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu
            500                 505                 510

Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr
        515                 520                 525

Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys
    530                 535                 540

Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
545                 550                 555                 560

Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser
                565                 570                 575

Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr
            580                 585                 590

Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn
        595                 600                 605

Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu
    610                 615                 620

Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His
625                 630                 635                 640

Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr
```

```
                645                 650                 655

Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys
            660                 665                 670

Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala
        675                 680                 685

Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys
    690                 695                 700

Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His
705                 710                 715                 720

Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
                725                 730                 735

Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg
            740                 745                 750

His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr
        755                 760                 765

Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu
    770                 775                 780

Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val
785                 790                 795                 800

Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln
                805                 810                 815

Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu
            820                 825                 830

Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp
        835                 840                 845
```

```
Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly
    850                 855                 860

Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn
865                 870                 875                 880

Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe
                885                 890                 895

Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys
            900                 905                 910

Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys
        915                 920                 925

His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu
    930                 935                 940

Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys
945                 950                 955                 960

Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu
                965                 970                 975

Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val
            980                 985                 990

Gly Thr Ala Leu Ile Lys Lys Tyr  Pro Lys Leu Glu Ser  Glu Phe Val
        995                 1000                1005

Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg Lys Met  Ile Ala Lys
    1010                1015                1020

Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala Lys Tyr  Phe Phe Tyr
    1025                1030                1035

Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu Ile Thr  Leu Ala Asn
```

```
    1040                1045                1050

Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu Thr Asn  Gly Glu Thr
    1055                1060                1065

Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp Phe Ala  Thr Val Arg
    1070                1075                1080

Lys Val  Leu Ser Met Pro Gln  Val Asn Ile Val Lys  Lys Thr Glu
    1085                1090                1095

Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser Ile Leu  Pro Lys Arg
    1100                1105                1110

Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys Asp Trp  Asp Pro Lys
    1115                1120                1125

Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val Ala Tyr  Ser Val Leu
    1130                1135                1140

Val Val  Ala Lys Val Glu Lys  Gly Lys Ser Lys Lys  Leu Lys Ser
    1145                1150                1155

Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met Glu Arg  Ser Ser Phe
    1160                1165                1170

Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala Lys Gly  Tyr Lys Glu
    1175                1180                1185

Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro Lys Tyr  Ser Leu Phe
    1190                1195                1200

Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu Ala Ser  Ala Gly Glu
    1205                1210                1215

Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro Ser Lys  Tyr Val Asn
    1220                1225                1230
```

```
Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys Leu Lys  Gly Ser Pro
    1235                1240                1245


Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val Glu Gln  His Lys His
    1250                1255                1260


Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser Glu Phe  Ser Lys Arg
    1265                1270                1275


Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys Val Leu  Ser Ala Tyr
    1280                1285                1290


Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu Gln Ala  Glu Asn Ile
    1295                1300                1305


Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly Ala Pro  Ala Ala Phe
    1310                1315                1320


Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys Arg Tyr  Thr Ser Thr
    1325                1330                1335


Lys Glu  Val Leu Asp Ala Thr  Leu Ile His Gln Ser  Ile Thr Gly
    1340                1345                1350


Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln Leu Gly  Gly Asp Gly
    1355                1360                1365


Ser Gly  Ser Pro Lys Lys Lys  Arg Lys Val Asp Gly  Ser Pro Lys
    1370                1375                1380


Lys Lys  Arg Lys Val Asp Ser  Gly
    1385                1390
```

<210> 77
<211> 4179
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 77
auggacaaga aguacagcau cggacuggca aucggaacaa acagcgucgg augggcaguc 60

aucacagacg aauacaaggu cccgagcaag aaguucaagg uccugggaaa cacagacaga 120

cacagcauca agaagaaccu gaucggagca cugcuguucg acagcggaga aacagcagaa 180

gcaacaagac ugaagagaac agcaagaaga agauacacaa gaagaaagaa cagaaucugc 240

uaccugcagg aaaucuucag caacgaaaug gcaaaggucg acgacagcuu cuuccacaga 300

cuggaagaaa gcuuccuggu cgaagaagac aagaagcacg aaagacaccc gaucuucgga 360

aacaucgucg acgaagucgc auaccacgaa aaguacccga caaucuacca ccugagaaag 420

aagcuggucg acagcacaga caaggcagac cugagacuga ucuaccuggc acuggcacac 480

augaucaagu ucagaggaca cuuccugauc gaaggagacc ugaacccgga caacagcgac 540

gucgacaagc uguucaucca gcugguccag acauacaacc agcuguucga agaaaacccg 600

aucaacgcaa gcggagucga cgcaaaggca auccugagcg caagacugag caagagcaga 660

agacuggaaa accugaucgc acagcugccg ggagaaaaga agaacggacu guucggaaac 720

cugaucgcac ugagccuggg acugacaccg aacuucaaga gcaacuucga ccuggcagaa 780

gacgcaaagc ugcagcugag caaggacaca uacgacgacg accuggacaa ccugcuggca 840

cagaucggag accaguacgc agaccuguuc cuggcagcaa agaaccugag cgacgcaauc 900

cugcugagcg acauccugag agucaacaca gaaaucacaa aggcaccgcu gagcgcaagc 960

augaucaaga gauacgacga acaccaccag gaccugacac ugcugaaggc acuggucaga 1020

cagcagcugc cggaaaagua caaggaaauc uucuucgacc agagcaagaa cggauacgca 1080

ggauacaucg acggaggagc aagccaggaa gaauucuaca aguucaucaa gccgauccug 1140

gaaaagaugg acggaacaga agaacugcug gucaagcuga acagagaaga ccugcugaga 1200

aagcagagaa cauucgacaa cggaagcauc ccgcaccaga uccaccuggg agaacugcac 1260

gcaauccuga gaagacagga agacuucuac ccguuccuga aggacaacag agaaaagauc 1320

gaaaagaucc ugacauucag aaucccguac uacgucggac cgcuggcaag aggaaacagc 1380

agauucgcau ggaugacaag aaagagcgaa gaaacaauca caccguggaa cuucgaagaa 1440

gucgucgaca agggagcaag cgcacagagc uucaucgaaa gaaugacaaa cuucgacaag 1500

aaccugccga acgaaaaggu ccugccgaag cacagccugc uguacgaaua cuucacaguc 1560

uacaacgaac ugacaaaggu caaguacguc acagaaggaa ugagaaagcc ggcauuccug 1620

agcggagaac agaagaaggc aaucgucgac cugcuguuca agacaaacag aaaggucaca 1680

gucaagcagc ugaaggaaga cuacuucaag aagaucgaau gcuucgacag cgucgaaauc 1740

agcggagucg aagacagauu caacgcaagc cugggaacau accacgaccu gcugaagauc 1800

aucaaggaca aggacuuccu ggacaacgaa gaaaacgaag acauccugga agacaucguc 1860

cugacacuga cacuguucga agacagagaa augaucgaag aaagacugaa gacauacgca 1920

caccuguucg acgacaaggu caugaagcag cugaagagaa gaagauacac aggaugggga 1980

agacugagca gaaagcugau caacggaauc agagacaagc agagcggaaa gacaauccug 2040

gacuuccuga agagcgacgg auucgcaaac agaaacuuca ugcagcugau ccacgacgac 2100

agccugacau ucaaggaaga cauccagaag gcacagguca gcggacaggg agacagccug 2160

cacgaacaca ucgcaaaccu ggcaggaagc ccggcaauca agaagggaau ccugcagaca 2220

gucaaggucg ucgacgaacu ggucaagguc augggaagac acaagccgga aaacaucguc 2280

aucgaaaugg caagagaaaa ccagacaaca cagaagggac agaagaacag cagagaaaga 2340

augaagagaa ucgaagaagg aaucaaggaa cugggaagcc agauccugaa ggaacacccg 2400

gucgaaaaca cacagcugca gaacgaaaag cuguaccugu acuaccugca gaacggaaga 2460

gacauguacg ucgaccagga acuggacauc aacagacuga gcgacuacga cgucgaccac 2520

aucgucccgc agagcuuccu gaaggacgac agcaucgaca acaagguccu gacaagaagc 2580

gacaagaaca gaggaaagag cgacaacguc ccgagcgaag aagucgucaa gaagaugaag 2640

aacuacugga gacagcugcu gaacgcaaag cugaucacac agagaaaguu cgacaaccug 2700

acaaaggcag agagaggagg acugagcgaa cuggacaagg caggauucau caagagacag 2760

cuggucgaaa caagacagau cacaaagcac gucgcacaga uccuggacag cagaaugaac 2820

acaaaguacg acgaaaacga caagcugauc agagaaguca aggucaucac acugaagagc 2880

aagcugguca gcgacuucag aaaggacuuc caguucuaca aggucagaga aaucaacaac 2940

uaccaccacg cacacgacgc auaccugaac gcagucgucg gaacagcacu gaucaagaag 3000

uacccgaagc uggaaagcga auucgucuac ggagacuaca aggucuacga cgucagaaag 3060

augaucgcaa agagcgaaca ggaaaucgga aaggcaacag caaaguacuu cuucuacagc 3120

aacaucauga acuucuucaa gacagaaauc acacuggcaa acggagaaau cagaaagaga 3180

ccgcugaucg aaacaaacgg agaaacagga gaaaucgucu gggacaaggg aagagacuuc 3240

gcaacaguca gaaagguccu gagcaugccg caggucaaca ucgucaagaa gacagaaguc 3300

cagacaggag gauucagcaa ggaaagcauc cugccgaaga gaaacagcga caagcugauc 3360

gcaagaaaga aggacuggga cccgaagaag uacggaggau ucgacagccc gacagucgca 3420

uacagcgucc uggucgucgc aaaggucgaa aagggaaaga gcaagaagcu gaagagcguc 3480

aaggaacugc ugggaaucac aaucauggaa agaagcagcu ucgaaaagaa cccgaucgac 3540

uuccuggaag caaagggaua caaggaaguc aagaaggacc ugaucaucaa gcugccgaag 3600

uacagccugu ucgaacugga aaacggaaga aagagaaugc uggcaagcgc aggagaacug 3660

cagaagggaa acgaacuggc acugccgagc aaguacguca acuuccugua ccuggcaagc 3720

cacuacgaaa agcugaaggg aagcccggaa gacaacgaac agaagcagcu guucgucgaa 3780

cagcacaagc acuaccugga cgaaaucauc gaacagauca gcgaauucag caagagaguc 3840

auccuggcag acgcaaaccu ggacaaggUc cugagcgcau acaacaagca cagagacaag 3900

ccgaucagag aacaggcaga aaacaucauc caccuguuca cacugacaaa ccugggagca 3960

ccggcagcau ucaaguacuu cgacacaaca aucgacagaa agagauacac aagcacaaag 4020

gaaguccugg acgcaacacu gauccaccag agcaucacag gacuguacga aacaagaauc 4080

gaccugagcc agcugggagg agacggaagc ggaagcccga agaagaagag aaaggucgac 4140

ggaagcccga agaagaagag aaaggucgac agcggauag 4179

<210> 78
<211> 4173
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 78
gacaagaagu acagcaucgg acuggcaauc ggaacaaaca gcgucggaug ggcagucauc 60

acagacgaau acaaggcccc gagcaagaag uucaaggucc ugggaaacac agacagacac 120

agcaucaaga agaaccugau cggagcacug cuguucgaca gcggagaaac agcagaagca 180

acaagacuga agagaacagc aagaagaaga uacacaagaa gaaagaacag aaucugcuac 240

cugcaggaaa ucuucagcaa cgaaauggca aaggucgacg acagcuucuu ccacagacug 300

gaagaaagcu uccuggucga agaagacaag aagcacgaaa gacacccgau cuucggaaac 360

aucgucgacg aagucgcaua ccacgaaaag uacccgacaa ucuaccaccu gagaaagaag 420

cuggucgaca gcacagacaa ggcagaccug agacugaucu accuggcacu ggcacacaug 480

aucaaguuca gaggacacuu ccugaucgaa ggagaccuga acccggacaa cagcgacguc 540

gacaagcugu ucauccagcu gguccagaca uacaaccagc uguucgaaga aaacccgauc 600

aacgcaagcg gagucgacgc aaaggcaauc cugagcgcaa gacugagcaa gagcagaaga 660

cuggaaaacc ugaucgcaca gcugccggga gaaaagaaga acggacuguu cggaaaccug 720

aucgcacuga gccugggacu gacaccgaac uucaagagca acuucgaccu ggcagaagac 780

gcaaagcugc agcugagcaa ggacacauac gacgacgacc uggacaaccu gcuggcacag 840

aucggagacc aguacgcaga ccuguuccug gcagcaaaga accugagcga cgcaauccug 900

cugagcgaca uccugagagu caacacagaa aucacaaagg caccgcugag cgcaagcaug 960

aucaagagau acgacgaaca ccaccaggac cugacacugc ugaaggcacu ggucagacag 1020

cagcugccgg aaaaguacaa ggaaaucuuc uucgaccaga gcaagaacgg auacgcagga 1080

uacaucgacg gaggagcaag ccaggaagaa uucuacaagu ucaucaagcc gauccuggaa 1140

aagauggacg gaacagaaga acugcugguc aagcugaaca gagaagaccu gcugagaaag 1200

cagagaacau ucgacaacgg aagcaucccg caccagaucc accugggaga acugcacgca 1260

auccugagaa gacaggaaga cuucuacccg uuccugaagg acaacagaga aaagaucgaa 1320

aagauccuga cauucagaau cccguacuac gucggaccgc uggcaagagg aaacagcaga 1380

uucgcaugga ugacaagaaa gagcgaagaa acaaucacac cguggaacuu cgaagaaguc 1440

gucgacaagg gagcaagcgc acagagcuuc aucgaaagaa ugacaaacuu cgacaagaac 1500

cugccgaacg aaaagguccu gccgaagcac agccugcugu acgaauacuu cacagucuac 1560

aacgaacuga caaaggucaa guacgucaca gaaggaauga gaaagccggc auuccugagc 1620

ggagaacaga agaaggcaau cgucgaccug cuguucaaga caaacagaaa ggucacaguc 1680

aagcagcuga aggaagacua cuucaagaag aucgaaugcu ucgacagcgu cgaaaucagc 1740

ggagucgaag acagauucaa cgcaagccug ggaacauacc acgaccugcu gaagaucauc 1800

aaggacaagg acuuccugga caacgaagaa aacgaagaca uccuggaaga caucguccug 1860

acacugacac uguucgaaga cagagaaaug aucgaagaaa gacugaagac auacgcacac 1920

cuguucgacg acaaggucau gaagcagcug aagagaagaa gauacacagg auggggaaga 1980

cugagcagaa agcugaucaa cggaaucaga gacaagcaga gcggaaagac aauccuggac 2040

uuccugaaga gcgacggauu cgcaaacaga aacuucaugc agcugaucca cgacgacagc 2100

cugacauuca aggaagacau ccagaaggca caggucagcg gacagggaga cagccugcac 2160

gaacacaucg caaaccuggc aggaagcccg gcaaucaaga agggaauccu gcagacaguc 2220

aaggucgucg acgaacuggu caaggucaug ggaagacaca agccggaaaa caucgucauc 2280

gaaauggcaa gagaaaacca gacaacacag aagggacaga agaacagcag agaaagaaug 2340

aagagaaucg aagaaggaau caaggaacug ggaagccaga uccugaagga acacccgguc 2400

gaaaacacac agcugcagaa cgaaaagcug uaccuguacu accugcagaa cggaagagac 2460

auguacgucg accaggaacu ggacaucaac agacugagcg acuacgacgu cgaccacauc 2520

gucccgcaga gcuuccugaa ggacgacagc aucgacaaca agguccugac aagaagcgac 2580

aagaacagag gaaagagcga caacgucccg agcgaagaag ucgucaagaa gaugaagaac 2640

uacuggagac agcugcugaa cgcaaagcug aucacacaga gaaaguucga caaccugaca 2700

aaggcagaga gaggaggacu gagcgaacug gacaaggcag gauucaucaa gagacagcug 2760

gucgaaacaa gacagaucac aaagcacguc gcacagaucc uggacagcag aaugaacaca 2820

aaguacgacg aaaacgacaa gcugaucaga gaagucaagg ucaucacacu gaagagcaag 2880

cuggucagcg acuucagaaa ggacuuccag uucuacaagg ucagagaaau caacaacuac 2940

caccacgcac acgacgcaua ccugaacgca gucgucggaa cagcacugau caagaaguac 3000

ccgaagcugg aaagcgaauu cgucuacgga gacuacaagg ucuacgacgu cagaaagaug 3060

aucgcaaaga gcgaacagga aaucggaaag gcaacagcaa aguacuucuu cuacagcaac 3120

aucaugaacu ucuucaagac agaaaucaca cuggcaaacg gagaaaucag aaagagaccg 3180

cugaucgaaa caaacggaga aacaggagaa aucgucuggg acaagggaag agacuucgca 3240

acagucagaa agguccugag caugccgcag gucaacaucg ucaagaagac agaaguccag 3300

acaggaggau ucagcaagga aagcauccug ccgaagagaa acagcgacaa gcugaucgca 3360

agaaagaagg acugggaccc gaagaaguac ggaggauucg acagcccgac agucgcauac 3420

agcguccugg ucgucgcaaa ggucgaaaag ggaaagagca agaagcugaa gagcgucaag 3480

gaacugcugg gaaucacaau cauggaaaga agcagcuucg aaaagaaccc gaucgacuuc 3540

cuggaagcaa agggauacaa ggaagucaag aaggaccuga ucaucaagcu gccgaaguac 3600

agccuguucg aacuggaaaa cggaagaaag agaaugcugg caagcgcagg agaacugcag 3660

aagggaaacg aacuggcacu gccgagcaag uacgucaacu uccuguaccu ggcaagccac 3720

uacgaaaagc ugaagggaag cccggaagac aacgaacaga agcagcuguu cgucgaacag 3780

cacaagcacu accuggacga aaucaucgaa cagaucagcg aauucagcaa gagagucauc 3840

cuggcagacg caaaccugga caagguccug agcgcauaca acaagcacag agacaagccg 3900

aucagagaac aggcagaaaa caucauccac cuguucacac ugacaaaccu gggagcaccg 3960

gcagcauuca aguacuucga cacaacaauc gacagaaaga gauacacaag cacaaaggaa 4020

guccuggacg caacacugau ccaccagagc aucacaggac uguacgaaac aagaaucgac 4080

cugagccagc ugggaggaga cggaagcgga agcccgaaga agaagagaaa ggucgacgga 4140

agcccgaaga agaagagaaa ggucgacagc gga 4173

<210> 79
<211> 4140
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 79
atggacaaga agtactccat cggcctggcc atcggcacca actccgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gccctccaag aagttcaagg tgctgggcaa caccgaccgg 120

cactccatca agaagaacct gatcggcgcc ctgctgttcg actccggcga gaccgccgag 180

gccacccggc tgaagcggac cgcccggcgg cggtacaccc ggcggaagaa ccggatctgc 240

tacctgcagg agatcttctc caacgagatg gccaaggtgg acgactcctt cttccaccgg 300

ctggaggagt ccttcctggt ggaggaggac aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcggaag 420

aagctggtgg actccaccga caaggccgac ctgcggctga tctacctggc cctggcccac 480

atgatcaagt tccggggcca cttcctgatc gagggcgacc tgaaccccga caactccgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcct ccggcgtgga cgccaaggcc atcctgtccg cccggctgtc caagtcccgg 660

cggctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgtccctggg cctgaccccc aacttcaagt ccaacttcga cctggccgag 780

gacgccaagc tgcagctgtc caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgtc cgacgccatc 900

ctgctgtccg acatcctgcg ggtgaacacc gagatcacca aggcccccct gtccgcctcc 960

atgatcaagc ggtacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgg 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agtccaagaa cggctacgcc 1080

ggctacatcg acggcggcgc ctcccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgggagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggctccatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc ggcggcagga ggacttctac cccttcctga aggacaaccg ggagaagatc 1320

gagaagatcc tgaccttccg gatcccctac tacgtgggcc ccctggcccg gggcaactcc 1380

cggttcgcct ggatgacccg gaagtccgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgcctc cgcccagtcc ttcatcgagc ggatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cactccctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcggaagcc cgccttcctg 1620

tccggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg gaaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgactc cgtggagatc 1740

tccggcgtgg aggaccggtt caacgcctcc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgggag atgatcgagg agcggctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcggc ggcggtacac cggctggggc 1980

cggctgtccc ggaagctgat caacggcatc cgggacaagc agtccggcaa gaccatcctg 2040

gacttcctga agtccgacgg cttcgccaac cggaacttca tgcagctgat ccacgacgac 2100

tccctgacct tcaaggagga catccagaag gcccaggtgt ccggccaggg cgactccctg 2160

cacgagcaca tcgccaacct ggccggctcc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccggc acaagcccga gaacatcgtg 2280

atcgagatgg cccgggagaa ccagaccacc cagaagggcc agaagaactc ccgggagcgg 2340

atgaagcgga tcgaggaggg catcaaggag ctgggctccc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgg 2460

gacatgtacg tggaccagga gctggacatc aaccggctgt ccgactacga cgtggaccac 2520

atcgtgcccc agtccttcct gaaggacgac tccatcgaca acaaggtgct gacccggtcc 2580

gacaagaacc ggggcaagtc cgacaacgtg ccctccgagg aggtggtgaa gaagatgaag 2640

aactactggc ggcagctgct gaacgccaag ctgatcaccc agcggaagtt cgacaacctg 2700

accaaggccg agcggggcgg cctgtccgag ctggacaagg ccggcttcat caagcggcag 2760

ctggtggaga cccggcagat caccaagcac gtggcccaga tcctggactc ccggatgaac 2820

accaagtacg acgagaacga caagctgatc cgggaggtga aggtgatcac cctgaagtcc 2880

aagctggtgt ccgacttccg gaaggacttc cagttctaca aggtgcggga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagtccga gttcgtgtac ggcgactaca aggtgtacga cgtgcggaag 3060

atgatcgcca agtccgagca ggagatcggc aaggccaccg ccaagtactt cttctactcc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccggaagcgg 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgggacttc 3240

gccaccgtgc ggaaggtgct gtccatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttctccaa ggagtccatc ctgcccaagc ggaactccga caagctgatc 3360

gcccggaaga aggactggga ccccaagaag tacggcggct tcgactcccc caccgtggcc 3420

tactccgtgc tggtggtggc caaggtggag aagggcaagt ccaagaagct gaagtccgtg 3480

aaggagctgc tgggcatcac catcatggag cggtcctcct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tactccctgt tcgagctgga gaacggccgg aagcggatgc tggcctccgc cggcgagctg 3660

cagaagggca acgagctggc cctgccctcc aagtacgtga acttcctgta cctggcctcc 3720

cactacgaga agctgaaggg ctcccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatct ccgagttctc caagcgggtg 3840

atcctggccg acgccaacct ggacaaggtg ctgtccgcct acaacaagca ccgggacaag 3900

cccatccggg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgga agcggtacac ctccaccaag 4020

gaggtgctgg acgccaccct gatccaccag tccatcaccg gcctgtacga gacccggatc 4080

gacctgtccc agctgggcgg cgacggcggc ggctccccca agaagaagcg gaaggtgtga 4140

<210> 80
<211> 4107
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 80
atggacaaga agtactccat cggcctggcc atcggcacca actccgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gccctccaag aagttcaagg tgctgggcaa caccgaccgg 120

cactccatca agaagaacct gatcggcgcc ctgctgttcg actccggcga gaccgccgag 180

gccacccggc tgaagcggac cgcccggcgg cggtacaccc ggcggaagaa ccggatctgc 240

tacctgcagg agatcttctc caacgagatg gccaaggtgg acgactcctt cttccaccgg 300

ctggaggagt ccttcctggt ggaggaggac aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcggaag 420

aagctggtgg actccaccga caaggccgac ctgcggctga tctacctggc cctggcccac 480

atgatcaagt tccggggcca cttcctgatc gagggcgacc tgaaccccga caactccgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcct ccggcgtgga cgccaaggcc atcctgtccg cccggctgtc caagtcccgg 660

cggctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgtccctggg cctgaccccc aacttcaagt ccaacttcga cctggccgag 780

gacgccaagc tgcagctgtc caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgtc cgacgccatc 900

ctgctgtccg acatcctgcg ggtgaacacc gagatcacca aggcccccct gtccgcctcc 960

atgatcaagc ggtacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgg 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agtccaagaa cggctacgcc 1080

ggctacatcg acggcggcgc ctcccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgggagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggctccatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc ggcggcagga ggacttctac cccttcctga aggacaaccg ggagaagatc 1320

gagaagatcc tgaccttccg gatcccctac tacgtgggcc ccctggcccg gggcaactcc 1380

cggttcgcct ggatgacccg gaagtccgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgcctc cgcccagtcc ttcatcgagc ggatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cactccctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcggaagcc cgccttcctg 1620

tccggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg gaaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgactc cgtggagatc 1740

tccggcgtgg aggaccggtt caacgcctcc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgggag atgatcgagg agcggctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcggc ggcggtacac cggctggggc 1980

cggctgtccc ggaagctgat caacggcatc cgggacaagc agtccggcaa gaccatcctg 2040

gacttcctga agtccgacgg cttcgccaac cggaacttca tgcagctgat ccacgacgac 2100

tccctgacct tcaaggagga catccagaag gcccaggtgt ccggccaggg cgactccctg 2160

cacgagcaca tcgccaacct ggccggctcc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccggc acaagcccga gaacatcgtg 2280

atcgagatgg cccgggagaa ccagaccacc cagaagggcc agaagaactc ccgggagcgg 2340

atgaagcgga tcgaggaggg catcaaggag ctgggctccc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgg 2460

gacatgtacg tggaccagga gctggacatc aaccggctgt ccgactacga cgtggaccac 2520

atcgtgcccc agtccttcct gaaggacgac tccatcgaca acaaggtgct gacccggtcc 2580

gacaagaacc ggggcaagtc cgacaacgtg ccctccgagg aggtggtgaa gaagatgaag 2640

aactactggc ggcagctgct gaacgccaag ctgatcaccc agcggaagtt cgacaacctg 2700

accaaggccg agcggggcgg cctgtccgag ctggacaagg ccggcttcat caagcggcag 2760

ctggtggaga cccggcagat caccaagcac gtggcccaga tcctggactc ccggatgaac 2820

accaagtacg acgagaacga caagctgatc cgggaggtga aggtgatcac cctgaagtcc 2880

aagctggtgt ccgacttccg gaaggacttc cagttctaca aggtgcggga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagtccga gttcgtgtac ggcgactaca aggtgtacga cgtgcggaag 3060

atgatcgcca agtccgagca ggagatcggc aaggccaccg ccaagtactt cttctactcc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccggaagcgg 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgggacttc 3240

gccaccgtgc ggaaggtgct gtccatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttctccaa ggagtccatc ctgcccaagc ggaactccga caagctgatc 3360

gcccggaaga aggactggga ccccaagaag tacggcggct tcgactcccc caccgtggcc 3420

tactccgtgc tggtggtggc caaggtggag aagggcaagt ccaagaagct gaagtccgtg 3480

aaggagctgc tgggcatcac catcatggag cggtcctcct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tactccctgt tcgagctgga gaacggccgg aagcggatgc tggcctccgc cggcgagctg 3660

cagaagggca acgagctggc cctgccctcc aagtacgtga acttcctgta cctggcctcc 3720

cactacgaga agctgaaggg ctcccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatct ccgagttctc caagcgggtg 3840

atcctggccg acgccaacct ggacaaggtg ctgtccgcct acaacaagca ccgggacaag 3900

cccatccggg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgga agcggtacac ctccaccaag 4020

gaggtgctgg acgccaccct gatccaccag tccatcaccg gcctgtacga gacccggatc 4080

gacctgtccc agctgggcgg cgactga 4107

<210> 81
<211> 4179
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 81

atggacaaga agtactccat cggcctggcc atcggcacca actccgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gccctccaag aagttcaagg tgctgggcaa caccgaccgg 120

cactccatca agaagaacct gatcggcgcc ctgctgttcg actccggcga gaccgccgag 180

gccacccggc tgaagcggac cgcccggcgg cggtacaccc ggcggaagaa ccggatctgc 240

tacctgcagg agatcttctc caacgagatg gccaaggtgg acgactcctt cttccaccgg 300

ctggaggagt ccttcctggt ggaggaggac aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcggaag 420

aagctggtgg actccaccga caaggccgac ctgcggctga tctacctggc cctggcccac 480

atgatcaagt tccggggcca cttcctgatc gagggcgacc tgaaccccga caactccgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcct ccggcgtgga cgccaaggcc atcctgtccg cccggctgtc caagtcccgg 660

cggctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgtccctggg cctgaccccc aacttcaagt ccaacttcga cctggccgag 780

gacgccaagc tgcagctgtc caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgtc cgacgccatc 900

ctgctgtccg acatcctgcg ggtgaacacc gagatcacca aggcccccct gtccgcctcc 960

atgatcaagc ggtacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgg 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agtccaagaa cggctacgcc 1080

ggctacatcg acggcggcgc ctcccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgggagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggctccatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc ggcggcagga ggacttctac cccttcctga aggacaaccg ggagaagatc 1320

gagaagatcc tgaccttccg gatcccctac tacgtgggcc ccctggcccg gggcaactcc 1380

cggttcgcct ggatgacccg gaagtccgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgcctc cgcccagtcc ttcatcgagc ggatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cactccctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcggaagcc cgccttcctg 1620

tccggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg gaaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgactc cgtggagatc 1740

tccggcgtgg aggaccggtt caacgcctcc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgggag atgatcgagg agcggctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcggc ggcggtacac cggctggggc 1980

cggctgtccc ggaagctgat caacggcatc cgggacaagc agtccggcaa gaccatcctg 2040

gacttcctga agtccgacgg cttcgccaac cggaacttca tgcagctgat ccacgacgac 2100

tccctgacct tcaaggagga catccagaag gcccaggtgt ccggccaggg cgactccctg 2160

cacgagcaca tcgccaacct ggccggctcc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccggc acaagcccga gaacatcgtg 2280

atcgagatgg cccgggagaa ccagaccacc cagaagggcc agaagaactc ccgggagcgg 2340

atgaagcgga tcgaggaggg catcaaggag ctgggctccc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgg 2460

gacatgtacg tggaccagga gctggacatc aaccggctgt ccgactacga cgtggaccac 2520

atcgtgcccc agtccttcct gaaggacgac tccatcgaca acaaggtgct gacccggtcc 2580

gacaagaacc ggggcaagtc cgacaacgtg ccctccgagg aggtggtgaa gaagatgaag 2640

aactactggc ggcagctgct gaacgccaag ctgatcaccc agcggaagtt cgacaacctg 2700

accaaggccg agcggggcgg cctgtccgag ctggacaagg ccggcttcat caagcggcag 2760

ctggtggaga cccggcagat caccaagcac gtggcccaga tcctggactc ccggatgaac 2820

accaagtacg acgagaacga caagctgatc cgggaggtga aggtgatcac cctgaagtcc 2880

aagctggtgt ccgacttccg gaaggacttc cagttctaca aggtgcggga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagtccga gttcgtgtac ggcgactaca aggtgtacga cgtgcggaag 3060

atgatcgcca agtccgagca ggagatcggc aaggccaccg ccaagtactt cttctactcc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccggaagcgg 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgggacttc 3240

gccaccgtgc ggaaggtgct gtccatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttctccaa ggagtccatc ctgcccaagc ggaactccga caagctgatc 3360

gcccggaaga aggactggga ccccaagaag tacggcggct tcgactcccc caccgtggcc 3420

tactccgtgc tggtggtggc caaggtggag aagggcaagt ccaagaagct gaagtccgtg 3480

aaggagctgc tgggcatcac catcatggag cggtcctcct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tactccctgt tcgagctgga gaacggccgg aagcggatgc tggcctccgc cggcgagctg 3660

cagaagggca acgagctggc cctgccctcc aagtacgtga acttcctgta cctggcctcc 3720

cactacgaga agctgaaggg ctcccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatct ccgagttctc caagcgggtg 3840

atcctggccg acgccaacct ggacaaggtg ctgtccgcct acaacaagca ccgggacaag 3900

cccatccggg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgga agcggtacac ctccaccaag 4020

gaggtgctgg acgccaccct gatccaccag tccatcaccg gcctgtacga gacccggatc 4080

gacctgtccc agctgggcgg cgacggctcc ggctccccca agaagaagcg gaaggtggac 4140

ggctccccca agaagaagcg gaaggtggac tccggctga 4179

<210> 82
<211> 4140
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 82

atggacaaga agtacagcat cggcctggcc atcggcacca acagcgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gcccagcaag aagttcaagg tgctgggcaa caccgaccgg 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag 180

gccacccggc tgaagcggac cgcccggcgg cggtacaccc ggcggaagaa ccggatctgc 240

tacctgcagg agatcttcag caacgagatg gccaaggtgg acgacagctt cttccaccgg 300

ctggaggaga gcttcctggt ggaggaggac aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcggaag 420

aagctggtgg acagcaccga caaggccgac ctgcggctga tctacctggc cctggcccac 480

atgatcaagt tccggggcca cttcctgatc gagggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg cccggctgag caagagccgg 660

cggctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgagcctggg cctgaccccc aacttcaaga gcaacttcga cctggccgag 780

gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgag cgacgccatc 900

ctgctgagcg acatcctgcg ggtgaacacc gagatcacca aggcccccct gagcgccagc 960

atgatcaagc ggtacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgg 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc 1080

ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgggagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggcagcatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc ggcggcagga ggacttctac cccttcctga aggacaaccg ggagaagatc 1320

gagaagatcc tgaccttccg gatcccctac tacgtgggcc ccctggcccg gggcaacagc 1380

cggttcgcct ggatgacccg gaagagcgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgccag cgcccagagc ttcatcgagc ggatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cacagcctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcggaagcc cgccttcctg 1620

agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg gaaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc 1740

agcggcgtgg aggaccggtt caacgccagc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgggag atgatcgagg agcggctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcggc ggcggtacac cggctggggc 1980

cggctgagcc ggaagctgat caacggcatc cgggacaagc agagcggcaa gaccatcctg 2040

gacttcctga agagcgacgg cttcgccaac cggaacttca tgcagctgat ccacgacgac 2100

agcctgacct tcaaggagga catccagaag gcccaggtga gcggccaggg cgacagcctg 2160

cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccggc acaagcccga gaacatcgtg 2280

atcgagatgg cccgggagaa ccagaccacc cagaagggcc agaagaacag ccgggagcgg 2340

atgaagcgga tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgg 2460

gacatgtacg tggaccagga gctggacatc aaccggctga gcgactacga cgtggaccac 2520

atcgtgcccc agagcttcct gaaggacgac agcatcgaca acaaggtgct gacccggagc 2580

gacaagaacc ggggcaagag cgacaacgtg cccagcgagg aggtggtgaa gaagatgaag 2640

aactactggc ggcagctgct gaacgccaag ctgatcaccc agcggaagtt cgacaacctg 2700

accaaggccg agcggggcgg cctgagcgag ctggacaagg ccggcttcat caagcggcag 2760

ctggtggaga cccggcagat caccaagcac gtggcccaga tcctggacag ccggatgaac 2820

accaagtacg acgagaacga caagctgatc cgggaggtga aggtgatcac cctgaagagc 2880

aagctggtga gcgacttccg gaaggacttc cagttctaca aggtgcggga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcggaag 3060

atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccggaagcgg 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgggacttc 3240

gccaccgtgc ggaaggtgct gagcatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttcagcaa ggagagcatc ctgcccaagc ggaacagcga caagctgatc 3360

gcccggaaga aggactggga ccccaagaag tacggcggct tcgacagccc caccgtggcc 3420

tacagcgtgc tggtggtggc caaggtggag aagggcaaga gcaagaagct gaagagcgtg 3480

aaggagctgc tgggcatcac catcatggag cggagcagct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tacagcctgt tcgagctgga gaacggccgg aagcggatgc tggccagcgc cggcgagctg 3660

cagaagggca acgagctggc cctgcccagc aagtacgtga acttcctgta cctggccagc 3720

cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgggtg 3840

atcctggccg acgccaacct ggacaaggtg ctgagcgcct acaacaagca ccgggacaag 3900

cccatccggg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgga agcggtacac cagcaccaag 4020

gaggtgctgg acgccaccct gatccaccag agcatcaccg gcctgtacga gacccggatc 4080

gacctgagcc agctgggcgg cgacggcggc ggcagcccca agaagaagcg gaaggtgtga 4140

<210> 83
<211> 4179
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 83
atggacaaga agtacagcat cggcctggcc atcggcacca acagcgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gcccagcaag aagttcaagg tgctgggcaa caccgaccgg 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag 180

gccacccggc tgaagcggac cgcccggcgg cggtacaccc ggcggaagaa ccggatctgc 240

tacctgcagg agatcttcag caacgagatg gccaaggtgg acgacagctt cttccaccgg 300

ctggaggaga gcttcctggt ggaggaggac aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcggaag 420

aagctggtgg acagcaccga caaggccgac ctgcggctga tctacctggc cctggcccac 480

atgatcaagt tccggggcca cttcctgatc gagggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg cccggctgag caagagccgg 660

cggctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgagcctggg cctgaccccc aacttcaaga gcaacttcga cctggccgag 780

gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgag cgacgccatc 900

ctgctgagcg acatcctgcg ggtgaacacc gagatcacca aggcccccct gagcgccagc 960

atgatcaagc ggtacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgg 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc 1080

ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgggagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggcagcatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc ggcggcagga ggacttctac cccttcctga aggacaaccg ggagaagatc 1320

gagaagatcc tgaccttccg gatcccctac tacgtgggcc ccctggcccg gggcaacagc 1380

cggttcgcct ggatgacccg gaagagcgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgccag cgcccagagc ttcatcgagc ggatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cacagcctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcggaagcc cgccttcctg 1620

agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg gaaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc 1740

agcggcgtgg aggaccggtt caacgccagc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgggag atgatcgagg agcggctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcggc ggcggtacac cggctggggc 1980

cggctgagcc ggaagctgat caacggcatc cgggacaagc agagcggcaa gaccatcctg 2040

gacttcctga agagcgacgg cttcgccaac cggaacttca tgcagctgat ccacgacgac 2100

agcctgacct tcaaggagga catccagaag gcccaggtga gcggccaggg cgacagcctg 2160

cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccggc acaagcccga gaacatcgtg 2280

atcgagatgg cccgggagaa ccagaccacc cagaagggcc agaagaacag ccgggagcgg 2340

atgaagcgga tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgg 2460

gacatgtacg tggaccagga gctggacatc aaccggctga gcgactacga cgtggaccac 2520

atcgtgcccc agagcttcct gaaggacgac agcatcgaca acaaggtgct gacccggagc 2580

gacaagaacc ggggcaagag cgacaacgtg cccagcgagg aggtggtgaa gaagatgaag 2640

aactactggc ggcagctgct gaacgccaag ctgatcaccc agcggaagtt cgacaacctg 2700

accaaggccg agcggggcgg cctgagcgag ctggacaagg ccggcttcat caagcggcag 2760

ctggtggaga cccggcagat caccaagcac gtggcccaga tcctggacag ccggatgaac 2820

accaagtacg acgagaacga caagctgatc cgggaggtga aggtgatcac cctgaagagc 2880

aagctggtga gcgacttccg gaaggacttc cagttctaca aggtgcggga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcggaag 3060

atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccggaagcgg 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgggacttc 3240

gccaccgtgc ggaaggtgct gagcatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttcagcaa ggagagcatc ctgcccaagc ggaacagcga caagctgatc 3360

gcccggaaga aggactggga ccccaagaag tacggcggct tcgacagccc caccgtggcc 3420

tacagcgtgc tggtggtggc caaggtggag aagggcaaga gcaagaagct gaagagcgtg 3480

aaggagctgc tgggcatcac catcatggag cggagcagct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tacagcctgt tcgagctgga gaacggccgg aagcggatgc tggccagcgc cggcgagctg 3660

cagaagggca acgagctggc cctgcccagc aagtacgtga acttcctgta cctggccagc 3720

cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgggtg 3840

atcctggccg acgccaacct ggacaaggtg ctgagcgcct acaacaagca ccgggacaag 3900

cccatccggg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgga agcggtacac cagcaccaag 4020

gaggtgctgg acgccaccct gatccaccag agcatcaccg gcctgtacga gacccggatc 4080

gacctgagcc agctgggcgg cgacggcagc ggcagcccca agaagaagcg gaaggtggac 4140

ggcagcccca agaagaagcg gaaggtggac agcggctga 4179

<210> 84
<211> 4107
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 84
atggacaaga agtacagcat cggcctggcc atcggcacca acagcgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gcccagcaag aagttcaagg tgctgggcaa caccgaccgg 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag 180

gccacccggc tgaagcggac cgcccggcgg cggtacaccc ggcggaagaa ccggatctgc 240

tacctgcagg agatcttcag caacgagatg gccaaggtgg acgacagctt cttccaccgg 300

ctggaggaga gcttcctggt ggaggaggac aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcggaag 420

aagctggtgg acagcaccga caaggccgac ctgcggctga tctacctggc cctggcccac 480

atgatcaagt tccggggcca cttcctgatc gagggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg cccggctgag caagagccgg 660

cggctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgagcctggg cctgaccccc aacttcaaga gcaacttcga cctggccgag 780

gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgag cgacgccatc 900

ctgctgagcg acatcctgcg ggtgaacacc gagatcacca aggcccccct gagcgccagc 960

atgatcaagc ggtacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgg 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc 1080

ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgggagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggcagcatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc ggcggcagga ggacttctac cccttcctga aggacaaccg ggagaagatc 1320

gagaagatcc tgaccttccg gatcccctac tacgtgggcc ccctggcccg gggcaacagc 1380

cggttcgcct ggatgacccg gaagagcgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgccag cgcccagagc ttcatcgagc ggatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cacagcctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcggaagcc cgccttcctg 1620

agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg gaaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc 1740

agcggcgtgg aggaccggtt caacgccagc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgggag atgatcgagg agcggctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcggc ggcggtacac cggctggggc 1980

cggctgagcc ggaagctgat caacggcatc cgggacaagc agagcggcaa gaccatcctg 2040

gacttcctga agagcgacgg cttcgccaac cggaacttca tgcagctgat ccacgacgac 2100

agcctgacct tcaaggagga catccagaag gcccaggtga gcggccaggg cgacagcctg 2160

cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccggc acaagcccga gaacatcgtg 2280

atcgagatgg cccgggagaa ccagaccacc cagaagggcc agaagaacag ccgggagcgg 2340

atgaagcgga tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgg 2460

gacatgtacg tggaccagga gctggacatc aaccggctga gcgactacga cgtggaccac 2520

atcgtgcccc agagcttcct gaaggacgac agcatcgaca acaaggtgct gacccggagc 2580

gacaagaacc ggggcaagag cgacaacgtg cccagcgagg aggtggtgaa gaagatgaag 2640

aactactggc ggcagctgct gaacgccaag ctgatcaccc agcggaagtt cgacaacctg 2700

accaaggccg agcggggcgg cctgagcgag ctggacaagg ccggcttcat caagcggcag 2760

ctggtggaga cccggcagat caccaagcac gtggcccaga tcctggacag ccggatgaac 2820

accaagtacg acgagaacga caagctgatc cgggaggtga aggtgatcac cctgaagagc 2880

aagctggtga gcgacttccg gaaggacttc cagttctaca aggtgcggga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcggaag 3060

atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccggaagcgg 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgggacttc 3240

gccaccgtgc ggaaggtgct gagcatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttcagcaa ggagagcatc ctgcccaagc ggaacagcga caagctgatc 3360

gcccggaaga aggactggga ccccaagaag tacggcggct tcgacagccc caccgtggcc 3420

tacagcgtgc tggtggtggc caaggtggag aagggcaaga gcaagaagct gaagagcgtg 3480

aaggagctgc tgggcatcac catcatggag cggagcagct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tacagcctgt tcgagctgga gaacggccgg aagcggatgc tggccagcgc cggcgagctg 3660

cagaagggca acgagctggc cctgcccagc aagtacgtga acttcctgta cctggccagc 3720

cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgggtg 3840

atcctggccg acgccaacct ggacaaggtg ctgagcgcct acaacaagca ccgggacaag 3900

cccatccggg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgga agcggtacac cagcaccaag 4020

gaggtgctgg acgccaccct gatccaccag agcatcaccg gcctgtacga gacccggatc 4080

gacctgagcc agctgggcgg cgactga 4107

<210> 85
<211> 4134
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 85
gacaagaagt actccatcgg cctggccatc ggcaccaact ccgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc ctccaagaag ttcaaggtgc tgggcaacac cgaccggcac 120

tccatcaaga agaacctgat cggcgccctg ctgttcgact ccggcgagac cgccgaggcc 180

acccggctga agcggaccgc ccggcggcgg tacacccggc ggaagaaccg gatctgctac 240

ctgcaggaga tcttctccaa cgagatggcc aaggtggacg actccttctt ccaccggctg 300

gaggagtcct tcctggtgga ggaggacaag aagcacgagc ggcaccccat cttcggcaac 360

atcgtggacg aggtggccta ccacgagaag taccccacca tctaccacct gcggaagaag 420

ctggtggact ccaccgacaa ggccgacctg cggctgatct acctggccct ggcccacatg 480

atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa ctccgacgtg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga gaaccccatc 600

aacgcctccg gcgtggacgc caaggccatc ctgtccgccc ggctgtccaa gtcccggcgg 660

ctggagaacc tgatcgccca gctgcccggc gagaagaaga acggcctgtt cggcaacctg 720

atcgccctgt ccctgggcct gacccccaac ttcaagtcca acttcgacct ggccgaggac 780

gccaagctgc agctgtccaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacgccga cctgttcctg gccgccaaga acctgtccga cgccatcctg 900

ctgtccgaca tcctgcgggt gaacaccgag atcaccaagg ccccccctgtc cgcctccatg 960

atcaagcggt acgacgagca ccaccaggac ctgaccctgc tgaaggccct ggtgcggcag 1020

cagctgcccg agaagtacaa ggagatcttc ttcgaccagt ccaagaacgg ctacgccggc 1080

tacatcgacg gcggcgcctc ccaggaggag ttctacaagt tcatcaagcc catcctggag 1140

aagatggacg gcaccgagga gctgctggtg aagctgaacc gggaggacct gctgcggaag 1200

cagcggacct tcgacaacgg ctccatcccc caccagatcc acctgggcga gctgcacgcc 1260

atcctgcggc ggcaggagga cttctacccc ttcctgaagg acaaccggga gaagatcgag 1320

aagatcctga ccttccggat cccctactac gtgggccccc tggcccgggg caactcccgg 1380

ttcgcctgga tgacccggaa gtccgaggag accatcaccc cctggaactt cgaggaggtg 1440

gtggacaagg gcgcctccgc ccagtccttc atcgagcgga tgaccaactt cgacaagaac 1500

ctgcccaacg agaaggtgct gcccaagcac tccctgctgt acgagtactt caccgtgtac 1560

aacgagctga ccaaggtgaa gtacgtgacc gagggcatgc ggaagcccgc cttcctgtcc 1620

ggcgagcaga agaaggccat cgtggacctg ctgttcaaga ccaaccggaa ggtgaccgtg 1680

aagcagctga aggaggacta cttcaagaag atcgagtgct tcgactccgt ggagatctcc 1740

ggcgtggagg accggttcaa cgcctccctg ggcacctacc acgacctgct gaagatcatc 1800

aaggacaagg acttcctgga caacgaggag aacgaggaca tcctggagga catcgtgctg 1860

accctgaccc tgttcgagga ccgggagatg atcgaggagc ggctgaagac ctacgcccac 1920

ctgttcgacg acaaggtgat gaagcagctg aagcggcggc ggtacaccgg ctggggccgg 1980

ctgtcccgga agctgatcaa cggcatccgg gacaagcagt ccggcaagac catcctggac 2040

ttcctgaagt ccgacggctt cgccaaccgg aacttcatgc agctgatcca cgacgactcc 2100

ctgaccttca aggaggacat ccagaaggcc caggtgtccg gccagggcga ctccctgcac 2160

gagcacatcg ccaacctggc cggctccccc gccatcaaga agggcatcct gcagaccgtg 2220

aaggtggtgg acgagctggt gaaggtgatg ggccggcaca agcccgagaa catcgtgatc 2280

gagatggccc gggagaacca gaccacccag aagggccaga agaactcccg ggagcggatg 2340

aagcggatcg aggagggcat caaggagctg ggctcccaga tcctgaagga gcaccccgtg 2400

gagaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa cggccgggac 2460

atgtacgtgg accaggagct ggacatcaac cggctgtccg actacgacgt ggaccacatc 2520

gtgccccagt ccttcctgaa ggacgactcc atcgacaaca aggtgctgac ccggtccgac 2580

aagaaccggg gcaagtccga caacgtgccc tccgaggagg tggtgaagaa gatgaagaac 2640

tactggcggc agctgctgaa cgccaagctg atcacccagc ggaagttcga caacctgacc 2700

aaggccgagc ggggcggcct gtccgagctg gacaaggccg gcttcatcaa gcggcagctg 2760

gtggagaccc ggcagatcac caagcacgtg gcccagatcc tggactcccg gatgaacacc 2820

aagtacgacg agaacgacaa gctgatccgg gaggtgaagg tgatcaccct gaagtccaag 2880

ctggtgtccg acttccggaa ggacttccag ttctacaagg tgcgggagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtggtgggca ccgccctgat caagaagtac 3000

cccaagctgg agtccgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060

atcgccaagt ccgagcagga gatcggcaag gccaccgcca agtacttctt ctactccaac 3120

atcatgaact tcttcaagac cgagatcacc ctggccaacg gcgagatccg gaagcggccc 3180

ctgatcgaga ccaacggcga gaccggcgag atcgtgtggg acaagggccg ggacttcgcc 3240

accgtgcgga aggtgctgtc catgccccag gtgaacatcg tgaagaagac cgaggtgcag 3300

accggcggct tctccaagga gtccatcctg cccaagcgga actccgacaa gctgatcgcc 3360

cggaagaagg actgggaccc caagaagtac ggcggcttcg actcccccac cgtggcctac 3420

tccgtgctgg tggtggccaa ggtggagaag ggcaagtcca agaagctgaa gtccgtgaag 3480

gagctgctgg gcatcaccat catggagcgg tcctccttcg agaagaaccc catcgacttc 3540

ctggaggcca agggctacaa ggaggtgaag aaggacctga tcatcaagct gcccaagtac 3600

tccctgttcg agctggagaa cggccggaag cggatgctgg cctccgccgg cgagctgcag 3660

aagggcaacg agctggccct gccctccaag tacgtgaact tcctgtacct ggcctcccac 3720

tacgagaagc tgaagggctc ccccgaggac aacgagcaga agcagctgtt cgtggagcag 3780

cacaagcact acctggacga gatcatcgag cagatctccg agttctccaa gcgggtgatc 3840

ctggccgacg ccaacctgga caaggtgctg tccgcctaca acaagcaccg ggacaagccc 3900

atccgggagc aggccgagaa catcatccac ctgttcaccc tgaccaacct gggcgccccc 3960

gccgccttca agtacttcga caccaccatc gaccggaagc ggtacacctc caccaaggag 4020

gtgctggacg ccaccctgat ccaccagtcc atcaccggcc tgtacgagac ccggatcgac 4080

ctgtcccagc tgggcggcga cggcggcggc tcccccaaga agaagcggaa ggtg 4134

<210> 86
<211> 4101
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 86
gacaagaagt actccatcgg cctggccatc ggcaccaact ccgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc ctccaagaag ttcaaggtgc tgggcaacac cgaccggcac 120

tccatcaaga agaacctgat cggcgccctg ctgttcgact ccggcgagac cgccgaggcc 180

acccggctga agcggaccgc ccggcggcgg tacacccggc ggaagaaccg gatctgctac 240

ctgcaggaga tcttctccaa cgagatggcc aaggtggacg actccttctt ccaccggctg 300

gaggagtcct tcctggtgga ggaggacaag aagcacgagc ggcaccccat cttcggcaac 360

atcgtggacg aggtggccta ccacgagaag taccccacca tctaccacct gcggaagaag 420

ctggtggact ccaccgacaa ggccgacctg cggctgatct acctggccct ggcccacatg 480

atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa ctccgacgtg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga gaaccccatc 600

aacgcctccg gcgtggacgc caaggccatc ctgtccgccc ggctgtccaa gtcccggcgg 660

ctggagaacc tgatcgccca gctgcccggc gagaagaaga acggcctgtt cggcaacctg 720

atcgccctgt ccctgggcct gacccccaac ttcaagtcca acttcgacct ggccgaggac 780

gccaagctgc agctgtccaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacgccga cctgttcctg gccgccaaga acctgtccga cgccatcctg 900

ctgtccgaca tcctgcgggt gaacaccgag atcaccaagg ccccccctgtc cgcctccatg 960

atcaagcggt acgacgagca ccaccaggac ctgaccctgc tgaaggccct ggtgcggcag 1020

cagctgcccg agaagtacaa ggagatcttc ttcgaccagt ccaagaacgg ctacgccggc 1080

tacatcgacg gcggcgcctc ccaggaggag ttctacaagt tcatcaagcc catcctggag 1140

aagatggacg gcaccgagga gctgctggtg aagctgaacc gggaggacct gctgcggaag 1200

cagcggacct tcgacaacgg ctccatcccc caccagatcc acctgggcga gctgcacgcc 1260

atcctgcggc ggcaggagga cttctacccc ttcctgaagg acaaccggga gaagatcgag 1320

aagatcctga ccttccggat cccctactac gtgggccccc tggcccgggg caactcccgg 1380

ttcgcctgga tgacccggaa gtccgaggag accatcaccc cctggaactt cgaggaggtg 1440

gtggacaagg gcgcctccgc ccagtccttc atcgagcgga tgaccaactt cgacaagaac 1500

ctgcccaacg agaaggtgct gcccaagcac tccctgctgt acgagtactt caccgtgtac 1560

aacgagctga ccaaggtgaa gtacgtgacc gagggcatgc ggaagcccgc cttcctgtcc 1620

ggcgagcaga agaaggccat cgtggacctg ctgttcaaga ccaaccggaa ggtgaccgtg 1680

aagcagctga aggaggacta cttcaagaag atcgagtgct tcgactccgt ggagatctcc 1740

ggcgtggagg accggttcaa cgcctccctg ggcacctacc acgacctgct gaagatcatc 1800

aaggacaagg acttcctgga caacgaggag aacgaggaca tcctggagga catcgtgctg 1860

accctgaccc tgttcgagga ccgggagatg atcgaggagc ggctgaagac ctacgcccac 1920

ctgttcgacg acaaggtgat gaagcagctg aagcggcggc ggtacaccgg ctggggccgg 1980

ctgtcccgga agctgatcaa cggcatccgg gacaagcagt ccggcaagac catcctggac 2040

ttcctgaagt ccgacggctt cgccaaccgg aacttcatgc agctgatcca cgacgactcc 2100

ctgaccttca aggaggacat ccagaaggcc caggtgtccg gccagggcga ctccctgcac 2160

gagcacatcg ccaacctggc cggctccccc gccatcaaga agggcatcct gcagaccgtg 2220

aaggtggtgg acgagctggt gaaggtgatg ggccggcaca agcccgagaa catcgtgatc 2280

gagatggccc gggagaacca gaccacccag aagggccaga agaactcccg ggagcggatg 2340

aagcggatcg aggagggcat caaggagctg ggctcccaga tcctgaagga gcaccccgtg 2400

gagaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa cggccgggac 2460

atgtacgtgg accaggagct ggacatcaac cggctgtccg actacgacgt ggaccacatc 2520

gtgccccagt ccttcctgaa ggacgactcc atcgacaaca aggtgctgac ccggtccgac 2580

aagaaccggg gcaagtccga caacgtgccc tccgaggagg tggtgaagaa gatgaagaac 2640

tactggcggc agctgctgaa cgccaagctg atcacccagc ggaagttcga caacctgacc 2700

aaggccgagc ggggcggcct gtccgagctg gacaaggccg gcttcatcaa gcggcagctg 2760

gtggagaccc ggcagatcac caagcacgtg gcccagatcc tggactcccg gatgaacacc 2820

aagtacgacg agaacgacaa gctgatccgg gaggtgaagg tgatcaccct gaagtccaag 2880

ctggtgtccg acttccggaa ggacttccag ttctacaagg tgcgggagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtggtgggca ccgccctgat caagaagtac 3000

cccaagctgg agtccgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060

atcgccaagt ccgagcagga gatcggcaag gccaccgcca agtacttctt ctactccaac 3120

atcatgaact tcttcaagac cgagatcacc ctggccaacg gcgagatccg gaagcggccc 3180

ctgatcgaga ccaacggcga gaccggcgag atcgtgtggg acaagggccg ggacttcgcc 3240

accgtgcgga aggtgctgtc catgccccag gtgaacatcg tgaagaagac cgaggtgcag 3300

accggcggct tctccaagga gtccatcctg cccaagcgga actccgacaa gctgatcgcc 3360

cggaagaagg actgggaccc caagaagtac ggcggcttcg actcccccac cgtggcctac 3420

tccgtgctgg tggtggccaa ggtggagaag ggcaagtcca agaagctgaa gtccgtgaag 3480

gagctgctgg gcatcaccat catggagcgg tcctccttcg agaagaaccc catcgacttc 3540

ctggaggcca agggctacaa ggaggtgaag aaggacctga tcatcaagct gcccaagtac 3600

tccctgttcg agctggagaa cggccggaag cggatgctgg cctccgccgg cgagctgcag 3660

aagggcaacg agctggccct gccctccaag tacgtgaact tcctgtacct ggcctcccac 3720

tacgagaagc tgaagggctc ccccgaggac aacgagcaga agcagctgtt cgtggagcag 3780

cacaagcact acctggacga gatcatcgag cagatctccg agttctccaa gcgggtgatc 3840

ctggccgacg ccaacctgga caaggtgctg tccgcctaca acaagcaccg ggacaagccc 3900

atccgggagc aggccgagaa catcatccac ctgttcaccc tgaccaacct gggcgccccc 3960

gccgccttca agtacttcga caccaccatc gaccggaagc ggtacacctc caccaaggag 4020

gtgctggacg ccaccctgat ccaccagtcc atcaccggcc tgtacgagac ccggatcgac 4080

ctgtcccagc tgggcggcga c 4101

<210> 87
<211> 4173
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 87
gacaagaagt actccatcgg cctggccatc ggcaccaact ccgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc ctccaagaag ttcaaggtgc tgggcaacac cgaccggcac 120

tccatcaaga agaacctgat cggcgccctg ctgttcgact ccggcgagac cgccgaggcc 180

acccggctga agcggaccgc ccggcggcgg tacacccggc ggaagaaccg gatctgctac 240

ctgcaggaga tcttctccaa cgagatggcc aaggtggacg actccttctt ccaccggctg 300

gaggagtcct tcctggtgga ggaggacaag aagcacgagc ggcaccccat cttcggcaac 360

atcgtggacg aggtggccta ccacgagaag taccccacca tctaccacct gcggaagaag 420

ctggtggact ccaccgacaa ggccgacctg cggctgatct acctggccct ggcccacatg 480

atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa ctccgacgtg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga gaaccccatc 600

aacgcctccg gcgtggacgc caaggccatc ctgtccgccc ggctgtccaa gtcccggcgg 660

ctggagaacc tgatcgccca gctgcccggc gagaagaaga acggcctgtt cggcaacctg 720

atcgccctgt ccctgggcct gacccccaac ttcaagtcca acttcgacct ggccgaggac 780

gccaagctgc agctgtccaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacgccga cctgttcctg gccgccaaga acctgtccga cgccatcctg 900

ctgtccgaca tcctgcgggt gaacaccgag atcaccaagg ccccccctgtc cgcctccatg 960

atcaagcggt acgacgagca ccaccaggac ctgaccctgc tgaaggccct ggtgcggcag 1020

cagctgcccg agaagtacaa ggagatcttc ttcgaccagt ccaagaacgg ctacgccggc 1080

tacatcgacg gcggcgcctc ccaggaggag ttctacaagt tcatcaagcc catcctggag 1140

aagatggacg gcaccgagga gctgctggtg aagctgaacc gggaggacct gctgcggaag 1200

cagcggacct tcgacaacgg ctccatcccc caccagatcc acctgggcga gctgcacgcc 1260

atcctgcggc ggcaggagga cttctacccc ttcctgaagg acaaccggga gaagatcgag 1320

aagatcctga ccttccggat cccctactac gtgggccccc tggcccgggg caactcccgg 1380

ttcgcctgga tgacccggaa gtccgaggag accatcaccc cctggaactt cgaggaggtg 1440

gtggacaagg gcgcctccgc ccagtccttc atcgagcgga tgaccaactt cgacaagaac 1500

ctgcccaacg agaaggtgct gcccaagcac tccctgctgt acgagtactt caccgtgtac 1560

aacgagctga ccaaggtgaa gtacgtgacc gagggcatgc ggaagcccgc cttcctgtcc 1620

ggcgagcaga agaaggccat cgtggacctg ctgttcaaga ccaaccggaa ggtgaccgtg 1680

aagcagctga aggaggacta cttcaagaag atcgagtgct tcgactccgt ggagatctcc 1740

ggcgtggagg accggttcaa cgcctccctg ggcacctacc acgacctgct gaagatcatc 1800

aaggacaagg acttcctgga caacgaggag aacgaggaca tcctggagga catcgtgctg 1860

accctgaccc tgttcgagga ccgggagatg atcgaggagc ggctgaagac ctacgcccac 1920

ctgttcgacg acaaggtgat gaagcagctg aagcggcggc ggtacaccgg ctggggccgg 1980

ctgtcccgga agctgatcaa cggcatccgg gacaagcagt ccggcaagac catcctggac 2040

ttcctgaagt ccgacggctt cgccaaccgg aacttcatgc agctgatcca cgacgactcc 2100

ctgaccttca aggaggacat ccagaaggcc caggtgtccg gccagggcga ctccctgcac 2160

gagcacatcg ccaacctggc cggctccccc gccatcaaga agggcatcct gcagaccgtg 2220

aaggtggtgg acgagctggt gaaggtgatg ggccggcaca agcccgagaa catcgtgatc 2280

gagatggccc gggagaacca gaccacccag aagggccaga agaactcccg ggagcggatg 2340

aagcggatcg aggagggcat caaggagctg ggctcccaga tcctgaagga gcaccccgtg 2400

gagaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa cggccgggac 2460

atgtacgtgg accaggagct ggacatcaac cggctgtccg actacgacgt ggaccacatc 2520

gtgccccagt ccttcctgaa ggacgactcc atcgacaaca aggtgctgac ccggtccgac 2580

aagaaccggg gcaagtccga caacgtgccc tccgaggagg tggtgaagaa gatgaagaac 2640

tactggcggc agctgctgaa cgccaagctg atcacccagc ggaagttcga caacctgacc 2700

aaggccgagc ggggcggcct gtccgagctg gacaaggccg gcttcatcaa gcggcagctg 2760

gtggagaccc ggcagatcac caagcacgtg gcccagatcc tggactcccg gatgaacacc 2820

aagtacgacg agaacgacaa gctgatccgg gaggtgaagg tgatcaccct gaagtccaag 2880

ctggtgtccg acttccggaa ggacttccag ttctacaagg tgcgggagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtggtgggca ccgccctgat caagaagtac 3000

cccaagctgg agtccgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060

atcgccaagt ccgagcagga gatcggcaag gccaccgcca agtacttctt ctactccaac 3120

atcatgaact tcttcaagac cgagatcacc ctggccaacg gcgagatccg gaagcggccc 3180

ctgatcgaga ccaacggcga gaccggcgag atcgtgtggg acaagggccg ggacttcgcc 3240

accgtgcgga aggtgctgtc catgccccag gtgaacatcg tgaagaagac cgaggtgcag 3300

accggcggct tctccaagga gtccatcctg cccaagcgga actccgacaa gctgatcgcc 3360

cggaagaagg actgggaccc caagaagtac ggcggcttcg actcccccac cgtggcctac 3420

tccgtgctgg tggtggccaa ggtggagaag ggcaagtcca agaagctgaa gtccgtgaag 3480

gagctgctgg gcatcaccat catggagcgg tcctccttcg agaagaaccc catcgacttc 3540

ctggaggcca agggctacaa ggaggtgaag aaggacctga tcatcaagct gcccaagtac 3600

tccctgttcg agctggagaa cggccggaag cggatgctgg cctccgccgg cgagctgcag 3660

aagggcaacg agctggccct gccctccaag tacgtgaact tcctgtacct ggcctcccac 3720

tacgagaagc tgaagggctc ccccgaggac aacgagcaga agcagctgtt cgtggagcag 3780

cacaagcact acctggacga gatcatcgag cagatctccg agttctccaa gcgggtgatc 3840

ctggccgacg ccaacctgga caaggtgctg tccgcctaca acaagcaccg ggacaagccc 3900

atccgggagc aggccgagaa catcatccac ctgttcaccc tgaccaacct gggcgccccc 3960

gccgccttca agtacttcga caccaccatc gaccggaagc ggtacacctc caccaaggag 4020

gtgctggacg ccaccctgat ccaccagtcc atcaccggcc tgtacgagac ccggatcgac 4080

ctgtcccagc tgggcggcga cggctccggc tcccccaaga agaagcggaa ggtggacggc 4140

tcccccaaga agaagcggaa ggtggactcc ggc 4173

<210> 88
<211> 4134
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 88
gacaagaagt acagcatcgg cctggccatc ggcaccaaca gcgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc cagcaagaag ttcaaggtgc tgggcaacac cgaccggcac 120

agcatcaaga agaacctgat cggcgccctg ctgttcgaca gcggcgagac cgccgaggcc 180

acccggctga agcggaccgc ccggcggcgg tacacccggc ggaagaaccg gatctgctac 240

ctgcaggaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt ccaccggctg 300

gaggagagct tcctggtgga ggaggacaag aagcacgagc ggcaccccat cttcggcaac 360

atcgtggacg aggtggccta ccacgagaag taccccacca tctaccacct gcggaagaag 420

ctggtggaca gcaccgacaa ggccgacctg cggctgatct acctggccct ggcccacatg 480

atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa cagcgacgtg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga gaaccccatc 600

aacgccagcg gcgtggacgc caaggccatc ctgagcgccc ggctgagcaa gagccggcgg 660

ctggagaacc tgatcgccca gctgcccggc gagaagaaga acggcctgtt cggcaacctg 720

atcgccctga gcctgggcct gacccccaac ttcaagagca acttcgacct ggccgaggac 780

gccaagctgc agctgagcaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacgccga cctgttcctg gccgccaaga acctgagcga cgccatcctg 900

ctgagcgaca tcctgcgggt gaacaccgag atcaccaagg ccccccctgag cgccagcatg 960

atcaagcggt acgacgagca ccaccaggac ctgaccctgc tgaaggccct ggtgcggcag 1020

cagctgcccg agaagtacaa ggagatcttc ttcgaccaga gcaagaacgg ctacgccggc 1080

tacatcgacg gcggcgccag ccaggaggag ttctacaagt tcatcaagcc catcctggag 1140

aagatggacg gcaccgagga gctgctggtg aagctgaacc gggaggacct gctgcggaag 1200

cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggcga gctgcacgcc 1260

atcctgcggc ggcaggagga cttctacccc ttcctgaagg acaaccggga gaagatcgag 1320

aagatcctga ccttccggat cccctactac gtgggccccc tggcccgggg caacagccgg 1380

ttcgcctgga tgacccggaa gagcgaggag accatcaccc cctggaactt cgaggaggtg 1440

gtggacaagg gcgccagcgc ccagagcttc atcgagcgga tgaccaactt cgacaagaac 1500

ctgcccaacg agaaggtgct gcccaagcac agcctgctgt acgagtactt caccgtgtac 1560

aacgagctga ccaaggtgaa gtacgtgacc gagggcatgc ggaagcccgc cttcctgagc 1620

ggcgagcaga agaaggccat cgtggacctg ctgttcaaga ccaaccggaa ggtgaccgtg 1680

aagcagctga aggaggacta cttcaagaag atcgagtgct tcgacagcgt ggagatcagc 1740

ggcgtggagg accggttcaa cgccagcctg ggcacctacc acgacctgct gaagatcatc 1800

aaggacaagg acttcctgga caacgaggag aacgaggaca tcctggagga catcgtgctg 1860

accctgaccc tgttcgagga ccgggagatg atcgaggagc ggctgaagac ctacgcccac 1920

ctgttcgacg acaaggtgat gaagcagctg aagcggcggc ggtacaccgg ctggggccgg 1980

ctgagccgga agctgatcaa cggcatccgg gacaagcaga gcggcaagac catcctggac 2040

ttcctgaaga gcgacggctt cgccaaccgg aacttcatgc agctgatcca cgacgacagc 2100

ctgaccttca aggaggacat ccagaaggcc caggtgagcg gccagggcga cagcctgcac 2160

gagcacatcg ccaacctggc cggcagcccc gccatcaaga agggcatcct gcagaccgtg 2220

aaggtggtgg acgagctggt gaaggtgatg ggccggcaca agcccgagaa catcgtgatc 2280

gagatggccc gggagaacca gaccacccag aagggccaga agaacagccg ggagcggatg 2340

aagcggatcg aggagggcat caaggagctg ggcagccaga tcctgaagga gcaccccgtg 2400

gagaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa cggccgggac 2460

atgtacgtgg accaggagct ggacatcaac cggctgagcg actacgacgt ggaccacatc 2520

gtgccccaga gcttcctgaa ggacgacagc atcgacaaca aggtgctgac ccggagcgac 2580

aagaaccggg gcaagagcga caacgtgccc agcgaggagg tggtgaagaa gatgaagaac 2640

tactggcggc agctgctgaa cgccaagctg atcacccagc ggaagttcga caacctgacc 2700

aaggccgagc ggggcggcct gagcgagctg gacaaggccg gcttcatcaa gcggcagctg 2760

gtggagaccc ggcagatcac caagcacgtg gcccagatcc tggacagccg gatgaacacc 2820

aagtacgacg agaacgacaa gctgatccgg gaggtgaagg tgatcaccct gaagagcaag 2880

ctggtgagcg acttccggaa ggacttccag ttctacaagg tgcgggagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtggtgggca ccgccctgat caagaagtac 3000

cccaagctgg agagcgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060

atcgccaaga gcgagcagga gatcggcaag gccaccgcca agtacttctt ctacagcaac 3120

atcatgaact tcttcaagac cgagatcacc ctggccaacg gcgagatccg gaagcggccc 3180

ctgatcgaga ccaacggcga gaccggcgag atcgtgtggg acaagggccg ggacttcgcc 3240

accgtgcgga aggtgctgag catgccccag gtgaacatcg tgaagaagac cgaggtgcag 3300

accggcggct tcagcaagga gagcatcctg cccaagcgga acagcgacaa gctgatcgcc 3360

cggaagaagg actgggaccc caagaagtac ggcggcttcg acagccccac cgtggcctac 3420

agcgtgctgg tggtggccaa ggtggagaag ggcaagagca agaagctgaa gagcgtgaag 3480

gagctgctgg gcatcaccat catggagcgg agcagcttcg agaagaaccc catcgacttc 3540

ctggaggcca agggctacaa ggaggtgaag aaggacctga tcatcaagct gcccaagtac 3600

agcctgttcg agctggagaa cggccggaag cggatgctgg ccagcgccgg cgagctgcag 3660

aagggcaacg agctggccct gcccagcaag tacgtgaact tcctgtacct ggccagccac 3720

tacgagaagc tgaagggcag ccccgaggac aacgagcaga agcagctgtt cgtggagcag 3780

cacaagcact acctggacga gatcatcgag cagatcagcg agttcagcaa gcgggtgatc 3840

ctggccgacg ccaacctgga caaggtgctg agcgcctaca acaagcaccg ggacaagccc 3900

atccgggagc aggccgagaa catcatccac ctgttcaccc tgaccaacct gggcgccccc 3960

gccgccttca agtacttcga caccaccatc gaccggaagc ggtacaccag caccaaggag 4020

gtgctggacg ccaccctgat ccaccagagc atcaccggcc tgtacgagac ccggatcgac 4080

ctgagccagc tgggcggcga cggcggcggc agccccaaga agaagcggaa ggtg 4134

<210> 89
<211> 4173
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 89
gacaagaagt acagcatcgg cctggccatc ggcaccaaca gcgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc cagcaagaag ttcaaggtgc tgggcaacac cgaccggcac 120

agcatcaaga agaacctgat cggcgccctg ctgttcgaca gcggcgagac cgccgaggcc 180

acccggctga agcggaccgc ccggcggcgg tacacccggc ggaagaaccg gatctgctac 240

ctgcaggaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt ccaccggctg 300

gaggagagct tcctggtgga ggaggacaag aagcacgagc ggcaccccat cttcggcaac 360

atcgtggacg aggtggccta ccacgagaag taccccacca tctaccacct gcggaagaag 420

ctggtggaca gcaccgacaa ggccgacctg cggctgatct acctggccct ggcccacatg 480

atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa cagcgacgtg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga gaaccccatc 600

aacgccagcg gcgtggacgc caaggccatc ctgagcgccc ggctgagcaa gagccggcgg 660

ctggagaacc tgatcgccca gctgcccggc gagaagaaga acggcctgtt cggcaacctg 720

atcgccctga gcctgggcct gacccccaac ttcaagagca acttcgacct ggccgaggac 780

gccaagctgc agctgagcaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacgccga cctgttcctg gccgccaaga acctgagcga cgccatcctg 900

ctgagcgaca tcctgcgggt gaacaccgag atcaccaagg ccccccctgag cgccagcatg 960

atcaagcggt acgacgagca ccaccaggac ctgaccctgc tgaaggccct ggtgcggcag 1020

cagctgcccg agaagtacaa ggagatcttc ttcgaccaga gcaagaacgg ctacgccggc 1080

tacatcgacg gcggcgccag ccaggaggag ttctacaagt tcatcaagcc catcctggag 1140

aagatggacg gcaccgagga gctgctggtg aagctgaacc gggaggacct gctgcggaag 1200

cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggcga gctgcacgcc 1260

atcctgcggc ggcaggagga cttctacccc ttcctgaagg acaaccggga gaagatcgag 1320

aagatcctga ccttccggat cccctactac gtgggccccc tggcccgggg caacagccgg 1380

ttcgcctgga tgacccggaa gagcgaggag accatcaccc cctggaactt cgaggaggtg 1440

gtggacaagg gcgccagcgc ccagagcttc atcgagcgga tgaccaactt cgacaagaac 1500

ctgcccaacg agaaggtgct gcccaagcac agcctgctgt acgagtactt caccgtgtac 1560

aacgagctga ccaaggtgaa gtacgtgacc gagggcatgc ggaagcccgc cttcctgagc 1620

ggcgagcaga agaaggccat cgtggacctg ctgttcaaga ccaaccggaa ggtgaccgtg 1680

aagcagctga aggaggacta cttcaagaag atcgagtgct tcgacagcgt ggagatcagc 1740

ggcgtggagg accggttcaa cgccagcctg ggcacctacc acgacctgct gaagatcatc 1800

aaggacaagg acttcctgga caacgaggag aacgaggaca tcctggagga catcgtgctg 1860

accctgaccc tgttcgagga ccgggagatg atcgaggagc ggctgaagac ctacgcccac 1920

ctgttcgacg acaaggtgat gaagcagctg aagcggcggc ggtacaccgg ctggggccgg 1980

ctgagccgga agctgatcaa cggcatccgg gacaagcaga gcggcaagac catcctggac 2040

ttcctgaaga gcgacggctt cgccaaccgg aacttcatgc agctgatcca cgacgacagc 2100

ctgaccttca aggaggacat ccagaaggcc caggtgagcg gccagggcga cagcctgcac 2160

gagcacatcg ccaacctggc cggcagcccc gccatcaaga agggcatcct gcagaccgtg 2220

aaggtggtgg acgagctggt gaaggtgatg ggccggcaca agcccgagaa catcgtgatc 2280

gagatggccc gggagaacca gaccacccag aagggccaga agaacagccg ggagcggatg 2340

aagcggatcg aggagggcat caaggagctg ggcagccaga tcctgaagga gcaccccgtg 2400

gagaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa cggccgggac 2460

atgtacgtgg accaggagct ggacatcaac cggctgagcg actacgacgt ggaccacatc 2520

gtgccccaga gcttcctgaa ggacgacagc atcgacaaca aggtgctgac ccggagcgac 2580

aagaaccggg gcaagagcga caacgtgccc agcgaggagg tggtgaagaa gatgaagaac 2640

tactggcggc agctgctgaa cgccaagctg atcacccagc ggaagttcga caacctgacc 2700

aaggccgagc ggggcggcct gagcgagctg gacaaggccg gcttcatcaa gcggcagctg 2760

gtggagaccc ggcagatcac caagcacgtg gcccagatcc tggacagccg gatgaacacc 2820

aagtacgacg agaacgacaa gctgatccgg gaggtgaagg tgatcaccct gaagagcaag 2880

ctggtgagcg acttccggaa ggacttccag ttctacaagg tgcgggagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtggtgggca ccgccctgat caagaagtac 3000

cccaagctgg agagcgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060

atcgccaaga gcgagcagga gatcggcaag gccaccgcca agtacttctt ctacagcaac 3120

atcatgaact tcttcaagac cgagatcacc ctggccaacg gcgagatccg gaagcggccc 3180

ctgatcgaga ccaacggcga gaccggcgag atcgtgtggg acaagggccg ggacttcgcc 3240

accgtgcgga aggtgctgag catgccccag gtgaacatcg tgaagaagac cgaggtgcag 3300

accggcggct tcagcaagga gagcatcctg cccaagcgga acagcgacaa gctgatcgcc 3360

cggaagaagg actgggaccc caagaagtac ggcggcttcg acagccccac cgtggcctac 3420

agcgtgctgg tggtggccaa ggtggagaag ggcaagagca agaagctgaa gagcgtgaag 3480

gagctgctgg gcatcaccat catggagcgg agcagcttcg agaagaaccc catcgacttc 3540

ctggaggcca agggctacaa ggaggtgaag aaggacctga tcatcaagct gcccaagtac 3600

agcctgttcg agctggagaa cggccggaag cggatgctgg ccagcgccgg cgagctgcag 3660

aagggcaacg agctggccct gcccagcaag tacgtgaact tcctgtacct ggccagccac 3720

tacgagaagc tgaagggcag ccccgaggac aacgagcaga agcagctgtt cgtggagcag 3780

cacaagcact acctggacga gatcatcgag cagatcagcg agttcagcaa gcgggtgatc 3840

ctggccgacg ccaacctgga caaggtgctg agcgcctaca acaagcaccg ggacaagccc 3900

atccgggagc aggccgagaa catcatccac ctgttcaccc tgaccaacct gggcgccccc 3960

gccgccttca agtacttcga caccaccatc gaccggaagc ggtacaccag caccaaggag 4020

gtgctggacg ccaccctgat ccaccagagc atcaccggcc tgtacgagac ccggatcgac 4080

ctgagccagc tgggcggcga cggcagcggc agccccaaga agaagcggaa ggtggacggc 4140

agccccaaga agaagcggaa ggtggacagc ggc 4173

<210> 90
<211> 4101
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 90
gacaagaagt acagcatcgg cctggccatc ggcaccaaca gcgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc cagcaagaag ttcaaggtgc tgggcaacac cgaccggcac 120

agcatcaaga agaacctgat cggcgccctg ctgttcgaca gcggcgagac cgccgaggcc 180

acccggctga agcggaccgc ccggcggcgg tacacccggc ggaagaaccg gatctgctac 240

ctgcaggaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt ccaccggctg 300

gaggagagct tcctggtgga ggaggacaag aagcacgagc ggcaccccat cttcggcaac 360

atcgtggacg aggtggccta ccacgagaag taccccacca tctaccacct gcggaagaag 420

ctggtggaca gcaccgacaa ggccgacctg cggctgatct acctggccct ggcccacatg 480

atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa cagcgacgtg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga gaaccccatc 600

aacgccagcg gcgtggacgc caaggccatc ctgagcgccc ggctgagcaa gagccggcgg 660

ctggagaacc tgatcgccca gctgcccggc gagaagaaga acggcctgtt cggcaacctg 720

atcgccctga gcctgggcct gacccccaac ttcaagagca acttcgacct ggccgaggac 780

gccaagctgc agctgagcaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacgccga cctgttcctg gccgccaaga acctgagcga cgccatcctg 900

ctgagcgaca tcctgcgggt gaacaccgag atcaccaagg ccccccctgag cgccagcatg 960

atcaagcggt acgacgagca ccaccaggac ctgaccctgc tgaaggccct ggtgcggcag 1020

cagctgcccg agaagtacaa ggagatcttc ttcgaccaga gcaagaacgg ctacgccggc 1080

tacatcgacg gcggcgccag ccaggaggag ttctacaagt tcatcaagcc catcctggag 1140

aagatggacg gcaccgagga gctgctggtg aagctgaacc gggaggacct gctgcggaag 1200

cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggcga gctgcacgcc 1260

atcctgcggc ggcaggagga cttctacccc ttcctgaagg acaaccggga gaagatcgag 1320

aagatcctga ccttccggat cccctactac gtgggccccc tggcccgggg caacagccgg 1380

ttcgcctgga tgacccggaa gagcgaggag accatcaccc cctggaactt cgaggaggtg 1440

gtggacaagg gcgccagcgc ccagagcttc atcgagcgga tgaccaactt cgacaagaac 1500

ctgcccaacg agaaggtgct gcccaagcac agcctgctgt acgagtactt caccgtgtac 1560

aacgagctga ccaaggtgaa gtacgtgacc gagggcatgc ggaagcccgc cttcctgagc 1620

ggcgagcaga agaaggccat cgtggacctg ctgttcaaga ccaaccggaa ggtgaccgtg 1680

aagcagctga aggaggacta cttcaagaag atcgagtgct tcgacagcgt ggagatcagc 1740

ggcgtggagg accggttcaa cgccagcctg ggcacctacc acgacctgct gaagatcatc 1800

aaggacaagg acttcctgga caacgaggag aacgaggaca tcctggagga catcgtgctg 1860

accctgaccc tgttcgagga ccgggagatg atcgaggagc ggctgaagac ctacgcccac 1920

ctgttcgacg acaaggtgat gaagcagctg aagcggcggc ggtacaccgg ctggggccgg 1980

ctgagccgga agctgatcaa cggcatccgg gacaagcaga gcggcaagac catcctggac 2040

ttcctgaaga gcgacggctt cgccaaccgg aacttcatgc agctgatcca cgacgacagc 2100

ctgaccttca aggaggacat ccagaaggcc caggtgagcg gccagggcga cagcctgcac 2160

gagcacatcg ccaacctggc cggcagcccc gccatcaaga agggcatcct gcagaccgtg 2220

aaggtggtgg acgagctggt gaaggtgatg ggccggcaca agcccgagaa catcgtgatc 2280

gagatggccc gggagaacca gaccacccag aagggccaga agaacagccg ggagcggatg 2340

aagcggatcg aggagggcat caaggagctg ggcagccaga tcctgaagga gcaccccgtg 2400

gagaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa cggccgggac 2460

atgtacgtgg accaggagct ggacatcaac cggctgagcg actacgacgt ggaccacatc 2520

gtgccccaga gcttcctgaa ggacgacagc atcgacaaca aggtgctgac ccggagcgac 2580

aagaaccggg gcaagagcga caacgtgccc agcgaggagg tggtgaagaa gatgaagaac 2640

tactggcggc agctgctgaa cgccaagctg atcacccagc ggaagttcga caacctgacc 2700

aaggccgagc ggggcggcct gagcgagctg gacaaggccg gcttcatcaa gcggcagctg 2760

gtggagaccc ggcagatcac caagcacgtg gcccagatcc tggacagccg gatgaacacc 2820

aagtacgacg agaacgacaa gctgatccgg gaggtgaagg tgatcaccct gaagagcaag 2880

ctggtgagcg acttccggaa ggacttccag ttctacaagg tgcgggagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtggtgggca ccgccctgat caagaagtac 3000

cccaagctgg agagcgagtt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060

atcgccaaga gcgagcagga gatcggcaag gccaccgcca agtacttctt ctacagcaac 3120

atcatgaact tcttcaagac cgagatcacc ctggccaacg gcgagatccg gaagcggccc 3180

ctgatcgaga ccaacggcga gaccggcgag atcgtgtggg acaagggccg ggacttcgcc 3240

accgtgcgga aggtgctgag catgccccag gtgaacatcg tgaagaagac cgaggtgcag 3300

accggcggct tcagcaagga gagcatcctg cccaagcgga acagcgacaa gctgatcgcc 3360

cggaagaagg actgggaccc caagaagtac ggcggcttcg acagccccac cgtggcctac 3420

agcgtgctgg tggtggccaa ggtggagaag ggcaagagca agaagctgaa gagcgtgaag 3480

gagctgctgg gcatcaccat catggagcgg agcagcttcg agaagaaccc catcgacttc 3540

ctggaggcca agggctacaa ggaggtgaag aaggacctga tcatcaagct gcccaagtac 3600

agcctgttcg agctggagaa cggccggaag cggatgctgg ccagcgccgg cgagctgcag 3660

aagggcaacg agctggccct gcccagcaag tacgtgaact tcctgtacct ggccagccac 3720

tacgagaagc tgaagggcag ccccgaggac aacgagcaga agcagctgtt cgtggagcag 3780

cacaagcact acctggacga gatcatcgag cagatcagcg agttcagcaa gcgggtgatc 3840

ctggccgacg ccaacctgga caaggtgctg agcgcctaca acaagcaccg ggacaagccc 3900

atccgggagc aggccgagaa catcatccac ctgttcaccc tgaccaacct gggcgccccc 3960

gccgccttca agtacttcga caccaccatc gaccggaagc ggtacaccag caccaaggag 4020

gtgctggacg ccaccctgat ccaccagagc atcaccggcc tgtacgagac ccggatcgac 4080

ctgagccagc tgggcggcga c 4101

<210> 91
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 91
acatttgctt ctgacacaac tgtgttcact agcaacctca aacagacacc 50

<210> 92
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 92

cataaaccct ggcgcgctcg cggcccggca ctcttctggt ccccacagac tcagagagaa 60

cccacc 66

<210> 93
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 93
aagctcagaa taaacgctca actttggcc 29

<210> 94
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 94
cagggtcctg tggacagctc accagct 27

<210> 95
<211> 61
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 95
tcccgcagtc ggcgtccagc ggctctgctt gttcgtgtgt gtgtcgttgc aggccttatt 60

c 61

<210> 96
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 96
cagatcgcct ggagacgcca tccacgctgt tttgacctcc at 42

<210> 97
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 97
agaagacacc gggaccgatc cagcctccgc ggccgggaac gg 42

<210> 98
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 98
tgcattggaa cgcggattcc ccgtgccaag agtgactcac cg 42

<210> 99
<211> 132
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 99
gctcgctttc ttgctgtcca atttctatta aaggttcctt tgttccctaa gtccaactac 60

taaactgggg gatattatga agggccttga gcatctggat tctgcctaat aaaaaacatt 120

tattttcatt gc 132

<210> 100
<211> 110
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 100
gctggagcct cggtggccat gcttcttgcc ccttgggcct ccccccagcc cctcctcccc 60

ttcctgcacc cgtacccccg tggtctttga ataaagtctg agtgggcggc 110

<210> 101
<211> 130
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 101
accagcctca agaacacccg aatggagtct ctaagctaca taataccaac ttacacttta 60

caaaatgttg tcccccaaaa tgtagccatt cgtatctgct cctaataaaa agaaagtttc 120

ttcacattct 130

<210> 102
<211> 102
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 102
ttgccagcca tctgttgttt gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac 60

tcccactgtc ctttcctaat aaaatgagga aattgcatcg ca 102

<210> 103
<211> 93
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 103
gctgccttct gcggggcttg ccttctggcc atgcccttct tctctccctt gcacctgtac 60

ctcttggtct ttgaataaag cctgagtagg aag 93

<210> 104
<211> 278
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 104
ctggtactgc atgcacgcaa tgctagctgc ccctttcccg tcctgggtac cccgagtctc 60

ccccgacctc gggtcccagg tatgctccca cctccacctg ccccactcac cacctctgct 120

agttccagac acctcccaag cacgcagcaa tgcagctcaa aacgcttagc ctagccacac 180

ccccacggga aacagcagtg attaaccttt agcaataaac gaaagtttaa ctaagctata 240

ctaaccccag ggttggtcaa tttcgtgcca gccacacc 278

<210> 105
<211> 278
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 105
caagcacgca gcaatgcagc tcaaaacgct tagcctagcc acacccccac gggaaacagc 60

agtgattaac ctttagcaat aaacgaaagt ttaactaagc tatactaacc ccagggttgg 120

tcaatttcgt gccagccaca ccctggtact gcatgcacgc aatgctagct gccccttccc 180

cgtcctgggt accccgagtc tcccccgacc tcgggtccca ggtatgctcc cacctccacc 240

tgccccactc accacctctg ctagttccag acacctcc 278

<210> 106
<211> 260
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 106
accagcctca agaacacccg aatggagtct ctaagctaca taataccaac ttacacttta 60

caaaatgttg tcccccaaaa tgtagccatt cgtatctgct cctaataaaa agaaagtttc 120

ttcacattct accagcctca agaacacccg aatggagtct ctaagctaca taataccaac 180

ttacacttta caaaatgttg tcccccaaaa tgtagccatt cgtatctgct cctaataaaa 240

agaaagtttc ttcacattct 260

<210> 107
<211> 10
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 107
gccrccaugg 10

<210> 108
<211> 13
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 108
gccgccrcca ugg 13

<210> 109
<211> 105
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 109
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gcgaaaaaaa aaaaaaaaaa aaaaaaaaaa 60

aaaccgaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa 105

<210> 110
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 110

Leu Ala Ala Lys Arg Ser Arg Thr Thr
1 5

<210> 111
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 111

Gln Ala Ala Lys Arg Ser Arg Thr Thr
1 5

<210> 112
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 112

Pro Ala Pro Ala Lys Arg Glu Arg Thr Thr
1 5 10

<210> 113
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 113

Gln Ala Ala Lys Arg Pro Arg Thr Thr
1 5

<210> 114
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 114

Arg Ala Ala Lys Arg Pro Arg Thr Thr
1 5

<210> 115
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 115

Ala Ala Ala Lys Arg Ser Trp Ser Met Ala Ala
1 5 10

<210> 116
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 116

Ala Ala Ala Lys Arg Val Trp Ser Met Ala Phe
1 5 10

<210> 117

<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 117

Ala Ala Ala Lys Arg Ser Trp Ser Met Ala Phe
1 5 10

<210> 118
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 118

Ala Ala Ala Lys Arg Lys Tyr Phe Ala Ala
1 5 10

<210> 119
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 119

Arg Ala Ala Lys Arg Lys Ala Phe Ala Ala
1 5 10

<210> 120
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 120

Arg Ala Ala Lys Arg Lys Tyr Phe Ala Val
1 5 10

<210> 121
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 121

Pro Lys Lys Lys Arg Arg Val
1 5

<210> 122
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 122

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 123
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 123
ccgaagaaga agagaaaggt c 21

<210> 124
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 124
ctggcagcaa agagaagcag aacaaca 27

<210> 125
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 125
caggcagcaa agagaagcag aacaaca 27

<210> 126
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 126
ccggcaccgg caaagagaga aagaacaaca 30

<210> 127
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 127
caggcagcaa agagaccgag aacaaca 27

<210> 128
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 128
agagcagcaa agagaccgag aacaaca 27

<210> 129
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 129
gcagcagcaa agagaagctg gagcatggca gca 33

<210> 130
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 130
gcagcagcaa agagagtctg gagcatggca ttc 33

<210> 131
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 131
gcagcagcaa agagaagctg gagcatggca ttc 33

<210> 132
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 132

gcagcagcaa agagaaagta cttcgcagca 30

<210> 133
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 133
agagcagcaa agagaaaggc attcgcagca 30

<210> 134
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 134
agagcagcaa agagaaagta cttcgcagtc 30

<210> 135
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 135
ccgaagaaga agagaagagt c 21

<210> 136

<400> 136
000

<210> 137

<400> 137
000

<210> 138

<400> 138
000

<210> 139
<211> 22
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 139
guuuuagagc uaugcuguuu ug 22

<210> 140
<211> 76
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 140
guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu 60

ggcaccgagu cggugc 76

<210> 141
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is a, c, g, or u

<400> 141
nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 142
<211> 80
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 142
guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu 60

ggcaccgagu cggugcuuuu 80

<210> 143
<211> 84
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(4)
<223> n is a, c, g, or u

<400> 143
nnnnguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau caacuugaaa 60

aaguggcacc gagucggugc uuuu 84

<210> 144
<211> 81
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(1)
<223> n is a, c, g, or u

<400> 144
nguuuuagag cuagaaauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag 60

uggcaccgag ucggugcuuu u 81

<210> 145
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is a, c, g, or u

<400> 145

nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 146

<400> 146
000

<210> 147

<400> 147
000

<210> 148

<400> 148
000

<210> 149

<400> 149
000

<210> 150

<400> 150
000

<210> 151
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 151
uaaggccagu ggaaagaauu 20

<210> 152
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 152
uuaccccacu uaacuaucuu 20

<210> 153
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 153
uuacagccac gucuacagca 20

<210> 154
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 154
uucaaaaccu gucagugauu 20

<210> 155
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 155
cgcugucaag uccaguucua 20

<210> 156
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 156
ccuuccgaaa gaggcccccc 20

<210> 157
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 157
ucccuggcug aggaucccca 20

<210> 158
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 158
acucacgaug aaauccugga 20

<210> 159
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 159
ccccccgccg uguuuguggg 20

<210> 160
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 160
gagcccccca cuguggugac 20

<210> 161
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 161
accggcucug caaaggccag 20

<210> 162
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 162
caccggcucu gcaaaggcca 20

<210> 163
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 163
ccaccggcuc ugcaaaggcc 20

<210> 164
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 164
cugcuccacc ggcucugcaa 20

<210> 165
<211> 20

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 165
cugugucacc cguuucaggu 20

<210> 166
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 166
ugugucaccc guuucaggug 20

<210> 167
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 167
acccguuuca ggugggguga 20

<210> 168
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 168
cccguuucag gugggugag 20

<210> 169
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 169
ugugcagacu cagaggugag 20

<210> 170
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 170
cagcgcaucc aggcugcagg 20

<210> 171
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 171
gcguccacau ccugcaaggg 20

<210> 172
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 172
ggcguccaca uccugcaagg 20

<210> 173
<211> 20
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 173
ugggcgucca cauccugcaa 20

<210> 174

<400> 174
000

<210> 175

<400> 175
000

<210> 176

<400> 176
000

<210> 177
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 177
uaggcagaca gacuugucac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 178
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 178
uuucaaaacc ugucagugau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 179
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 179
acucacgcug gauagccucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 180
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 180

cuuaccccac uuaacuaucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 181
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 181
uaaggccagu ggaaagaauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 182
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 182
uuaccccacu uaacuaucuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 183
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 183
uuacagccac gucuacagca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 184
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 184
uucaaaaccu gucagugauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 185
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 185
cgcugucaag uccaguucua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 186
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 186
ccuuccgaaa gaggcccccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 187
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 187
ucccuggcug aggaucccca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 188
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 188
acucacgaug aaauccugga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 189
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 189
ccccccgccg uguuuguggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 190
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 190
gagcccccca cuguggugac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 191
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 191
accggcucug caaaggccag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 192
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 192
caccggcucu gcaaaggcca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 193
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 193
ccaccggcuc ugcaaaggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 194
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 194
cugcuccacc ggcucugcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 195
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 195
cugugucacc cguuucaggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 196
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 196
ugugucaccc guuucaggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 197
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 197
acccguuuca gguggggguga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 198
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 198
cccguuucag gugggggugag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 199
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 199
ugugcagacu cagaggugag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 200
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 200
cagcgcaucc aggcugcagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 201
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 201
gcguccacau ccugcaaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 202
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 202
ggcguccaca uccugcaagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 203
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 203
ugggcgucca cauccugcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 204

<400> 204
000

<210> 205

<400> 205
000

<210> 206

<400> 206
000

<210> 207

<400> 207
000

<210> 208

<400> 208
000

<210> 209

<400> 209
000

<210> 210

<400> 210
000

<210> 211
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 211

Gly Gly Ser
1

<210> 212
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 212

Gly Gly Gly Gly Ser
1   5

<210> 213
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 213

Glu Ala Ala Ala Lys
1   5

<210> 214
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 214

Ser Glu Gly Ser Ala
1   5

<210> 215
<211> 9
<212> PRT
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 215

Ser Glu Gly Ser Ala Gly Thr Ser Thr
1 5

<210> 216
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 216

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10

<210> 217
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 217

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys
1 5 10

<210> 218
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 218

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser
1 5 10

<210> 219
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 219

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys
1   5   10

<210> 220
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 220

Ser Glu Gly Ser Ala Gly Thr Ser Thr Glu Ser Glu Gly Ser Ala
1   5   10   15

<210> 221
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 221

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1   5   10   15

<210> 222
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 222

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys
1 5 10 15

<210> 223
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 223

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser
1 5 10 15

<210> 224
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 224

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys
1 5 10 15

<210> 225
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 225

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser
1 5 10 15

<210> 226
<211> 20

<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 226

Ser Glu Gly Ser Ala Gly Thr Ser Thr Glu Ser Glu Gly Ser Ala Gly
1   5   10   15

Thr Ser Thr Glu
            20

<210> 227
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 227

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys
            20

<210> 228
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 228

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly
1   5   10   15

Gly Gly Gly Ser
            20

<210> 229
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 229

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys
            20

<210> 230
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 230

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly
1   5   10   15

Gly Gly Gly Ser
            20

<210> 231
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 231

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys
            20

<210> 232
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 232

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1   5   10   15

Gly Gly Gly Ser
            20

<210> 233
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 233

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys
            20

<210> 234
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 234

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly
1 5 10 15

Gly Gly Gly Ser
20

<210> 235
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 235

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu
1 5 10 15

Ala Ala Ala Lys
20

<210> 236
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 236

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser
20

<210> 237
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 237

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys
            20

<210> 238
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 238

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly
1   5   10   15

Gly Gly Gly Ser
            20

<210> 239
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 239

Ser Glu Gly Ser Ala Gly Thr Ser Thr Glu Ser Glu Gly Ser Ala Gly
1   5   10   15

Thr Ser Thr Glu Ser Glu Gly Ser Ala
            20   25

<210> 240
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 240

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1   5   10   15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
            20   25

<210> 241
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 241

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1   5   10   15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
            20   25

<210> 242
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 242

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
            20   25

<210> 243
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 243

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 244
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 244

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly
1   5   10   15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
            20   25

<210> 245
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 245

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly
1   5   10   15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
            20   25

<210> 246
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 246

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
            20   25

<210> 247
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 247

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 248
<211> 25
<212> PRT
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 248

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25

<210> 249
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 249

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
20 25

<210> 250
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 250

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu
1 5 10 15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
20 25

<210> 251
<211> 25

<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 251

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 252
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 252

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly
1   5   10   15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
            20   25

<210> 253
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 253

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
            20   25

<210> 254
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 254

Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 255
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 255

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1   5   10   15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
            20   25

<210> 256
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 256

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1   5   10   15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
            20   25

<210> 257
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 257

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
            20   25

<210> 258
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 258

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 259
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 259

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25

<210> 260
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 260

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Gly
1 5 10 15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
20 25

<210> 261
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 261

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu
1 5 10 15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
20 25

<210> 262
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 262

Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 263
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 263

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
            20   25

<210> 264
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 264

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 265
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 265

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Gly
1   5   10   15

Gly Gly Gly Ser Glu Ala Ala Ala Lys
            20   25

<210> 266
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 266

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys Gly Gly Gly Gly Ser
            20   25

<210> 267
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 267

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu
1   5   10   15

Ala Ala Ala Lys Glu Ala Ala Ala Lys
            20   25

<210> 268
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 268

Gly Thr Lys Asp Ser Thr Lys Asp Ile Pro Glu Thr Pro Ser Lys Asp
1   5   10   15

<210> 269
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 269

Gly Arg Asp Val Arg Gln Pro Glu Val Lys Glu Glu Lys Pro Glu Ser
1   5   10   15

<210> 270
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 270

Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Ser Thr Ala Gly
1   5   10   15

<210> 271
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 271

Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr
1   5   10   15

<210> 272
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 272

Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Thr
1   5   10   15

<210> 273

<400> 273
000

<210> 274

<400> 274
000

<210> 275

<400> 275
000

<210> 276

<400> 276
000

<210> 277

<400> 277
000

<210> 278

<400> 278
000

<210> 279

<400> 279
000

<210> 280

<400> 280
000

<210> 281

<400> 281
000

<210> 282

<400> 282
000

<210> 283

<400> 283
000

<210> 284

<400> 284
000

<210> 285

<400> 285
000

<210> 286

<400> 286
000

<210> 287

<400> 287
000

<210> 288

<400> 288
000

<210> 289

<400> 289
000

<210> 290

<400> 290
000

<210> 291

<400> 291
000

<210> 292

<400> 292
000

<210> 293

<400> 293
000

<210> 294

<400> 294
000

<210> 295

<400> 295
000

<210> 296

<400> 296
000

<210> 297

<400> 297
000

<210> 298

<400> 298
000

<210> 299

<400> 299
000

<210> 300

<400> 300
000

<210> 301
<211> 4136
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 301
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc gagacccccg 660

gcaccuccga guccgccacc cccgaguccg cagcguucaa accaaauccc aucaacuaca 720

uccugggccu ggccaucggc aucgccuccg ugggcugggc caugguggag aucgacgagg 780

aggagaaccc cauccggcug aucgaccugg gcgugcgggu guucgagcgg gccgaggugc 840

ccaagaccgg cgacucccug gccauggccc ggcggcuggc ccgguccgug cggcggcuga 900

cccggcggcg ggcccaccgg cugcugcggg cccggcggcu gcugaagcgg gagggcgugc 960

ugcaggccgc cgacuucgac gagaacggcc ugaucaaguc ccugcccaac acccccuggc 1020

agcugcgggc cgccgcccug gaccggaagc ugacccccu ggaguggucc gccgugcugc 1080

ugcaccugau caagcaccgg ggcuaccugu cccagcggaa gaacgagggc gagaccgccg 1140

acaaggagcu gggcgcccug cugaagggcg uggccaacaa cgcccacgcc cugcagaccg 1200

gcgacuuccg gacccccgcc gagcuggccc ugaacaaguu cgagaaggag uccggccaca 1260

uccggaacca gcggggcgac uacucccaca ccuucucccg gaaggaccug caggccgagc 1320

ugauccugcu guucgagaag cagaaggagu ucggcaaccc ccacgugucc ggcggccuga 1380

aggagggcau cgagacccug cugaugaccc agcggcccgc ccuguccggc gacgccgugc 1440

agaagaugcu gggccacugc accuucgagc ccgccgagcc caaggccgcc aagaacaccu 1500

acaccgccga gcgguucauc uggcugacca agcugaacaa ccugcggauc cuggagcagg 1560

gcuccgagcg gccccugacc gacaccgagc gggccacccu gauggacgag cccuaccgga 1620

aguccaagcu gaccuacgcc caggcccgga agcugcuggg ccuggaggac accgccuucu 1680

ucaagggccu gcgguacggc aaggacaacg ccgaggccuc cacccugaug gagaugaagg 1740

ccuaccacgc caucucccgg gcccuggaga aggagggccu gaaggacaag aaguccccc 1800

ugaaccuguc cuccgagcug caggacgaga ucggcaccgc cuucucccug uucaagaccg 1860

acgaggacau caccggccgg cugaaggacc gggugcagcc cgagauccug gaggcccugc 1920

ugaagcacau cuccuucgac aaguucgugc agaucucccu gaaggcccug cggcggaucg 1980

ugccccugau ggagcagggc aagcgguacg acgaggccug cgccgagauc uacggcgacc 2040

acuacggcaa gaagaacacc gaggagaaga ucuaccugcc ccccaucccc gccgacgaga 2100

uccggaaccc cguggugcug cgggcccugu cccaggcccg gaaggugauc aacggcgugg 2160

ugcggcggua cggcuccccc gcccggaucc acaucgagac cgcccgggag gugggcaagu 2220

ccuucaagga ccggaaggag aucgagaagc ggcaggagga gaaccggaag gaccgggaga 2280

aggccgccgc caaguuccgg gaguacuucc ccaacuucgu gggcgagccc aaguccaagg 2340

acauccugaa gcugcggcug uacgagcagc agcacggcaa gugccuguac uccggcaagg 2400

agaucaaccu ggugcggcug aacgagaagg gcuacgugga gaucgaccac gcccugcccu 2460

ucucccggac cugggacgac uccuucaaca acaaggugcu ggugcugggc uccgagaacc 2520

agaacaaggg caaccagacc cccuacgagu acuucaacgg caaggacaac ucccgggagu 2580

ggcaggaguu caaggcccgg guggagaccu cccgguuccc ccgguccaag aagcagcgga 2640

uccugcugca gaaguucgac gaggacggcu ucaaggagug caaccugaac gacacccggu 2700

acgugaaccg cuuccugugc caguucgugg ccgaccacau ccugcugacc ggcaagggca 2760

agcggcgggu guucgccucc aacggccaga ucaccaaccu gcugcggggc uucuggggcc 2820

ugcggaaggu gcgggccgag aacgaccggc accacgcccu ggacgccgug gugguggccu 2880

gcuccaccgu ggccaugcag cagaagauca cccgguucgu gcgguacaag gagaugaacg 2940

ccuucgacgg caagaccauc gacaaggaga ccggcaaggu gcugcaccag aagacccacu 3000

ucccccagcc cugggaguuc uucgcccagg aggugaugau ccggguguuc ggcaagcccg 3060

acggcaagcc cgaguucgag gaggccgaca cccccgagaa gcugcggacc cugcuggccg 3120

agaagcuguc cucccggccc gaggccgugc acgaguacgu gacccccccug uucgugucccc 3180

gggcccccaa ccggaagaug uccggcgccc acaaggacac ccugcggucc gccaagcggu 3240

ucgugaagca caacgagaag aucuccguga agcgggugug gcugaccgag aucaagcugg 3300

ccgaccugga gaacauggug aacuacaaga acggccggga gaucgagcug uacgaggccc 3360

ugaaggcccg gcuggaggcc uacggcggca acgccaagca ggccuucgac cccaaggaca 3420

accccuucua caagaagggc ggccagcugg ugaaggccgu gcgggguggag aagacccagg 3480

aguccggcgu gcugcugaac aagaagaacg ccuacaccau cgccgacaac ggcgacaugg 3540

ugcgggugga cguguucugc aagguggaca agaagggcaa gaaccaguac uucaucgugc 3600

ccaucuacgc cuggcaggug gccgagaaca uccugcccga caucgacugc aagggcuacc 3660

ggaucgacga cuccuacacc uucugcuucu cccugcacaa guacgaccug aucgccuucc 3720

agaaggacga gaaguccaag guggaguucg ccuacuacau caacugcgac uccuccaacg 3780

gccgguucua ccuggccugg cacgacaagg gcuccaagga gcagcaguuc cggaucucca 3840

cccagaaccu ggugcugauc cagaaguacc aggugaacga gcugggcaag gagauccggc 3900

ccugccggcu gaagaagcgg ccccccgugc gguccggaaa gcggaccgcc gacggcuccg 3960

aguucgaguc ccccaagaag aagcggaagg uggaguagug acuagcacca gccucaagaa 4020

cacccgaaug gagucucuaa gcuacauaau accaacuuac acuuuacaaa auguuguccc 4080

ccaaaaugua gccauucgua ucugcuccua auaaaaagaa aguuucuuca cauucu 4136

<210> 302
<211> 3954
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 302
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacucc 600

ggcuccgaga cccccggcac cuccgagucc gccacccccg aguccgcagc guucaaacca 660

aaucccauca acuacauccu gggccuggcc aucggcaucg ccuccguggg cugggccaug 720

guggagaucg acgaggagga gaaccccauc cggcugaucg accugggcgu gcggguguuc 780

gagcgggccg aggugcccaa gaccggcgac ucccuggcca uggcccggcg gcuggcccgg 840

uccgugcggc ggcugacccg gcggcgggcc caccggcugc ugcgggcccg gcggcugcug 900

aagcgggagg gcgugcugca ggccgccgac uucgacgaga acggccugau caagucccug 960

cccaacaccc ccuggcagcu gcgggccgcc gcccuggacc ggaagcugac ccccccuggag 1020

ugguccgccg ugcugcugca ccugaucaag caccggggcu accuguccca gcggaagaac 1080

gagggcgaga ccgccgacaa ggagcugggc gcccugcuga agggcguggc caacaacgcc 1140

cacgcccugc agaccggcga cuuccggacc cccgccgagc uggcccugaa caaguucgag 1200

aaggaguccg gccacauccg gaaccagcgg ggcgacuacu cccacaccuu cucccggaag 1260

gaccugcagg ccgagcugau ccugcuguuc gagaagcaga aggaguucgg caacccccac 1320

guguccggcg gccugaagga gggcaucgag acccugcuga ugacccagcg gcccgcccug 1380

uccggcgacg ccgugcagaa gaugcugggc cacugcaccu ucgagcccgc cgagcccaag 1440

gccgccaaga acaccuacac cgccgagcgg uucaucuggc ugaccaagcu gaacaaccug 1500

cggauccugg agcagggcuc cgagcggccc cugaccgaca ccgagcgggc cacccugaug 1560

gacgagcccu accggaaguc caagcugacc uacgcccagg cccggaagcu gcugggccug 1620

gaggacaccg ccuucuucaa gggccugcgg uacggcaagg acaacgccga ggccuccacc 1680

cugauggaga ugaaggccua ccacgccauc ucccgggccc uggagaagga gggccugaag 1740

gacaagaagu ccccccugaa ccuguccucc gagcugcagg acgagaucgg caccgccuuc 1800

ucccuguuca agaccgacga ggacaucacc ggccggcuga aggaccgggu gcagcccgag 1860

auccuggagg cccugcugaa gcacaucucc uucgacaagu ucgugcagau cucccugaag 1920

gcccugcggc ggaucgugcc ccugauggag cagggcaagc gguacgacga ggccugcgcc 1980

gagaucuacg gcgaccacua cggcaagaag aacaccgagg agaagaucua ccugcccccc 2040

aucccgccg acgagauccg gaaccccgug gugcugcggg cccuguccca ggcccggaag 2100

gugaucaacg gcguggugcg gcgguacggc ucccccgccc ggauccacau cgagaccgcc 2160

cgggaggugg gcaaguccuu caaggaccgg aaggagaucg agaagcggca ggaggagaac 2220

cggaaggacc gggagaaggc cgccgccaag uuccgggagu acuuccccaa cuucgugggc 2280

gagcccaagu ccaaggacau ccugaagcug cggcuguacg agcagcagca cggcaagugc 2340

cuguacuccg gcaaggagau caaccuggug cggcugaacg agaagggcua cguggagauc 2400

gaccacgccc ugcccuucuc ccggaccugg gacgacuccu ucaacaacaa ggugcuggug 2460

cugggcuccg agaaccagaa caagggcaac cagacccccu acgaguacuu caacggcaag 2520

gacaacuccc gggaguggca ggaguucaag gcccgggugg agaccucccg guuccccggg 2580

uccaagaagc agcggauccu gcugcagaag uucgacgagg acggcuucaa ggagugcaac 2640

cugaacgaca cccgguacgu gaaccgcuuc cugugccagu ucguggccga ccacauccug 2700

cugaccggca agggcaagcg gcggguguuc gccuccaacg gccagaucac caaccugcug 2760

cggggcuucu ggggccugcg gaaggugcgg gccgagaacg accggcacca cgcccuggac 2820

gccguggugg uggccugcuc caccguggcc augcagcaga agaucacccg guucgugcgg 2880

uacaaggaga ugaacgccuu cgacggcaag accaucgaca aggagaccgg caaggugcug 2940

caccagaaga cccacuuccc ccagcccugg gaguucuucg cccaggaggu gaugauccgg 3000

guguucggca agcccgacgg caagcccgag uucgaggagg ccgacacccc cgagaagcug 3060

cggacccugc uggccgagaa gcuguccucc cggcccgagg ccgugcacga guacgugacc 3120

ccccuguucg ugucccgggc ccccaaccgg aagauguccg gcgcccacaa ggacacccug 3180

cgguccgcca agcgguucgu gaagcacaac gagaagaucu ccgugaagcg gguguggcug 3240

accgagauca agcuggccga ccuggagaac auggugaacu acaagaacgg ccgggagauc 3300

gagcuguacg aggcccugaa ggcccggcug gaggccuacg gcggcaacgc caagcaggcc 3360

uucgacccca aggacaaccc cuucuacaag aagggcggcc agcuggugaa ggccgugcgg 3420

guggagaaga cccaggaguc cggcgugcug cugaacaaga agaacgccua caccaucgcc 3480

gacaacggcg acauggugcg gguggacgug uucugcaagg uggacaagaa gggcaagaac 3540

caguacuuca ucgugcccau cuacgccugg cagguggccg agaacauccu gcccgacauc 3600

gacugcaagg gcuaccggau cgacgacucc uacaccuucu gcuucucccu gcacaaguac 3660

gaccugaucg ccuuccagaa ggacgagaag uccaaggugg aguucgccua cuacaucaac 3720

ugcgacuccu ccaacggccg guucuaccug gccuggcacg acaagggcuc caaggagcag 3780

caguuccgga ucuccaccca gaaccuggug cugauccaga aguaccaggu gaacgagcug 3840

ggcaaggaga uccggcccug ccggcugaag aagcggcccc ccgugcgguc cggaaagcgg 3900

accgccgacg gcuccgaguu cgaguccccc aagaagaagc ggaaggugga guag 3954

<210> 303
<211> 1317
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 303

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
20 25 30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
35 40 45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
50 55 60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65 70 75 80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
85 90 95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
100 105 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
115 120 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
130 135 140

```
Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser Glu Thr Pro Gly Thr Ser
        195                 200                 205

Glu Ser Ala Thr Pro Glu Ser Ala Ala Phe Lys Pro Asn Pro Ile Asn
    210                 215                 220

Tyr Ile Leu Gly Leu Ala Ile Gly Ile Ala Ser Val Gly Trp Ala Met
225                 230                 235                 240

Val Glu Ile Asp Glu Glu Glu Asn Pro Ile Arg Leu Ile Asp Leu Gly
                245                 250                 255

Val Arg Val Phe Glu Arg Ala Glu Val Pro Lys Thr Gly Asp Ser Leu
            260                 265                 270

Ala Met Ala Arg Arg Leu Ala Arg Ser Val Arg Arg Leu Thr Arg Arg
        275                 280                 285

Arg Ala His Arg Leu Leu Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly
    290                 295                 300

Val Leu Gln Ala Ala Asp Phe Asp Glu Asn Gly Leu Ile Lys Ser Leu
305                 310                 315                 320

Pro Asn Thr Pro Trp Gln Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu
                325                 330                 335

Thr Pro Leu Glu Trp Ser Ala Val Leu Leu His Leu Ile Lys His Arg
            340                 345                 350
```

```
Gly Tyr Leu Ser Gln Arg Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu
        355                 360                 365

Leu Gly Ala Leu Leu Lys Gly Val Ala Asn Asn Ala His Ala Leu Gln
    370                 375                 380

Thr Gly Asp Phe Arg Thr Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu
385                 390                 395                 400

Lys Glu Ser Gly His Ile Arg Asn Gln Arg Gly Asp Tyr Ser His Thr
                405                 410                 415

Phe Ser Arg Lys Asp Leu Gln Ala Glu Leu Ile Leu Leu Phe Glu Lys
            420                 425                 430

Gln Lys Glu Phe Gly Asn Pro His Val Ser Gly Gly Leu Lys Glu Gly
        435                 440                 445

Ile Glu Thr Leu Leu Met Thr Gln Arg Pro Ala Leu Ser Gly Asp Ala
    450                 455                 460

Val Gln Lys Met Leu Gly His Cys Thr Phe Glu Pro Ala Glu Pro Lys
465                 470                 475                 480

Ala Ala Lys Asn Thr Tyr Thr Ala Glu Arg Phe Ile Trp Leu Thr Lys
                485                 490                 495

Leu Asn Asn Leu Arg Ile Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr
            500                 505                 510

Asp Thr Glu Arg Ala Thr Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys
        515                 520                 525

Leu Thr Tyr Ala Gln Ala Arg Lys Leu Leu Gly Leu Glu Asp Thr Ala
    530                 535                 540
```

```
Phe Phe Lys Gly Leu Arg Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr
545                 550                 555                 560

Leu Met Glu Met Lys Ala Tyr His Ala Ile Ser Arg Ala Leu Glu Lys
                565                 570                 575

Glu Gly Leu Lys Asp Lys Lys Ser Pro Leu Asn Leu Ser Ser Glu Leu
            580                 585                 590

Gln Asp Glu Ile Gly Thr Ala Phe Ser Leu Phe Lys Thr Asp Glu Asp
        595                 600                 605

Ile Thr Gly Arg Leu Lys Asp Arg Val Gln Pro Glu Ile Leu Glu Ala
    610                 615                 620

Leu Leu Lys His Ile Ser Phe Asp Lys Phe Val Gln Ile Ser Leu Lys
625                 630                 635                 640

Ala Leu Arg Arg Ile Val Pro Leu Met Glu Gln Gly Lys Arg Tyr Asp
                645                 650                 655

Glu Ala Cys Ala Glu Ile Tyr Gly Asp His Tyr Gly Lys Lys Asn Thr
            660                 665                 670

Glu Glu Lys Ile Tyr Leu Pro Pro Ile Pro Ala Asp Glu Ile Arg Asn
        675                 680                 685

Pro Val Val Leu Arg Ala Leu Ser Gln Ala Arg Lys Val Ile Asn Gly
    690                 695                 700

Val Val Arg Arg Tyr Gly Ser Pro Ala Arg Ile His Ile Glu Thr Ala
705                 710                 715                 720

Arg Glu Val Gly Lys Ser Phe Lys Asp Arg Lys Glu Ile Glu Lys Arg
                725                 730                 735

Gln Glu Glu Asn Arg Lys Asp Arg Glu Lys Ala Ala Ala Lys Phe Arg
            740                 745                 750
```

```
Glu Tyr Phe Pro Asn Phe Val Gly Glu Pro Lys Ser Lys Asp Ile Leu
        755                 760                 765

Lys Leu Arg Leu Tyr Glu Gln Gln His Gly Lys Cys Leu Tyr Ser Gly
    770                 775                 780

Lys Glu Ile Asn Leu Val Arg Leu Asn Glu Lys Gly Tyr Val Glu Ile
785                 790                 795                 800

Asp His Ala Leu Pro Phe Ser Arg Thr Trp Asp Asp Ser Phe Asn Asn
                805                 810                 815

Lys Val Leu Val Leu Gly Ser Glu Asn Gln Asn Lys Gly Asn Gln Thr
            820                 825                 830

Pro Tyr Glu Tyr Phe Asn Gly Lys Asp Asn Ser Arg Glu Trp Gln Glu
        835                 840                 845

Phe Lys Ala Arg Val Glu Thr Ser Arg Phe Pro Arg Ser Lys Lys Gln
    850                 855                 860

Arg Ile Leu Leu Gln Lys Phe Asp Glu Asp Gly Phe Lys Glu Cys Asn
865                 870                 875                 880

Leu Asn Asp Thr Arg Tyr Val Asn Arg Phe Leu Cys Gln Phe Val Ala
                885                 890                 895

Asp His Ile Leu Leu Thr Gly Lys Gly Lys Arg Arg Val Phe Ala Ser
            900                 905                 910

Asn Gly Gln Ile Thr Asn Leu Leu Arg Gly Phe Trp Gly Leu Arg Lys
        915                 920                 925

Val Arg Ala Glu Asn Asp Arg His His Ala Leu Asp Ala Val Val Val
    930                 935                 940
```

```
Ala Cys Ser Thr Val Ala Met Gln Gln Lys Ile Thr Arg Phe Val Arg
945                 950                 955                 960


Tyr Lys Glu Met Asn Ala Phe Asp Gly Lys Thr Ile Asp Lys Glu Thr
                965                 970                 975


Gly Lys Val Leu His Gln Lys Thr His Phe Pro Gln Pro Trp Glu Phe
            980                 985                 990


Phe Ala Gln Glu Val Met Ile Arg  Val Phe Gly Lys Pro  Asp Gly Lys
        995                 1000                1005


Pro Glu  Phe Glu Glu Ala Asp  Thr Pro Glu Lys Leu  Arg Thr Leu
    1010                1015                1020


Leu Ala  Glu Lys Leu Ser Ser  Arg Pro Glu Ala Val  His Glu Tyr
    1025                1030                1035


Val Thr  Pro Leu Phe Val Ser  Arg Ala Pro Asn Arg  Lys Met Ser
    1040                1045                1050


Gly Ala  His Lys Asp Thr Leu  Arg Ser Ala Lys Arg  Phe Val Lys
    1055                1060                1065


His Asn  Glu Lys Ile Ser Val  Lys Arg Val Trp Leu  Thr Glu Ile
    1070                1075                1080


Lys Leu  Ala Asp Leu Glu Asn  Met Val Asn Tyr Lys  Asn Gly Arg
    1085                1090                1095


Glu Ile  Glu Leu Tyr Glu Ala  Leu Lys Ala Arg Leu  Glu Ala Tyr
    1100                1105                1110


Gly Gly  Asn Ala Lys Gln Ala  Phe Asp Pro Lys Asp  Asn Pro Phe
    1115                1120                1125


Tyr Lys  Lys Gly Gly Gln Leu  Val Lys Ala Val Arg  Val Glu Lys
    1130                1135                1140
```

```
Thr Gln  Glu Ser Gly Val Leu  Leu Asn Lys Lys Asn  Ala Tyr Thr
    1145                1150                1155

Ile Ala  Asp Asn Gly Asp Met  Val Arg Val Asp Val  Phe Cys Lys
    1160                1165                1170

Val Asp  Lys Lys Gly Lys Asn  Gln Tyr Phe Ile Val  Pro Ile Tyr
    1175                1180                1185

Ala Trp  Gln Val Ala Glu Asn  Ile Leu Pro Asp Ile  Asp Cys Lys
    1190                1195                1200

Gly Tyr  Arg Ile Asp Asp Ser  Tyr Thr Phe Cys Phe  Ser Leu His
    1205                1210                1215

Lys Tyr  Asp Leu Ile Ala Phe  Gln Lys Asp Glu Lys  Ser Lys Val
    1220                1225                1230

Glu Phe  Ala Tyr Tyr Ile Asn  Cys Asp Ser Ser Asn  Gly Arg Phe
    1235                1240                1245

Tyr Leu  Ala Trp His Asp Lys  Gly Ser Lys Glu Gln  Gln Phe Arg
    1250                1255                1260

Ile Ser  Thr Gln Asn Leu Val  Leu Ile Gln Lys Tyr  Gln Val Asn
    1265                1270                1275

Glu Leu  Gly Lys Glu Ile Arg  Pro Cys Arg Leu Lys  Lys Arg Pro
    1280                1285                1290

Pro Val  Arg Ser Gly Lys Arg  Thr Ala Asp Gly Ser  Glu Phe Glu
    1295                1300                1305

Ser Pro  Lys Lys Lys Arg Lys  Val Glu
    1310                1315
```

<210> 304
<211> 4250
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 304
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggac ggcuccggcg 60

gcggcucccc caagaagaag cggaaggugg aggacaagcg gcccgccgcc accaagaagg 120

ccggccaggc caagaagaag aagggcggcu ccggcggcgg cgaggccucc cccgccuccg 180

gcccccggca ccugauggac ccccacaucu ucaccuccaa cuucaacaac ggcaucggcc 240

ggcacaagac cuaccugugc uacgaggugg agcggcugga caacggcacc uccgugaaga 300

uggaccagca ccggggcuuc cugcacaacc aggccaagaa ccugcugugc ggcuucuacg 360

gccggcacgc cgagcugcgg uuccuggacc uggugcccuc ccugcagcug gaccccgccc 420

agaucuaccg ggugaccugg uucaucuccu gguccccccug cuucuccugg ggcugcgccg 480

gcgaggugcg ggccuuccug caggagaaca cccacgugcg gcugcggauc uucgccgccc 540

ggaucuacga cuacgacccc cuguacaagg aggcccugca gaugcugcgg gacgccggcg 600

cccagguguc caucaugacc uacgacgagu ucaagcacug cugggacacc uucguggacc 660

accagggcug ccccuuccag cccugggacg gccuggacga gcacucccag gcccuguccg 720

gccggcugcg ggccauccug cagaaccagg gcaacuccgg cuccgagacc cccggcaccu 780

ccgaguccgc cacccccgag uccgcagcgu ucaaaccaaa ucccaucaac uacauccugg 840

gccuggccau cggcaucgcc uccgugggcu gggccauggu ggagaucgac gaggaggaga 900

accccauccg gcugaucgac cugggcgugc ggguguucga gcgggccgag gugcccaaga 960

ccggcgacuc ccuggccaug gcccggcggc uggcccgguc cgugcggcgg cugacccggc 1020

ggcgggccca ccggcugcug cgggcccggc ggcugcugaa gcgggagggc gugcugcagg 1080

ccgccgacuu cgacgagaac ggccugauca agucccugcc caacaccccc uggcagcugc 1140

gggccgccgc ccuggaccgg aagcugaccc cccuggagug guccgccgug cugcugcacc 1200

ugaucaagca ccggggcuac cugucccagc ggaagaacga gggcgagacc gccgacaagg 1260

agcugggcgc ccugcugaag ggcguggcca acaacgccca cgcccugcag accggcgacu 1320

uccggacccc cgccgagcug gcccugaaca aguucgagaa ggaguccggc cacauccgga 1380

accagcgggg cgacuacucc cacaccuucu cccggaagga ccugcaggcc gagcugaucc 1440

ugcuguucga gaagcagaag gaguucggca acccccacgu guccggcggc cugaaggagg 1500

gcaucgagac ccugcugaug acccagcggc ccgcccuguc cggcgacgcc gugcagaaga 1560

ugcugggcca cugcaccuuc gagcccgccg agcccaaggc cgccaagaac accuacaccg 1620

ccgagcgguu caucuggcug accaagcuga acaaccugcg gauccuggag cagggcuccg 1680

agcggccccu gaccgacacc gagcgggcca cccugaugga cgagcccuac cggaagucca 1740

agcugaccua cgcccaggcc cggaagcugc ugggccugga ggacaccgcc uucuucaagg 1800

gccugcggua cggcaaggac aacgccgagg ccuccacccu gauggagaug aaggccuacc 1860

acgccaucuc ccgggcccug gagaaggagg gccugaagga caagaagucc ccccugaacc 1920

uguccuccga gcugcaggac gagaucggca ccgccuucuc ccuguucaag accgacgagg 1980

acaucaccgg ccggcugaag gaccgggugc agcccgagau ccuggaggcc cugcugaagc 2040

acaucuccuu cgacaaguuc gugcagaucu cccugaaggc ccugcggcgg aucgugcccc 2100

ugauggagca gggcaagcgg uacgacgagg ccugcgccga gaucuacggc gaccacuacg 2160

gcaagaagaa caccgaggag aagaucuacc ugccccccau ccccgccgac gagauccgga 2220

accccguggu gcugcgggcc cugucccagg cccggaaggu gaucaacggc guggugcggc 2280

gguacggcuc ccccgcccgg auccacaucg agaccgcccg ggaggugggc aaguccuuca 2340

aggaccggaa ggagaucgag aagcggcagg aggagaaccg gaaggaccgg gagaaggccg 2400

ccgccaaguu ccgggaguac uuccccaacu ucgugggcga gcccaagucc aaggacaucc 2460

ugaagcugcg gcuguacgag cagcagcacg gcaagugccu guacuccggc aaggagauca 2520

accuggugcg gcugaacgag aagggcuacg uggagaucga ccacgcccug cccuucuccc 2580

ggaccuggga cgacuccuuc aacaacaagg ugcuggugcu gggcuccgag aaccagaaca 2640

agggcaacca gacccccuac gaguacuuca acggcaagga caacucccgg gaguggcagg 2700

aguucaaggc ccggguggag accucccggu ucccccgguc caagaagcag cggauccugc 2760

ugcagaaguu cgacgaggac ggcuucaagg agugcaaccu gaacgacacc cgguacguga 2820

accgcuuccu gugccaguuc guggccgacc acauccugcu gaccggcaag ggcaagcggc 2880

ggguguucgc cuccaacggc cagaucacca accugcugcg gggcuucugg ggccugcgga 2940

aggugcgggc cgagaacgac cggcaccacg cccuggacgc cgugguggug gccugcucca 3000

ccguggccau gcagcagaag aucacccggu ucgugcggua caaggagaug aacgccuucg 3060

acggcaagac caucgacaag gagaccggca aggugcugca ccagaagacc cacuuccccc 3120

agcccuggga guucuucgcc caggagguga ugauccgggu guucggcaag cccgacggca 3180

agcccgaguu cgaggaggcc gacacccccg agaagcugcg gacccugcug gccgagaagc 3240

uguccucccg gcccgaggcc gugcacgagu acgugacccc ccuguucgug ucccgggccc 3300

ccaaccggaa gauguccggc gcccacaagg acacccugcg guccgccaag cgguucguga 3360

agcacaacga gaagaucucc gugaagcggg uguggcugac cgagaucaag cuggccgacc 3420

uggagaacau ggugaacuac aagaacggcc gggagaucga gcuguacgag gcccugaagg 3480

cccggcugga ggccuacggc ggcaacgcca agcaggccuu cgaccccaag gacaaccccu 3540

ucuacaagaa gggcggccag cuggugaagg ccgugcgggu ggagaagacc caggaguccg 3600

gcgugcugcu gaacaagaag aacgccuaca ccaucgccga caacggcgac auggugcggg 3660

uggacguguu cugcaaggug gacaagaagg gcaagaacca guacuucauc gugcccaucu 3720

acgccuggca gguggccgag aacauccugc ccgacaucga cugcaagggc uaccggaucg 3780

acgacuccua caccuucugc uucucccugc acaaguacga ccugaucgcc uuccagaagg 3840

acgagaaguc caagguggag uucgccuacu acaucaacug cgacuccucc aacggccggu 3900

ucuaccuggc cuggcacgac aagggcucca aggagcagca guuccggauc uccacccaga 3960

accuggugcu gauccagaag uaccagguga acgagcuggg caaggagauc cggcccugcc 4020

ggcugaagaa gcggcccccc gugcgguccg gaaagcggac cgccgacggc uccgaguucg 4080

aguccccaa gaagaagcgg aagguggagu agugacuagc accagccuca agaacacccg 4140

aauggagucu cuaagcuaca uaauaccaac uuacacuuua caaaauguug ucccccaaaa 4200

uguagccauu cguaucugcu ccuaauaaaa agaaaguuuc uucacauucu 4250

<210> 305
<211> 4068
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 305
auggacggcu ccggcggcgg cucccccaag aagaagcgga agguggagga caagcggccc 60

gccgccacca agaaggccgg ccaggccaag aagaagaagg gcggcuccgg cggcggcgag 120

gccucccccg ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc 180

aacaacggca ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac 240

ggcaccuccg ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug 300

cugugcggcu ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug 360

cagcuggacc ccgcccagau cuaccgggug accuggguca ucuccugguc ccccugcuuc 420

uccuggggcu gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug 480

cggaucuucg ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug 540

cugcgggacg ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg 600

gacaccuucg uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac 660

ucccaggccc uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc 720

gagacccccg gcaccuccga guccgccacc cccgaguccg cagcguucaa accaaauccc 780

aucaacuaca uccugggccu ggccaucggc aucgccuccg ugggcugggc caugguggag 840

aucgacgagg aggagaaccc cauccggcug aucgaccugg gcgugcgggu guucgagcgg 900

gccgaggugc ccaagaccgg cgacucccug gccauggccc ggcggcuggc ccgguccgug 960

cggcggcuga cccggcggcg ggcccaccgg cugcugcggg cccggcggcu gcugaagcgg 1020

gagggcgugc ugcaggccgc cgacuucgac gagaacggcc ugaucaaguc ccugcccaac 1080

acccccuggc agcugcgggc cgccgcccug gaccggaagc ugacccccu ggaguggucc 1140

gccgugcugc ugcaccugau caagcaccgg ggcuaccugu cccagcggaa gaacgagggc 1200

gagaccgccg acaaggagcu gggcgcccug cugaagggcg uggccaacaa cgcccacgcc 1260

cugcagaccg gcgacuuccg gacccccgcc gagcuggccc ugaacaaguu cgagaaggag 1320

uccggccaca uccggaacca gcggggcgac uacucccaca ccuucucccg gaaggaccug 1380

caggccgagc ugauccugcu guucgagaag cagaaggagu ucggcaaccc ccacgugucc 1440

ggcggccuga aggagggcau cgagacccug cugaugaccc agcggcccgc ccuguccggc 1500

gacgccgugc agaagaugcu gggccacugc accuucgagc ccgccgagcc caaggccgcc 1560

aagaacaccu acaccgccga gcgguucauc uggcugacca agcugaacaa ccugcggauc 1620

cuggagcagg gcuccgagcg gccccugacc gacaccgagc gggccacccu gauggacgag 1680

cccuaccgga aguccaagcu gaccuacgcc caggcccgga agcugcuggg ccuggaggac 1740

accgccuucu ucaagggccu gcgguacggc aaggacaacg ccgaggccuc cacccugaug 1800

gagaugaagg ccuaccacgc caucucccgg gcccuggaga aggagggccu gaaggacaag 1860

aaguccccc ugaaccuguc cuccgagcug caggacgaga ucggcaccgc cuucucccug 1920

uucaagaccg acgaggacau caccggccgg cugaaggacc gggugcagcc cgagauccug 1980

gaggcccugc ugaagcacau cuccuucgac aaguucgugc agaucucccu gaaggcccug 2040

cggcggaucg ugccccugau ggagcagggc aagcgguacg acgaggccug cgccgagauc 2100

uacggcgacc acuacggcaa gaagaacacc gaggagaaga ucuaccugcc ccccaucccc 2160

gccgacgaga uccggaaccc cguggugcug cgggcccugu cccaggcccg gaaggugauc 2220

aacggcgugg ugcggcggua cggcucccc gcccggaucc acaucgagac cgcccgggag 2280

gugggcaagu ccuucaagga ccggaaggag aucgagaagc ggcaggagga gaaccggaag 2340

gaccgggaga aggccgccgc caaguuccgg gaguacuucc ccaacuucgu gggcgagccc 2400

aaguccaagg acauccugaa gcugcggcug uacgagcagc agcacggcaa gugccuguac 2460

uccggcaagg agaucaaccu ggugcggcug aacgagaagg gcuacgugga gaucgaccac 2520

gcccugcccu ucucccggac cugggacgac uccuucaaca acaaggugcu ggugcugggc 2580

uccgagaacc agaacaaggg caaccagacc cccuacgagu acuucaacgg caaggacaac 2640

ucccgggagu ggcaggaguu caaggcccgg guggagaccu cccgguuccc ccgguccaag 2700

aagcagcgga uccugcugca gaaguucgac gaggacggcu ucaaggagug caaccugaac 2760

gacacccggu acgugaaccg cuuccugugc caguucgugg ccgaccacau ccugcugacc 2820

ggcaagggca agcggcgggu guucgccucc aacggccaga ucaccaaccu gcugcggggc 2880

uucuggggcc ugcggaaggu gcgggccgag aacgaccggc accacgcccu ggacgccgug 2940

gugguggccu gcuccaccgu ggccaugcag cagaagauca cccgguucgu gcgguacaag 3000

gagaugaacg ccuucgacgg caagaccauc gacaaggaga ccggcaaggu gcugcaccag 3060

aagacccacu ucccccagcc cugggaguuc uucgcccagg aggugaugau ccggguguuc 3120

ggcaagcccg acggcaagcc cgaguucgag gaggccgaca cccccgagaa gcugcggacc 3180

cugcuggccg agaagcuguc cucccggccc gaggccgugc acgaguacgu gacccccug 3240

uucguguccc gggccccaa ccggaagaug uccggcgccc acaaggacac ccugcggucc 3300

gccaagcggu ucgugaagca caacgagaag aucuccguga agcgggugug gcugaccgag 3360

aucaagcugg ccgaccugga gaacauggug aacuacaaga acggccggga gaucgagcug 3420

uacgaggccc ugaaggcccg gcuggaggcc uacggcggca acgccaagca ggccuucgac 3480

cccaaggaca accccuucua caagaagggc ggccagcugg ugaaggccgu gcggguggag 3540

aagacccagg aguccggcgu gcugcugaac aagaagaacg ccuacaccau cgccgacaac 3600

ggcgacaugg ugcgggugga cguguucugc aagguggaca agaagggcaa gaaccaguac 3660

uucaucgugc ccaucuacgc cuggcaggug gccgagaaca uccugcccga caucgacugc 3720

aagggcuacc ggaucgacga cuccuacacc uucugcuucu cccugcacaa guacgaccug 3780

aucgccuucc agaaggacga gaaguccaag guggaguucg ccuacuacau caacugcgac 3840

uccuccaacg gccgguucua ccuggccugg cacgacaagg gcuccaagga gcagcaguuc 3900

cggaucucca cccagaaccu ggugcugauc cagaaguacc aggugaacga gcugggcaag 3960

gagauccggc ccugccggcu gaagaagcgg ccccccgugc gguccggaaa gcggaccgcc 4020

gacggcuccg aguucgaguc ccccaagaag aagcggaagg uggaguag 4068

<210> 306
<211> 1355

<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 306

Met Asp Gly Ser Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val Glu
1 5 10 15

Asp Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
20 25 30

Lys Gly Gly Ser Gly Gly Gly Glu Ala Ser Pro Ala Ser Gly Pro Arg
35 40 45

His Leu Met Asp Pro His Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile
50 55 60

Gly Arg His Lys Thr Tyr Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn
65 70 75 80

Gly Thr Ser Val Lys Met Asp Gln His Arg Gly Phe Leu His Asn Gln
85 90 95

Ala Lys Asn Leu Leu Cys Gly Phe Tyr Gly Arg His Ala Glu Leu Arg
100 105 110

Phe Leu Asp Leu Val Pro Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr
115 120 125

Arg Val Thr Trp Phe Ile Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys
130 135 140

Ala Gly Glu Val Arg Ala Phe Leu Gln Glu Asn Thr His Val Arg Leu
145 150 155 160

Arg Ile Phe Ala Ala Arg Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu
165 170 175

```
Ala Leu Gln Met Leu Arg Asp Ala Gly Ala Gln Val Ser Ile Met Thr
            180                 185                 190

Tyr Asp Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp His Gln Gly
        195                 200                 205

Cys Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser Gln Ala Leu
    210                 215                 220

Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser
225                 230                 235                 240

Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Ala Ala Phe
                245                 250                 255

Lys Pro Asn Pro Ile Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Ala
            260                 265                 270

Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Glu Glu Asn Pro Ile
        275                 280                 285

Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg Ala Glu Val Pro
    290                 295                 300

Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu Ala Arg Ser Val
305                 310                 315                 320

Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu Arg Ala Arg Arg
                325                 330                 335

Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp Phe Asp Glu Asn
            340                 345                 350

Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln Leu Arg Ala Ala
        355                 360                 365
```

```
Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser Ala Val Leu Leu
    370                 375                 380

His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg Lys Asn Glu Gly
385                 390                 395                 400

Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys Gly Val Ala Asn
                405                 410                 415

Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr Pro Ala Glu Leu
            420                 425                 430

Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile Arg Asn Gln Arg
        435                 440                 445

Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu Gln Ala Glu Leu
    450                 455                 460

Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn Pro His Val Ser
465                 470                 475                 480

Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met Thr Gln Arg Pro
                485                 490                 495

Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly His Cys Thr Phe
            500                 505                 510

Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr Thr Ala Glu Arg
        515                 520                 525

Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile Leu Glu Gln Gly
    530                 535                 540

Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr Leu Met Asp Glu
545                 550                 555                 560

Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala Arg Lys Leu Leu
                565                 570                 575
```

```
Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg Tyr Gly Lys Asp
            580                 585                 590

Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala Tyr His Ala Ile
        595                 600                 605

Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys Lys Ser Pro Leu
    610                 615                 620

Asn Leu Ser Ser Glu Leu Gln Asp Glu Ile Gly Thr Ala Phe Ser Leu
625                 630                 635                 640

Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys Asp Arg Val Gln
                645                 650                 655

Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser Phe Asp Lys Phe
            660                 665                 670

Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val Pro Leu Met Glu
        675                 680                 685

Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile Tyr Gly Asp His
    690                 695                 700

Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu Pro Pro Ile Pro
705                 710                 715                 720

Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala Leu Ser Gln Ala
                725                 730                 735

Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly Ser Pro Ala Arg
            740                 745                 750

Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser Phe Lys Asp Arg
        755                 760                 765
```

```
Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys Asp Arg Glu Lys
    770                 775                 780

Ala Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe Val Gly Glu Pro
785                 790                 795                 800

Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu Gln Gln His Gly
                805                 810                 815

Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Val Arg Leu Asn Glu
            820                 825                 830

Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe Ser Arg Thr Trp
        835                 840                 845

Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly Ser Glu Asn Gln
    850                 855                 860

Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn Gly Lys Asp Asn
865                 870                 875                 880

Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu Thr Ser Arg Phe
                885                 890                 895

Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys Phe Asp Glu Asp
            900                 905                 910

Gly Phe Lys Glu Cys Asn Leu Asn Asp Thr Arg Tyr Val Asn Arg Phe
        915                 920                 925

Leu Cys Gln Phe Val Ala Asp His Ile Leu Leu Thr Gly Lys Gly Lys
    930                 935                 940

Arg Arg Val Phe Ala Ser Asn Gly Gln Ile Thr Asn Leu Leu Arg Gly
945                 950                 955                 960

Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp Arg His His Ala
                965                 970                 975
```

```
Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala Met Gln Gln Lys
            980                 985                 990


Ile Thr Arg Phe Val Arg Tyr Lys  Glu Met Asn Ala Phe  Asp Gly Lys
        995                 1000                1005


Thr Ile  Asp Lys Glu Thr Gly  Lys Val Leu His Gln  Lys Thr His
    1010                1015                1020


Phe Pro  Gln Pro Trp Glu Phe  Phe Ala Gln Glu Val  Met Ile Arg
    1025                1030                1035


Val Phe  Gly Lys Pro Asp Gly  Lys Pro Glu Phe Glu  Glu Ala Asp
    1040                1045                1050


Thr Pro  Glu Lys Leu Arg Thr  Leu Leu Ala Glu Lys  Leu Ser Ser
    1055                1060                1065


Arg Pro  Glu Ala Val His Glu  Tyr Val Thr Pro Leu  Phe Val Ser
    1070                1075                1080


Arg Ala  Pro Asn Arg Lys Met  Ser Gly Ala His Lys  Asp Thr Leu
    1085                1090                1095


Arg Ser  Ala Lys Arg Phe Val  Lys His Asn Glu Lys  Ile Ser Val
    1100                1105                1110


Lys Arg  Val Trp Leu Thr Glu  Ile Lys Leu Ala Asp  Leu Glu Asn
    1115                1120                1125


Met Val  Asn Tyr Lys Asn Gly  Arg Glu Ile Glu Leu  Tyr Glu Ala
    1130                1135                1140


Leu Lys  Ala Arg Leu Glu Ala  Tyr Gly Gly Asn Ala  Lys Gln Ala
    1145                1150                1155
```

```
Phe Asp  Pro Lys Asp Asn Pro  Phe Tyr Lys Lys Gly  Gly Gln Leu
    1160                1165                1170

Val Lys  Ala Val Arg Val Glu  Lys Thr Gln Glu Ser  Gly Val Leu
    1175                1180                1185

Leu Asn  Lys Lys Asn Ala Tyr  Thr Ile Ala Asp Asn  Gly Asp Met
    1190                1195                1200

Val Arg  Val Asp Val Phe Cys  Lys Val Asp Lys Lys  Gly Lys Asn
    1205                1210                1215

Gln Tyr  Phe Ile Val Pro Ile  Tyr Ala Trp Gln Val  Ala Glu Asn
    1220                1225                1230

Ile Leu  Pro Asp Ile Asp Cys  Lys Gly Tyr Arg Ile  Asp Asp Ser
    1235                1240                1245

Tyr Thr  Phe Cys Phe Ser Leu  His Lys Tyr Asp Leu  Ile Ala Phe
    1250                1255                1260

Gln Lys  Asp Glu Lys Ser Lys  Val Glu Phe Ala Tyr  Tyr Ile Asn
    1265                1270                1275

Cys Asp  Ser Ser Asn Gly Arg  Phe Tyr Leu Ala Trp  His Asp Lys
    1280                1285                1290

Gly Ser  Lys Glu Gln Gln Phe  Arg Ile Ser Thr Gln  Asn Leu Val
    1295                1300                1305

Leu Ile  Gln Lys Tyr Gln Val  Asn Glu Leu Gly Lys  Glu Ile Arg
    1310                1315                1320

Pro Cys  Arg Leu Lys Lys Arg  Pro Pro Val Arg Ser  Gly Lys Arg
    1325                1330                1335

Thr Ala  Asp Gly Ser Glu Phe  Glu Ser Pro Lys Lys  Lys Arg Lys
    1340                1345                1350
```

Val Glu
1355

<210> 307
<211> 4368
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 307
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggac ggcuccggcg 60

gcggcucccc caagaagaag cggaaggugg aggacaagcg gcccgccgcc accaagaagg 120

ccggccaggc caagaagaag aagggcggcu ccggcggcgg cgaggccucc cccgccuccg 180

gcccccggca ccugauggac ccccacaucu ucaccuccaa cuucaacaac ggcaucggcc 240

ggcacaagac cuaccugugc uacgaggugg agcggcugga caacggcacc uccgugaaga 300

uggaccagca ccggggcuuc cugcacaacc aggccaagaa ccugcugugc ggcuucuacg 360

gccggcacgc cgagcugcgg uuccuggacc uggugcccuc ccugcagcug gaccccgccc 420

agaucuaccg ggugaccugg uucaucuccu gguccccug cuucuccugg ggcugcgccg 480

gcgaggugcg ggccuuccug caggagaaca cccacgugcg gcugcggauc uucgccgccc 540

ggaucuacga cuacgacccc cuguacaagg aggcccugca gaugcugcgg gacgccggcg 600

cccagguguc caucaugacc uacgacgagu ucaagcacug cugggacacc uucguggacc 660

accagggcug ccccuuccag cccugggacg gccuggacga gcacucccag gcccuguccg 720

gccggcugcg ggccauccug cagaaccagg gcaacuccgg cuccgagacc cccggcaccu 780

ccgaguccgc cacccccgag uccgcagcgu ucaaaccaaa ucccaucaac uacauccugg 840

gccuggccau cggcaucgcc uccgugggcu gggccauggu ggagaucgac gaggaggaga 900

accccauccg gcugaucgac cugggcgugc ggguguucga gcgggccgag gugcccaaga 960

ccggcgacuc ccuggccaug gcccggcggc uggcccgguc cgugcggcgg cugacccggc 1020

ggcgggccca ccggcugcug cgggcccggc ggcugcugaa gcgggagggc gugcugcagg 1080

ccgccgacuu cgacgagaac ggccugauca agucccugcc caacaccccc uggcagcugc 1140

gggccgccgc ccuggaccgg aagcugaccc cccuggagug guccgccgug cugcugcacc 1200

ugaucaagca ccggggcuac cugucccagc ggaagaacga gggcgagacc gccgacaagg 1260

agcugggcgc ccugcugaag ggcguggcca acaacgccca cgcccugcag accggcgacu 1320

uccggacccc cgccgagcug gcccugaaca aguucgagaa ggaguccggc cacauccgga 1380

accagcgggg cgacuacucc cacaccuucu cccggaagga ccugcaggcc gagcugaucc 1440

ugcuguucga gaagcagaag gaguucggca acccccacgu guccggcggc cugaaggagg 1500

gcaucgagac ccugcugaug acccagcggc ccgcccuguc cggcgacgcc gugcagaaga 1560

ugcugggcca cugcaccuuc gagcccgccg agcccaaggc cgccaagaac accuacaccg 1620

ccgagcgguu caucuggcug accaagcuga acaaccugcg gauccuggag cagggcuccg 1680

agcggccccu gaccgacacc gagcgggcca cccugaugga cgagcccuac cggaagucca 1740

agcugaccua cgcccaggcc cggaagcugc ugggccugga ggacaccgcc uucuucaagg 1800

gccugcggua cggcaaggac aacgccgagg ccuccacccu gauggagaug aaggccuacc 1860

acgccaucuc ccgggcccug gagaaggagg gccugaagga caagaagucc ccccugaacc 1920

uguccuccga gcugcaggac gagaucggca ccgccuucuc ccuguucaag accgacgagg 1980

acaucaccgg ccggcugaag gaccgggugc agcccgagau ccuggaggcc cugcugaagc 2040

acaucuccuu cgacaaguuc gugcagaucu cccugaaggc ccugcggcgg aucgugcccc 2100

ugauggagca gggcaagcgg uacgacgagg ccugcgccga gaucuacggc gaccacuacg 2160

gcaagaagaa caccgaggag aagaucuacc ugccccccau ccccgccgac gagauccgga 2220

accccguggu gcugcgggcc cugucccagg cccggaaggu gaucaacggc guggugcggc 2280

gguacggcuc ccccgcccgg auccacaucg agaccgcccg ggaggugggc aaguccuuca 2340

aggaccggaa ggagaucgag aagcggcagg aggagaaccg gaaggaccgg gagaaggccg 2400

ccgccaaguu ccgggaguac uuccccaacu ucgugggcga gcccaagucc aaggacaucc 2460

ugaagcugcg gcuguacgag cagcagcacg gcaagugccu guacuccggc aaggagauca 2520

accuggugcg gcugaacgag aagggcuacg uggagaucga ccacgcccug cccuucuccc 2580

ggaccuggga cgacuccuuc aacaacaagg ugcuggugcu gggcuccgag aaccagaaca 2640

agggcaacca gacccccuac gaguacuuca acggcaagga caacucccgg gaguggcagg 2700

aguucaaggc ccggguggag accucccggu ucccccgguc caagaagcag cggauccugc 2760

ugcagaaguu cgacgaggac ggcuucaagg agugcaaccu gaacgacacc cgguacguga 2820

accgcuuccu gugccaguuc guggccgacc acauccugcu gaccggcaag ggcaagcggc 2880

ggguguucgc cuccaacggc cagaucacca accugcugcg gggcuucugg ggccugcgga 2940

aggugcgggc cgagaacgac cggcaccacg cccuggacgc cgugguggug gccugcucca 3000

ccguggccau gcagcagaag aucacccggu ucgugcggua caaggagaug aacgccuucg 3060

acggcaagac caucgacaag gagaccggca aggugcugca ccagaagacc cacuuccccc 3120

agcccuggga guucuucgcc caggagguga ugauccgggu guucggcaag cccgacggca 3180

agcccgaguu cgaggaggcc gacacccccg agaagcugcg gacccugcug gccgagaagc 3240

uguccucccg gcccgaggcc gugcacgagu acgugacccc ccuguucgug ucccgggccc 3300

ccaaccggaa gauguccggc gcccacaagg acacccugcg guccgccaag cgguucguga 3360

agcacaacga gaagaucucc gugaagcggg uguggcugac cgagaucaag cuggccgacc 3420

uggagaacau ggugaacuac aagaacggcc gggagaucga gcuguacgag gcccugaagg 3480

cccggcugga ggccuacggc ggcaacgcca agcaggccuu cgaccccaag gacaaccccu 3540

ucuacaagaa gggcggccag cuggugaagg ccgugcgggu ggagaagacc caggaguccg 3600

gcgugcugcu gaacaagaag aacgccuaca ccaucgccga caacggcgac auggugcggg 3660

uggacguguu cugcaaggug gacaagaagg gcaagaacca guacuucauc gugcccaucu 3720

acgccuggca gguggccgag aacauccugc ccgacaucga cugcaagggc uaccggaucg 3780

acgacuccua caccuucugc uucucccugc acaaguacga ccugaucgcc uuccagaagg 3840

acgagaaguc caagguggag uucgccuacu acaucaacug cgacuccucc aacggccggu 3900

ucuaccuggc cuggcacgac aagggcucca aggagcagca guuccggauc uccacccaga 3960

accuggugcu gauccagaag uaccagguga acgagcuggg caaggagauc cggcccugcc 4020

ggcugaagaa gcggcccccc gugcgggagg acaagcggcc cgccgccacc aagaaggccg 4080

gccaggccaa gaagaagaag uaccccuacg acgugcccga cuacgccggc uaccccuacg 4140

acgugcccga cuacgccggc uccuaccccu acgacgugcc cgacuacgcc gccgcccccg 4200

ccgccaagaa gaagaagcug gacuagcuag ugacuagcac cagccucaag aacacccgaa 4260

uggagucucu aagcuacaua auaccaacuu acacuuuaca aaauguuguc ccccaaaaug 4320

uagccauucg uaucugcucc uaauaaaaag aaaguuucuu cacauucu 4368

<210> 308
<211> 4180
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 308
auggacggcu ccggcggcgg cucccccaag aagaagcgga agguggagga caagcggccc 60

gccgccacca agaaggccgg ccaggccaag aagaagaagg gcggcuccgg cggcggcgag 120

gccucccccg ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc 180

aacaacggca ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac 240

ggcaccuccg ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug 300

cugugcggcu ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug 360

cagcuggacc ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc 420

uccuggggcu gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug 480

cggaucuucg ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug 540

cugcgggacg ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg 600

gacaccuucg uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac 660

ucccaggccc uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc 720

gagacccccg gcaccuccga guccgccacc cccgaguccg cagcguucaa accaaauccc 780

aucaacuaca uccugggccu ggccaucggc aucgccuccg ugggcugggc caugguggag 840

aucgacgagg aggagaaccc cauccggcug aucgaccugg gcgugcgggu guucgagcgg 900

gccgaggugc ccaagaccgg cgacucccug gccauggccc ggcggcuggc ccgguccgug 960

cggcggcuga cccggcggcg ggcccaccgg cugcugcggg cccggcggcu gcugaagcgg 1020

gagggcgugc ugcaggccgc cgacuucgac gagaacggcc ugaucaaguc ccugcccaac 1080

acccccuggc agcugcgggc cgccgcccug gaccggaagc ugacccccu ggaguggucc 1140

gccgugcugc ugcaccugau caagcaccgg ggcuaccugu cccagcggaa gaacgagggc 1200

gagaccgccg acaaggagcu gggcgcccug cugaagggcg uggccaacaa cgcccacgcc 1260

cugcagaccg gcgacuuccg gacccccgcc gagcuggccc ugaacaaguu cgagaaggag 1320

uccggccaca uccggaacca gcggggcgac uacucccaca ccuucucccg gaaggaccug 1380

caggccgagc ugauccugcu guucgagaag cagaaggagu ucggcaaccc ccacgugucc 1440

ggcggccuga aggagggcau cgagacccug cugaugaccc agcggcccgc ccuguccggc 1500

gacgccgugc agaagaugcu gggccacugc accuucgagc ccgccgagcc caaggccgcc 1560

aagaacaccu acaccgccga gcgguucauc uggcugacca agcugaacaa ccugcggauc 1620

cuggagcagg gcuccgagcg gccccugacc gacaccgagc gggccacccu gauggacgag 1680

cccuaccgga aguccaagcu gaccuacgcc caggcccgga agcugcuggg ccuggaggac 1740

accgccuucu ucaagggccu gcgguacggc aaggacaacg ccgaggccuc cacccugaug 1800

gagaugaagg ccuaccacgc caucucccgg gcccuggaga aggagggccu gaaggacaag 1860

aagucccccc ugaaccuguc cuccgagcug caggacgaga ucggcaccgc cuucucccug 1920

uucaagaccg acgaggacau caccggccgg cugaaggacc gggugcagcc cgagauccug 1980

gaggcccugc ugaagcacau cuccuucgac aaguucgugc agaucucccu gaaggcccug 2040

cggcggaucg ugccccugau ggagcagggc aagcgguacg acgaggccug cgccgagauc 2100

uacggcgacc acuacggcaa gaagaacacc gaggagaaga ucuaccugcc ccccaucccc 2160

gccgacgaga uccggaaccc cguggugcug cgggcccugu cccaggcccg gaaggugauc 2220

aacggcgugg ugcggcggua cggcuccccc gcccggaucc acaucgagac cgcccgggag 2280

gugggcaagu ccuucaagga ccggaaggag aucgagaagc ggcaggagga gaaccggaag 2340

gaccgggaga aggccgccgc caaguuccgg gaguacuucc ccaacuucgu gggcgagccc 2400

aaguccaagg acauccugaa gcugcggcug uacgagcagc agcacggcaa gugccuguac 2460

uccggcaagg agaucaaccu ggugcggcug aacgagaagg gcuacgugga gaucgaccac 2520

gcccugcccu ucucccggac cugggacgac uccuucaaca acaaggugcu ggugcugggc 2580

uccgagaacc agaacaaggg caaccagacc cccuacgagu acuucaacgg caaggacaac 2640

ucccgggagu ggcaggaguu caaggcccgg guggagaccu cccgguuccc ccgguccaag 2700

aagcagcgga uccugcugca gaaguucgac gaggacggcu ucaaggagug caaccugaac 2760

gacacccggu acgugaaccg cuuccugugc caguucgugg ccgaccacau ccugcugacc 2820

ggcaagggca agcggcgggu guucgccucc aacggccaga ucaccaaccu gcugcggggc 2880

uucuggggcc ugcggaaggu gcgggccgag aacgaccggc accacgcccu ggacgccgug 2940

gugguggccu gcuccaccgu ggccaugcag cagaagauca cccgguucgu gcgguacaag 3000

gagaugaacg ccuucgacgg caagaccauc gacaaggaga ccggcaaggu gcugcaccag 3060

aagacccacu ucccccagcc cugggaguuc uucgcccagg aggugaugau ccggguguuc 3120

ggcaagcccg acggcaagcc cgaguucgag gaggccgaca cccccgagaa gcugcggacc 3180

cugcuggccg agaagcuguc cucccggccc gaggccgugc acgaguacgu gacccccug 3240

uucguguccc gggccccaa ccggaagaug uccggcgccc acaaggacac ccugcggucc 3300

gccaagcggu ucgugaagca caacgagaag aucuccguga agcgggugug gcugaccgag 3360

aucaagcugg ccgaccugga gaacauggug aacuacaaga acggccggga gaucgagcug 3420

uacgaggccc ugaaggcccg gcuggaggcc uacggcggca acgccaagca ggccuucgac 3480

cccaaggaca accccuucua caagaagggc ggccagcugg ugaaggccgu gcggguggag 3540

aagacccagg aguccggcgu gcugcugaac aagaagaacg ccuacaccau cgccgacaac 3600

ggcgacaugg ugcgggugga cguguucugc aagguggaca agaagggcaa gaaccaguac 3660

uucaucgugc ccaucuacgc cuggcaggug gccgagaaca uccugcccga caucgacugc 3720

aagggcuacc ggaucgacga cuccuacacc uucugcuucu cccugcacaa guacgaccug 3780

aucgccuucc agaaggacga gaaguccaag guggaguucg ccuacuacau caacugcgac 3840

uccuccaacg gccgguucua ccuggccugg cacgacaagg gcuccaagga gcagcaguuc 3900

cggaucucca cccagaaccu ggugcugauc cagaaguacc aggugaacga gcugggcaag 3960

gagauccggc ccugccggcu gaagaagcgg ccccccgugc gggaggacaa gcggcccgcc 4020

gccaccaaga aggccggcca ggccaagaag aagaaguacc ccuacgacgu gcccgacuac 4080

gccggcuacc ccuacgacgu gcccgacuac gccggcuccu accccuacga cgugcccgac 4140

uacgccgccg cccccgccgc caagaagaag aagcuggacu 4180

<210> 309
<211> 1393
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 309

Met Asp Gly Ser Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val Glu
1 5 10 15

Asp Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
20 25 30

Lys Gly Gly Ser Gly Gly Gly Glu Ala Ser Pro Ala Ser Gly Pro Arg
35 40 45

His Leu Met Asp Pro His Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile
50 55 60

Gly Arg His Lys Thr Tyr Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn
65 70 75 80

Gly Thr Ser Val Lys Met Asp Gln His Arg Gly Phe Leu His Asn Gln
85 90 95

Ala Lys Asn Leu Leu Cys Gly Phe Tyr Gly Arg His Ala Glu Leu Arg
100 105 110

```
Phe Leu Asp Leu Val Pro Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr
        115                 120                 125

Arg Val Thr Trp Phe Ile Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys
    130                 135                 140

Ala Gly Glu Val Arg Ala Phe Leu Gln Glu Asn Thr His Val Arg Leu
145                 150                 155                 160

Arg Ile Phe Ala Ala Arg Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu
                165                 170                 175

Ala Leu Gln Met Leu Arg Asp Ala Gly Ala Gln Val Ser Ile Met Thr
            180                 185                 190

Tyr Asp Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp His Gln Gly
        195                 200                 205

Cys Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser Gln Ala Leu
    210                 215                 220

Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser
225                 230                 235                 240

Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Ala Ala Phe
                245                 250                 255

Lys Pro Asn Pro Ile Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Ala
            260                 265                 270

Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Glu Glu Asn Pro Ile
        275                 280                 285

Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg Ala Glu Val Pro
    290                 295                 300

Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu Ala Arg Ser Val
305                 310                 315                 320
```

```
Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu Arg Ala Arg Arg
                325                 330                 335

Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp Phe Asp Glu Asn
            340                 345                 350

Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln Leu Arg Ala Ala
        355                 360                 365

Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser Ala Val Leu Leu
    370                 375                 380

His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg Lys Asn Glu Gly
385                 390                 395                 400

Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys Gly Val Ala Asn
                405                 410                 415

Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr Pro Ala Glu Leu
            420                 425                 430

Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile Arg Asn Gln Arg
        435                 440                 445

Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu Gln Ala Glu Leu
    450                 455                 460

Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn Pro His Val Ser
465                 470                 475                 480

Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met Thr Gln Arg Pro
                485                 490                 495

Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly His Cys Thr Phe
            500                 505                 510
```

```
Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr Thr Ala Glu Arg
        515                 520                 525

Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile Leu Glu Gln Gly
    530                 535                 540

Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr Leu Met Asp Glu
545                 550                 555                 560

Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala Arg Lys Leu Leu
                565                 570                 575

Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg Tyr Gly Lys Asp
            580                 585                 590

Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala Tyr His Ala Ile
        595                 600                 605

Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys Lys Ser Pro Leu
    610                 615                 620

Asn Leu Ser Ser Glu Leu Gln Asp Glu Ile Gly Thr Ala Phe Ser Leu
625                 630                 635                 640

Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys Asp Arg Val Gln
                645                 650                 655

Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser Phe Asp Lys Phe
            660                 665                 670

Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val Pro Leu Met Glu
        675                 680                 685

Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile Tyr Gly Asp His
    690                 695                 700

Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu Pro Pro Ile Pro
705                 710                 715                 720
```

```
Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala Leu Ser Gln Ala
                725                 730                 735

Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly Ser Pro Ala Arg
            740                 745                 750

Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser Phe Lys Asp Arg
        755                 760                 765

Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys Asp Arg Glu Lys
    770                 775                 780

Ala Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe Val Gly Glu Pro
785                 790                 795                 800

Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu Gln Gln His Gly
                805                 810                 815

Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Val Arg Leu Asn Glu
            820                 825                 830

Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe Ser Arg Thr Trp
        835                 840                 845

Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly Ser Glu Asn Gln
    850                 855                 860

Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn Gly Lys Asp Asn
865                 870                 875                 880

Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu Thr Ser Arg Phe
                885                 890                 895

Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys Phe Asp Glu Asp
            900                 905                 910
```

```
Gly Phe Lys Glu Cys Asn Leu Asn Asp Thr Arg Tyr Val Asn Arg Phe
        915                 920                 925

Leu Cys Gln Phe Val Ala Asp His Ile Leu Leu Thr Gly Lys Gly Lys
    930                 935                 940

Arg Arg Val Phe Ala Ser Asn Gly Gln Ile Thr Asn Leu Leu Arg Gly
945                 950                 955                 960

Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp Arg His His Ala
                965                 970                 975

Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala Met Gln Gln Lys
            980                 985                 990

Ile Thr Arg Phe Val Arg Tyr Lys  Glu Met Asn Ala Phe  Asp Gly Lys
        995                 1000                1005

Thr Ile  Asp Lys Glu Thr Gly  Lys Val Leu His Gln  Lys Thr His
    1010                1015                1020

Phe Pro  Gln Pro Trp Glu Phe  Phe Ala Gln Glu Val  Met Ile Arg
    1025                1030                1035

Val Phe  Gly Lys Pro Asp Gly  Lys Pro Glu Phe Glu  Glu Ala Asp
    1040                1045                1050

Thr Pro  Glu Lys Leu Arg Thr  Leu Leu Ala Glu Lys  Leu Ser Ser
    1055                1060                1065

Arg Pro  Glu Ala Val His Glu  Tyr Val Thr Pro Leu  Phe Val Ser
    1070                1075                1080

Arg Ala  Pro Asn Arg Lys Met  Ser Gly Ala His Lys  Asp Thr Leu
    1085                1090                1095

Arg Ser  Ala Lys Arg Phe Val  Lys His Asn Glu Lys  Ile Ser Val
    1100                1105                1110
```

```
Lys Arg  Val Trp Leu Thr Glu  Ile Lys Leu Ala Asp  Leu Glu Asn
    1115                1120                1125

Met Val  Asn Tyr Lys Asn Gly  Arg Glu Ile Glu Leu  Tyr Glu Ala
    1130                1135                1140

Leu Lys  Ala Arg Leu Glu Ala  Tyr Gly Gly Asn Ala  Lys Gln Ala
    1145                1150                1155

Phe Asp  Pro Lys Asp Asn Pro  Phe Tyr Lys Lys Gly  Gly Gln Leu
    1160                1165                1170

Val Lys  Ala Val Arg Val Glu  Lys Thr Gln Glu Ser  Gly Val Leu
    1175                1180                1185

Leu Asn  Lys Lys Asn Ala Tyr  Thr Ile Ala Asp Asn  Gly Asp Met
    1190                1195                1200

Val Arg  Val Asp Val Phe Cys  Lys Val Asp Lys Lys  Gly Lys Asn
    1205                1210                1215

Gln Tyr  Phe Ile Val Pro Ile  Tyr Ala Trp Gln Val  Ala Glu Asn
    1220                1225                1230

Ile Leu  Pro Asp Ile Asp Cys  Lys Gly Tyr Arg Ile  Asp Asp Ser
    1235                1240                1245

Tyr Thr  Phe Cys Phe Ser Leu  His Lys Tyr Asp Leu  Ile Ala Phe
    1250                1255                1260

Gln Lys  Asp Glu Lys Ser Lys  Val Glu Phe Ala Tyr  Tyr Ile Asn
    1265                1270                1275

Cys Asp  Ser Ser Asn Gly Arg  Phe Tyr Leu Ala Trp  His Asp Lys
    1280                1285                1290
```

```
Gly Ser  Lys Glu Gln Gln Phe  Arg Ile Ser Thr Gln  Asn Leu Val
    1295                1300                1305

Leu Ile  Gln Lys Tyr Gln Val  Asn Glu Leu Gly Lys  Glu Ile Arg
    1310                1315                1320

Pro Cys  Arg Leu Lys Lys Arg  Pro Pro Val Arg Glu  Asp Lys Arg
    1325                1330                1335

Pro Ala  Ala Thr Lys Lys Ala  Gly Gln Ala Lys Lys  Lys Lys Tyr
    1340                1345                1350

Pro Tyr  Asp Val Pro Asp Tyr  Ala Gly Tyr Pro Tyr  Asp Val Pro
    1355                1360                1365

Asp Tyr  Ala Gly Ser Tyr Pro  Tyr Asp Val Pro Asp  Tyr Ala Ala
    1370                1375                1380

Ala Pro  Ala Ala Lys Lys Lys  Lys Leu Asp
    1385                1390
```

<210> 310
<211> 4187
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 310

```
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggac ggcuccggcg       60

gcggcucccc caagaagaag cggaaggugg aggacaagcg gcccgccgcc accaagaagg      120

ccggccaggc caagaagaag aagggcggcu ccggcggcgg cgaggccucc cccgccuccg      180

gcccccggca ccugauggac ccccacaucu ucaccuccaa cuucaacaac ggcaucggcc      240

ggcacaagac cuaccugugc uacgaggugg agcggcugga caacggcacc uccgugaaga      300

uggaccagca ccggggcuuc cugcacaacc aggccaagaa ccugcugugc ggcuucuacg      360

gccggcacgc cgagcugcgg uuccuggacc uggugcccuc ccugcagcug gaccccgccc      420
```

agaucuaccg ggugaccugg uucaucuccu gguccccug cuucuccugg ggcugcgccg 480

gcgaggugcg ggccuuccug caggagaaca cccacgugcg gcugcggauc uucgccgccc 540

ggaucuacga cuacgacccc cuguacaagg aggcccugca gaugcugcgg gacgccggcg 600

cccagguguc caucaugacc uacgacgagu ucaagcacug cugggacacc uucguggacc 660

accagggcug ccccuuccag cccugggacg gccuggacga gcacucccag gcccuguccg 720

gccggcugcg ggccauccug cagaaccagg gcaacuccgg cuccgagacc cccggcaccu 780

ccgaguccgc cacccccgag uccgcagcgu ucaaaccaaa ucccaucaac uacauccugg 840

gccuggccau cggcaucgcc uccgugggcu gggccauggu ggagaucgac gaggaggaga 900

accccauccg gcugaucgac cugggcgugc ggguguucga gcgggccgag gugcccaaga 960

ccggcgacuc ccuggccaug gcccggcggc uggcccgguc cgugcggcgg cugacccggc 1020

ggcgggccca ccggcugcug cgggcccggc ggcugcugaa gcgggagggc gugcugcagg 1080

ccgccgacuu cgacgagaac ggccugauca agucccugcc caacaccccc uggcagcugc 1140

gggccgccgc ccuggaccgg aagcugaccc cccuggagug guccgccgug cugcugcacc 1200

ugaucaagca ccggggcuac cugucccagc ggaagaacga gggcgagacc gccgacaagg 1260

agcugggcgc ccugcugaag ggcguggcca acaacgccca cgcccugcag accggcgacu 1320

uccggacccc cgccgagcug gcccugaaca aguucgagaa ggaguccggc cacauccgga 1380

accagcgggg cgacuacucc cacaccuucu cccggaagga ccugcaggcc gagcugaucc 1440

ugcuguucga gaagcagaag gaguucggca acccccacgu guccggcggc cugaaggagg 1500

gcaucgagac ccugcugaug acccagcggc ccgcccuguc cggcgacgcc gugcagaaga 1560

ugcugggcca cugcaccuuc gagcccgccg agcccaaggc cgccaagaac accuacaccg 1620

ccgagcgguu caucuggcug accaagcuga acaaccugcg gauccuggag cagggcuccg 1680

agcggccccu gaccgacacc gagcgggcca cccugaugga cgagcccuac cggaagucca 1740

agcugaccua cgcccaggcc cggaagcugc ugggccugga ggacaccgcc uucuucaagg 1800

gccugcggua cggcaaggac aacgccgagg ccuccacccu gauggagaug aaggccuacc 1860

acgccaucuc ccgggcccug gagaaggagg gccugaagga caagaagucc ccccugaacc 1920

uguccuccga gcugcaggac gagaucggca ccgccuucuc ccuguucaag accgacgagg 1980

acaucaccgg ccggcugaag gaccgggugc agcccgagau ccuggaggcc cugcugaagc 2040

acaucuccuu cgacaaguuc gugcagaucu cccugaaggc ccugcggcgg aucgugcccc 2100

ugauggagca gggcaagcgg uacgacgagg ccugcgccga gaucuacggc gaccacuacg 2160

gcaagaagaa caccgaggag aagaucuacc ugccccccau ccccgccgac gagauccgga 2220

accccguggu gcugcgggcc cugucccagg cccggaaggu gaucaacggc guggugcggc 2280

gguacggcuc ccccgcccgg auccacaucg agaccgcccg ggaggugggc aaguccuuca 2340

aggaccggaa ggagaucgag aagcggcagg aggagaaccg gaaggaccgg gagaaggccg 2400

ccgccaaguu ccgggaguac uuccccaacu ucgugggcga gcccaagucc aaggacaucc 2460

ugaagcugcg gcuguacgag cagcagcacg gcaagugccu guacuccggc aaggagauca 2520

accuggugcg gcugaacgag aagggcuacg uggagaucga ccacgcccug cccuucuccc 2580

ggaccuggga cgacuccuuc aacaacaagg ugcuggugcu gggcuccgag aaccagaaca 2640

agggcaacca gacccccuac gaguacuuca acggcaagga caacucccgg gaguggcagg 2700

aguucaaggc ccggguggag accucccggu ucccccgguc caagaagcag cggauccugc 2760

ugcagaaguu cgacgaggac ggcuucaagg agugcaaccu gaacgacacc cgguacguga 2820

accgcuuccu gugccaguuc guggccgacc acauccugcu gaccggcaag ggcaagcggc 2880

ggguguucgc cuccaacggc cagaucacca accugcugcg gggcuucugg ggccugcgga 2940

aggugcgggc cgagaacgac cggcaccacg cccuggacgc cgugguggug gccugcucca 3000

ccguggccau gcagcagaag aucacccggu ucgugcggua caaggagaug aacgccuucg 3060

acggcaagac caucgacaag gagaccggca aggugcugca ccagaagacc cacuuccccc 3120

agcccuggga guucuucgcc caggagguga ugauccgggu guucggcaag cccgacggca 3180

agcccgaguu cgaggaggcc gacacccccg agaagcugcg gacccugcug gccgagaagc 3240

uguccucccg gcccgaggcc gugcacgagu acgugacccc ccuguucgug ucccgggccc 3300

ccaaccggaa gauguccggc gcccacaagg acacccugcg guccgccaag cgguucguga 3360

agcacaacga gaagaucucc gugaagcggg uguggcugac cgagaucaag cuggccgacc 3420

uggagaacau ggugaacuac aagaacggcc gggagaucga gcuguacgag gcccugaagg 3480

cccggcugga ggccuacggc ggcaacgcca agcaggccuu cgaccccaag gacaaccccu 3540

ucuacaagaa gggcggccag cuggugaagg ccgugcgggu ggagaagacc caggaguccg 3600

gcgugcugcu gaacaagaag aacgccuaca ccaucgccga caacggcgac auggugcggg 3660

uggacguguu cugcaaggug gacaagaagg gcaagaacca guacuucauc gugcccaucu 3720

acgccuggca gguggccgag aacauccugc ccgacaucga cugcaagggc uaccggaucg 3780

acgacuccua caccuucugc uucucccugc acaaguacga ccugaucgcc uuccagaagg 3840

acgagaaguc caagguggag uucgccuacu acaucaacug cgacuccucc aacggccggu 3900

ucuaccuggc cuggcacgac aagggcucca aggagcagca guuccggauc uccacccaga 3960

accuggugcu gauccagaag uaccagguga acgagcuggg caaggagauc cggcccugcc 4020

ggcugaagaa gcggcccccc gugcgguagu gacuagcacc agccucaaga acacccgaau 4080

ggagucucua agcuacauaa uaccaacuua cacuuuacaa aauguugucc cccaaaaugu 4140

agccauucgu aucugcuccu aauaaaaaga aaguuucuuc acauucu 4187

<210> 311
<211> 4005
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 311
auggacggcu ccggcggcgg cucccccaag aagaagcgga agguggagga caagcggccc 60

gccgccacca agaaggccgg ccaggccaag aagaagaagg gcggcuccgg cggcggcgag 120

gccucccccg ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc 180

aacaacggca ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac 240

ggcaccuccg ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug 300

cugugcggcu ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug 360

cagcuggacc ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc 420

uccuggggcu gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug 480

cggaucuucg ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug 540

cugcgggacg ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg 600

gacaccuucg uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac 660

ucccaggccc uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc 720

gagacccccg gcaccuccga guccgccacc cccgaguccg cagcguucaa accaaauccc 780

aucaacuaca uccugggccu ggccaucggc aucgccuccg ugggcugggc caugguggag 840

aucgacgagg aggagaaccc cauccggcug aucgaccugg gcgugcgggu guucgagcgg 900

gccgaggugc ccaagaccgg cgacucccug gccauggccc ggcggcuggc ccgguccgug 960

cggcggcuga cccggcggcg ggcccaccgg cugcugcggg cccggcggcu gcugaagcgg 1020

gagggcgugc ugcaggccgc cgacuucgac gagaacggcc ugaucaaguc ccugcccaac 1080

acccccuggc agcugcgggc cgccgcccug gaccggaagc ugacccccu ggaguggucc 1140

gccgugcugc ugcaccugau caagcaccgg ggcuaccugu cccagcggaa gaacgagggc 1200

gagaccgccg acaaggagcu gggcgcccug cugaagggcg uggccaacaa cgcccacgcc 1260

cugcagaccg gcgacuuccg gacccccgcc gagcuggccc ugaacaaguu cgagaaggag 1320

uccggccaca uccggaacca gcggggcgac uacucccaca ccuucucccg gaaggaccug 1380

caggccgagc ugauccugcu guucgagaag cagaaggagu ucggcaaccc ccacgugucc 1440

ggcggccuga aggagggcau cgagacccug cugaugaccc agcggcccgc ccuguccggc 1500

gacgccgugc agaagaugcu gggccacugc accuucgagc ccgccgagcc caaggccgcc 1560

aagaacaccu acaccgccga gcgguucauc uggcugacca agcugaacaa ccugcggauc 1620

cuggagcagg gcuccgagcg gccccugacc gacaccgagc gggccacccu gauggacgag 1680

cccuaccgga aguccaagcu gaccuacgcc caggcccgga agcugcuggg ccuggaggac 1740

accgccuucu ucaagggccu gcgguacggc aaggacaacg ccgaggccuc cacccugaug 1800

gagaugaagg ccuaccacgc caucucccgg gcccuggaga aggagggccu gaaggacaag 1860

aagucccccc ugaaccuguc cuccgagcug caggacgaga ucggcaccgc cuucucccug 1920

uucaagaccg acgaggacau caccggccgg cugaaggacc gggugcagcc cgagauccug 1980

gaggcccugc ugaagcacau cuccuucgac aaguucgugc agaucucccu gaaggcccug 2040

cggcggaucg ugccccugau ggagcagggc aagcgguacg acgaggccug cgccgagauc 2100

uacggcgacc acuacggcaa gaagaacacc gaggagaaga ucuaccugcc ccccaucccc 2160

gccgacgaga uccggaaccc cguggugcug cgggcccugu cccaggcccg gaaggugauc 2220

aacggcgugg ugcggcggua cggcuccccc gcccggaucc acaucgagac cgcccgggag 2280

gugggcaagu ccuucaagga ccggaaggag aucgagaagc ggcaggagga gaaccggaag 2340

gaccgggaga aggccgccgc caaguuccgg gaguacuucc ccaacuucgu gggcgagccc 2400

aaguccaagg acauccugaa gcugcggcug uacgagcagc agcacggcaa gugccuguac 2460

uccggcaagg agaucaaccu ggugcggcug aacgagaagg gcuacgugga gaucgaccac 2520

gcccugcccu ucucccggac cugggacgac uccuucaaca acaaggugcu ggugcugggc 2580

uccgagaacc agaacaaggg caaccagacc cccuacgagu acuucaacgg caaggacaac 2640

ucccgggagu ggcaggaguu caaggcccgg guggagaccu cccgguuccc ccgguccaag 2700

aagcagcgga uccugcugca gaaguucgac gaggacggcu ucaaggagug caaccugaac 2760

gacacccggu acgugaaccg cuuccugugc caguucgugg ccgaccacau ccugcugacc 2820

ggcaagggca agcggcgggu guucgccucc aacggccaga ucaccaaccu gcugcggggc 2880

uucuggggcc ugcggaaggu gcgggccgag aacgaccggc accacgcccu ggacgccgug 2940

gugguggccu gcuccaccgu ggccaugcag cagaagauca cccgguucgu gcgguacaag 3000

gagaugaacg ccuucgacgg caagaccauc gacaaggaga ccggcaaggu gcugcaccag 3060

aagacccacu ucccccagcc cugggaguuc uucgcccagg aggugaugau ccggguguuc 3120

ggcaagcccg acggcaagcc cgaguucgag gaggccgaca cccccgagaa gcugcggacc 3180

cugcuggccg agaagcuguc cucccggccc gaggccgugc acgaguacgu gacccccug 3240

uucguguccc gggccccaa ccggaagaug uccggcgccc acaaggacac ccugcggucc 3300

gccaagcggu ucgugaagca caacgagaag aucuccguga agcgggugug gcugaccgag 3360

aucaagcugg ccgaccugga gaacauggug aacuacaaga acggccggga gaucgagcug 3420

uacgaggccc ugaaggcccg gcuggaggcc uacggcggca acgccaagca ggccuucgac 3480

cccaaggaca accccuucua caagaagggc ggccagcugg ugaaggccgu gcggguggag 3540

aagacccagg aguccggcgu gcugcugaac aagaagaacg ccuacaccau cgccgacaac 3600

ggcgacaugg ugcgggugga cguguucugc aagguggaca agaagggcaa gaaccaguac 3660

uucaucgugc ccaucuacgc cuggcaggug gccgagaaca uccugcccga caucgacugc 3720

aagggcuacc ggaucgacga cuccuacacc uucugcuucu cccugcacaa guacgaccug 3780

aucgccuucc agaaggacga gaaguccaag guggaguucg ccuacuacau caacugcgac 3840

uccuccaacg gccgguucua ccuggccugg cacgacaagg gcuccaagga gcagcaguuc 3900

cggaucucca cccagaaccu ggugcugauc cagaaguacc aggugaacga gcugggcaag 3960

gagauccggc ccugccggcu gaagaagcgg ccccccgugc gguag 4005

<210> 312
<211> 1334
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 312

Met Asp Gly Ser Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val Glu
1 5 10 15

Asp Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
20 25 30

Lys Gly Gly Ser Gly Gly Gly Glu Ala Ser Pro Ala Ser Gly Pro Arg
35 40 45

His Leu Met Asp Pro His Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile
50 55 60

Gly Arg His Lys Thr Tyr Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn
65 70 75 80

```
Gly Thr Ser Val Lys Met Asp Gln His Arg Gly Phe Leu His Asn Gln
                85                  90                  95

Ala Lys Asn Leu Leu Cys Gly Phe Tyr Gly Arg His Ala Glu Leu Arg
            100                 105                 110

Phe Leu Asp Leu Val Pro Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr
        115                 120                 125

Arg Val Thr Trp Phe Ile Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys
    130                 135                 140

Ala Gly Glu Val Arg Ala Phe Leu Gln Glu Asn Thr His Val Arg Leu
145                 150                 155                 160

Arg Ile Phe Ala Ala Arg Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu
                165                 170                 175

Ala Leu Gln Met Leu Arg Asp Ala Gly Ala Gln Val Ser Ile Met Thr
            180                 185                 190

Tyr Asp Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp His Gln Gly
        195                 200                 205

Cys Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser Gln Ala Leu
    210                 215                 220

Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn Ser Gly Ser
225                 230                 235                 240

Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Ala Ala Phe
                245                 250                 255

Lys Pro Asn Pro Ile Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Ala
            260                 265                 270
```

```
Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Glu Glu Asn Pro Ile
        275                 280                 285

Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg Ala Glu Val Pro
    290                 295                 300

Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu Ala Arg Ser Val
305                 310                 315                 320

Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu Arg Ala Arg Arg
                325                 330                 335

Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp Phe Asp Glu Asn
            340                 345                 350

Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln Leu Arg Ala Ala
        355                 360                 365

Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser Ala Val Leu Leu
    370                 375                 380

His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg Lys Asn Glu Gly
385                 390                 395                 400

Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys Gly Val Ala Asn
                405                 410                 415

Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr Pro Ala Glu Leu
            420                 425                 430

Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile Arg Asn Gln Arg
        435                 440                 445

Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu Gln Ala Glu Leu
    450                 455                 460

Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn Pro His Val Ser
465                 470                 475                 480
```

```
Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met Thr Gln Arg Pro
                485                 490                 495

Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly His Cys Thr Phe
            500                 505                 510

Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr Thr Ala Glu Arg
        515                 520                 525

Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile Leu Glu Gln Gly
    530                 535                 540

Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr Leu Met Asp Glu
545                 550                 555                 560

Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala Arg Lys Leu Leu
                565                 570                 575

Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg Tyr Gly Lys Asp
            580                 585                 590

Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala Tyr His Ala Ile
        595                 600                 605

Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys Lys Ser Pro Leu
    610                 615                 620

Asn Leu Ser Ser Glu Leu Gln Asp Glu Ile Gly Thr Ala Phe Ser Leu
625                 630                 635                 640

Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys Asp Arg Val Gln
                645                 650                 655

Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser Phe Asp Lys Phe
            660                 665                 670
```

```
Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val Pro Leu Met Glu
        675                 680                 685

Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile Tyr Gly Asp His
    690                 695                 700

Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu Pro Pro Ile Pro
705                 710                 715                 720

Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala Leu Ser Gln Ala
                725                 730                 735

Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly Ser Pro Ala Arg
            740                 745                 750

Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser Phe Lys Asp Arg
        755                 760                 765

Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys Asp Arg Glu Lys
    770                 775                 780

Ala Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe Val Gly Glu Pro
785                 790                 795                 800

Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu Gln Gln His Gly
                805                 810                 815

Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Val Arg Leu Asn Glu
            820                 825                 830

Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe Ser Arg Thr Trp
        835                 840                 845

Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly Ser Glu Asn Gln
    850                 855                 860

Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn Gly Lys Asp Asn
865                 870                 875                 880
```

```
Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu Thr Ser Arg Phe
               885                 890                 895


Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys Phe Asp Glu Asp
            900                 905                 910


Gly Phe Lys Glu Cys Asn Leu Asn Asp Thr Arg Tyr Val Asn Arg Phe
        915                 920                 925


Leu Cys Gln Phe Val Ala Asp His Ile Leu Leu Thr Gly Lys Gly Lys
    930                 935                 940


Arg Arg Val Phe Ala Ser Asn Gly Gln Ile Thr Asn Leu Leu Arg Gly
945                 950                 955                 960


Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp Arg His His Ala
               965                 970                 975


Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala Met Gln Gln Lys
            980                 985                 990


Ile Thr Arg Phe Val Arg Tyr Lys  Glu Met Asn Ala Phe  Asp Gly Lys
        995                 1000                1005


Thr Ile  Asp Lys Glu Thr Gly  Lys Val Leu His Gln  Lys Thr His
    1010                1015                1020


Phe Pro  Gln Pro Trp Glu Phe  Phe Ala Gln Glu Val  Met Ile Arg
    1025                1030                1035


Val Phe  Gly Lys Pro Asp Gly  Lys Pro Glu Phe Glu  Glu Ala Asp
    1040                1045                1050


Thr Pro  Glu Lys Leu Arg Thr  Leu Leu Ala Glu Lys  Leu Ser Ser
    1055                1060                1065
```

```
Arg Pro  Glu Ala Val His Glu  Tyr Val Thr Pro Leu  Phe Val Ser
    1070                1075                1080

Arg Ala  Pro Asn Arg Lys Met  Ser Gly Ala His Lys  Asp Thr Leu
    1085                1090                1095

Arg Ser  Ala Lys Arg Phe Val  Lys His Asn Glu Lys  Ile Ser Val
    1100                1105                1110

Lys Arg  Val Trp Leu Thr Glu  Ile Lys Leu Ala Asp  Leu Glu Asn
    1115                1120                1125

Met Val  Asn Tyr Lys Asn Gly  Arg Glu Ile Glu Leu  Tyr Glu Ala
    1130                1135                1140

Leu Lys  Ala Arg Leu Glu Ala  Tyr Gly Gly Asn Ala  Lys Gln Ala
    1145                1150                1155

Phe Asp  Pro Lys Asp Asn Pro  Phe Tyr Lys Lys Gly  Gly Gln Leu
    1160                1165                1170

Val Lys  Ala Val Arg Val Glu  Lys Thr Gln Glu Ser  Gly Val Leu
    1175                1180                1185

Leu Asn  Lys Lys Asn Ala Tyr  Thr Ile Ala Asp Asn  Gly Asp Met
    1190                1195                1200

Val Arg  Val Asp Val Phe Cys  Lys Val Asp Lys Lys  Gly Lys Asn
    1205                1210                1215

Gln Tyr  Phe Ile Val Pro Ile  Tyr Ala Trp Gln Val  Ala Glu Asn
    1220                1225                1230

Ile Leu  Pro Asp Ile Asp Cys  Lys Gly Tyr Arg Ile  Asp Asp Ser
    1235                1240                1245

Tyr Thr  Phe Cys Phe Ser Leu  His Lys Tyr Asp Leu  Ile Ala Phe
    1250                1255                1260
```

```
Gln Lys  Asp Glu Lys Ser Lys  Val Glu Phe Ala Tyr  Tyr Ile Asn
    1265                1270                1275


Cys Asp  Ser Ser Asn Gly Arg  Phe Tyr Leu Ala Trp  His Asp Lys
    1280                1285                1290


Gly Ser  Lys Glu Gln Gln Phe  Arg Ile Ser Thr Gln  Asn Leu Val
    1295                1300                1305


Leu Ile  Gln Lys Tyr Gln Val  Asn Glu Leu Gly Lys  Glu Ile Arg
    1310                1315                1320


Pro Cys  Arg Leu Lys Lys Arg  Pro Pro Val Arg
    1325                1330


<210>  313
<211>  1334
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  313

Met Asp Gly Ser Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val Glu
1               5                   10                  15


Asp Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
            20                  25                  30


Lys Gly Gly Ser Gly Gly Gly Glu Ala Ser Pro Ala Ser Gly Pro Arg
        35                  40                  45


His Leu Met Asp Pro His Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile
    50                  55                  60


Gly Arg His Lys Thr Tyr Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn
65                  70                  75                  80
```

```
Gly Thr Ser Val Lys Met Asp Gln His Arg Gly Phe Leu His Asn Gln
                85                  90                  95

Ala Lys Asn Leu Leu Cys Gly Phe Tyr Gly Arg His Ala Glu Leu Arg
            100                 105                 110

Phe Leu Asp Leu Val Pro Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr
        115                 120                 125

Arg Val Thr Trp Phe Ile Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys
    130                 135                 140

Ala Gly Glu Val Arg Ala Phe Leu Gln Glu Asn Thr His Val Arg Leu
145                 150                 155                 160

Arg Ile Phe Ala Ala Arg Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu
                165                 170                 175

Ala Leu Gln Met Leu Arg Asp Ala Gly Ala Gln Val Ser Ile Met Thr
            180                 185                 190

Tyr Asp Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp His Gln Gly
        195                 200                 205

Cys Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser Gln Ala Leu
    210                 215                 220

Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn Gly Thr Lys
225                 230                 235                 240

Asp Ser Thr Lys Asp Ile Pro Glu Thr Pro Ser Lys Asp Ala Ala Phe
                245                 250                 255

Lys Pro Asn Pro Ile Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Ala
            260                 265                 270
```

```
Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Glu Glu Asn Pro Ile
        275                 280                 285

Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg Ala Glu Val Pro
    290                 295                 300

Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu Ala Arg Ser Val
305                 310                 315                 320

Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu Arg Ala Arg Arg
                325                 330                 335

Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp Phe Asp Glu Asn
            340                 345                 350

Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln Leu Arg Ala Ala
        355                 360                 365

Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser Ala Val Leu Leu
    370                 375                 380

His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg Lys Asn Glu Gly
385                 390                 395                 400

Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys Gly Val Ala Asn
                405                 410                 415

Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr Pro Ala Glu Leu
            420                 425                 430

Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile Arg Asn Gln Arg
        435                 440                 445

Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu Gln Ala Glu Leu
    450                 455                 460

Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn Pro His Val Ser
465                 470                 475                 480
```

```
Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met Thr Gln Arg Pro
                485                 490                 495

Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly His Cys Thr Phe
            500                 505                 510

Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr Thr Ala Glu Arg
        515                 520                 525

Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile Leu Glu Gln Gly
    530                 535                 540

Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr Leu Met Asp Glu
545                 550                 555                 560

Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala Arg Lys Leu Leu
                565                 570                 575

Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg Tyr Gly Lys Asp
            580                 585                 590

Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala Tyr His Ala Ile
        595                 600                 605

Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys Lys Ser Pro Leu
    610                 615                 620

Asn Leu Ser Ser Glu Leu Gln Asp Glu Ile Gly Thr Ala Phe Ser Leu
625                 630                 635                 640

Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys Asp Arg Val Gln
                645                 650                 655

Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser Phe Asp Lys Phe
            660                 665                 670
```

```
Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val Pro Leu Met Glu
        675                 680                 685

Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile Tyr Gly Asp His
    690                 695                 700

Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu Pro Pro Ile Pro
705                 710                 715                 720

Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala Leu Ser Gln Ala
                725                 730                 735

Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly Ser Pro Ala Arg
            740                 745                 750

Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser Phe Lys Asp Arg
        755                 760                 765

Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys Asp Arg Glu Lys
    770                 775                 780

Ala Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe Val Gly Glu Pro
785                 790                 795                 800

Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu Gln Gln His Gly
                805                 810                 815

Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Val Arg Leu Asn Glu
            820                 825                 830

Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe Ser Arg Thr Trp
        835                 840                 845

Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly Ser Glu Asn Gln
    850                 855                 860

Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn Gly Lys Asp Asn
865                 870                 875                 880
```

```
Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu Thr Ser Arg Phe
               885                 890                 895


Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys Phe Asp Glu Asp
            900                 905                 910


Gly Phe Lys Glu Cys Asn Leu Asn Asp Thr Arg Tyr Val Asn Arg Phe
        915                 920                 925


Leu Cys Gln Phe Val Ala Asp His Ile Leu Leu Thr Gly Lys Gly Lys
    930                 935                 940


Arg Arg Val Phe Ala Ser Asn Gly Gln Ile Thr Asn Leu Leu Arg Gly
945                 950                 955                 960


Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp Arg His His Ala
               965                 970                 975


Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala Met Gln Gln Lys
            980                 985                 990


Ile Thr Arg Phe Val Arg Tyr Lys  Glu Met Asn Ala Phe  Asp Gly Lys
        995                 1000                1005


Thr Ile  Asp Lys Glu Thr Gly  Lys Val Leu His Gln  Lys Thr His
    1010                1015                1020


Phe Pro  Gln Pro Trp Glu Phe  Phe Ala Gln Glu Val  Met Ile Arg
    1025                1030                1035


Val Phe  Gly Lys Pro Asp Gly  Lys Pro Glu Phe Glu  Glu Ala Asp
    1040                1045                1050


Thr Pro  Glu Lys Leu Arg Thr  Leu Leu Ala Glu Lys  Leu Ser Ser
    1055                1060                1065
```

```
Arg Pro  Glu Ala Val His Glu  Tyr Val Thr Pro Leu  Phe Val Ser
    1070                1075                1080

Arg Ala  Pro Asn Arg Lys Met  Ser Gly Ala His Lys  Asp Thr Leu
    1085                1090                1095

Arg Ser  Ala Lys Arg Phe Val  Lys His Asn Glu Lys  Ile Ser Val
    1100                1105                1110

Lys Arg  Val Trp Leu Thr Glu  Ile Lys Leu Ala Asp  Leu Glu Asn
    1115                1120                1125

Met Val  Asn Tyr Lys Asn Gly  Arg Glu Ile Glu Leu  Tyr Glu Ala
    1130                1135                1140

Leu Lys  Ala Arg Leu Glu Ala  Tyr Gly Gly Asn Ala  Lys Gln Ala
    1145                1150                1155

Phe Asp  Pro Lys Asp Asn Pro  Phe Tyr Lys Lys Gly  Gly Gln Leu
    1160                1165                1170

Val Lys  Ala Val Arg Val Glu  Lys Thr Gln Glu Ser  Gly Val Leu
    1175                1180                1185

Leu Asn  Lys Lys Asn Ala Tyr  Thr Ile Ala Asp Asn  Gly Asp Met
    1190                1195                1200

Val Arg  Val Asp Val Phe Cys  Lys Val Asp Lys Lys  Gly Lys Asn
    1205                1210                1215

Gln Tyr  Phe Ile Val Pro Ile  Tyr Ala Trp Gln Val  Ala Glu Asn
    1220                1225                1230

Ile Leu  Pro Asp Ile Asp Cys  Lys Gly Tyr Arg Ile  Asp Asp Ser
    1235                1240                1245

Tyr Thr  Phe Cys Phe Ser Leu  His Lys Tyr Asp Leu  Ile Ala Phe
    1250                1255                1260
```

Gln Lys Asp Glu Lys Ser Lys Val Glu Phe Ala Tyr Tyr Ile Asn
1265 1270 1275

Cys Asp Ser Ser Asn Gly Arg Phe Tyr Leu Ala Trp His Asp Lys
1280 1285 1290

Gly Ser Lys Glu Gln Gln Phe Arg Ile Ser Thr Gln Asn Leu Val
1295 1300 1305

Leu Ile Gln Lys Tyr Gln Val Asn Glu Leu Gly Lys Glu Ile Arg
1310 1315 1320

Pro Cys Arg Leu Lys Lys Arg Pro Pro Val Arg
1325 1330

<210> 314
<211> 5990
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 314
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cuccggcucc gagacccccg 660

gcaccuccga guccgccacc cccgaguccg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccca 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuacccca 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

ugguggggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucguguggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac uccggcggcu 4800

ccggcggcuc cggcggcucc accaaccugu ccgacaucau cgagaaggag accggcaagc 4860

agcuggugau ccaggagucc auccugaugc ugcccgagga gguggaggag gugaucggca 4920

acaagcccga guccgacauc cuggugcaca ccgccuacga cgaguccacc gacgagaacg 4980

ugaugcugcu gaccuccgac gcccccgagu acaagcccug ggcccuggug auccaggacu 5040

ccaacggcga gaacaagauc aagaugcugu ccggcggcuc cggcggcucc ggcggcucca 5100

ccaaccuguc cgacaucauc gagaaggaga ccggcaagca gcuggugauc caggagucca 5160

uccugaugcu gcccgaggag guggaggagg ugaucggcaa caagcccgag uccgacaucc 5220

uggugcacac cgccuacgac gaguccaccg acgagaacgu gaugcugcug accuccgacg 5280

cccccgagua caagcccugg gcccugguga uccaggacuc caacggcgag aacaagauca 5340

agaugcuguc cggcggcucc aagcggaccg ccgacggcuc cgaguucgag cccaagaaga 5400

agcggaaggu guccgagucc gccacccccg aguccguguc cggcuggcgg cuguucaaga 5460

agaucuccug auagcuagca ccagccucaa gaacacccga auggagucuc uaagcuacau 5520

aauaccaacu uacacuuuac aaaauguugu cccccaaaau guagccauuc guaucugcuc 5580

cuaauaaaaa gaaaguuucu ucacauucuc ucgagaaaaa aaaaaaaugg aaaaaaaaaa 5640

aacggaaaaa aaaaaaaggu aaaaaaaaaa aauauaaaaa aaaaaaacau aaaaaaaaaa 5700

aacgaaaaaa aaaaaacgua aaaaaaaaaa acucaaaaaa aaaaaagaua aaaaaaaaaa 5760

accuaaaaaa aaaaaaugua aaaaaaaaaa agggaaaaaa aaaaaacgca aaaaaaaaaa 5820

acacaaaaaa aaaaaaugca aaaaaaaaaa aucgaaaaaa aaaaaaucua aaaaaaaaaa 5880

acgaaaaaaa aaaaacccaa aaaaaaaaaa gacaaaaaaa aaaaauagaa aaaaaaaaaa 5940

guuaaaaaaa aaaaacugaa aaaaaaaaaa uuuaaaaaaa aaaaaucuag 5990

<210> 315
<211> 4563
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 315
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggac ggcuccggcg 60

gcggcucccc caagaagaag cggaaggugg aggacaagcg gcccgccgcc accaagaagg 120

ccggccaggc caagaagaag aagggcggcu ccggcggcgg cgaggccucc cccgccuccg 180

gcccccggca ccugauggac ccccacaucu ucaccuccaa cuucaacaac ggcaucggcc 240

ggcacaagac cuaccugugc uacgaggugg agcggcugga caacggcacc uccgugaaga 300

uggaccagca ccggggcuuc cugcacaacc aggccaagaa ccugcugugc ggcuucuacg 360

gccggcacgc cgagcugcgg uuccuggacc uggugcccuc ccugcagcug gaccccgccc 420

agaucuaccg ggugaccugg uucaucuccu gguccccccug cuucuccugg ggcugcgccg 480

gcgaggugcg ggccuuccug caggagaaca cccacgugcg gcugcggauc uucgccgccc 540

ggaucuacga cuacgacccc cuguacaagg aggcccugca gaugcugcgg gacgccggcg 600

cccagguguc caucaugacc uacgacgagu ucaagcacug cugggacacc uucguggacc 660

accagggcug ccccuuccag cccugggacg gccuggacga gcacucccag gcccuguccg 720

gccggcugcg ggccauccug cagaaccagg gcaacuccgg cuccgagacc cccggcaccu 780

ccgaguccgc cacccccgag uccgcagcgu ucaaaccaaa ucccaucaac uacauccugg 840

gccuggccau cggcaucgcc uccgugggcu gggccauggu ggagaucgac gaggaggaga 900

accccauccg gcugaucgac cugggcgugc ggguguucga gcgggccgag gugcccaaga 960

ccggcgacuc ccuggccaug gcccggcggc uggcccgguc cgugcggcgg cugacccggc 1020

ggcgggccca ccggcugcug cgggcccggc ggcugcugaa gcgggagggc gugcugcagg 1080

ccgccgacuu cgacgagaac ggccugauca agucccugcc caacaccccc uggcagcugc 1140

gggccgccgc ccuggaccgg aagcugaccc cccuggagug guccgccgug cugcugcacc 1200

ugaucaagca ccggggcuac cugucccagc ggaagaacga gggcgagacc gccgacaagg 1260

agcugggcgc ccugcugaag ggcguggcca acaacgccca cgcccugcag accggcgacu 1320

uccggacccc cgccgagcug gcccugaaca aguucgagaa ggaguccggc cacauccgga 1380

accagcgggg cgacuacucc cacaccuucu cccggaagga ccugcaggcc gagcugaucc 1440

ugcuguucga gaagcagaag gaguucggca acccccacgu guccggcggc cugaaggagg 1500

gcaucgagac ccugcugaug acccagcggc ccgcccuguc cggcgacgcc gugcagaaga 1560

ugcugggcca cugcaccuuc gagcccgccg agcccaaggc cgccaagaac accuacaccg 1620

ccgagcgguu caucuggcug accaagcuga acaaccugcg gauccuggag cagggcuccg 1680

agcggccccu gaccgacacc gagcgggcca cccugaugga cgagcccuac cggaagucca 1740

agcugaccua cgcccaggcc cggaagcugc ugggccugga ggacaccgcc uucuucaagg 1800

gccugcggua cggcaaggac aacgccgagg ccuccacccu gauggagaug aaggccuacc 1860

acgccaucuc ccgggcccug gagaaggagg gccugaagga caagaagucc ccccugaacc 1920

uguccuccga gcugcaggac gagaucggca ccgccuucuc ccuguucaag accgacgagg 1980

acaucaccgg ccggcugaag gaccgggugc agcccgagau ccuggaggcc cugcugaagc 2040

acaucuccuu cgacaaguuc gugcagaucu cccugaaggc ccugcggcgg aucgugcccc 2100

ugauggagca gggcaagcgg uacgacgagg ccugcgccga gaucuacggc gaccacuacg 2160

gcaagaagaa caccgaggag aagaucuacc ugccccccau ccccgccgac gagauccgga 2220

accccguggu gcugcgggcc cugucccagg cccggaaggu gaucaacggc guggugcggc 2280

gguacggcuc ccccgcccgg auccacaucg agaccgcccg ggaggugggc aaguccuuca 2340

aggaccggaa ggagaucgag aagcggcagg aggagaaccg gaaggaccgg gagaaggccg 2400

ccgccaaguu ccgggaguac uuccccaacu ucgugggcga gcccaagucc aaggacaucc 2460

ugaagcugcg gcuguacgag cagcagcacg gcaagugccu guacuccggc aaggagauca 2520

accuggugcg gcugaacgag aagggcuacg uggagaucga ccacgcccug cccuucuccc 2580

ggaccuggga cgacuccuuc aacaacaagg ugcuggugcu gggcuccgag aaccagaaca 2640

agggcaacca gacccccuac gaguacuuca acggcaagga caacucccgg gaguggcagg 2700

aguucaaggc ccggguggag accucccggu ucccccgguc caagaagcag cggauccugc 2760

ugcagaaguu cgacgaggac ggcuucaagg agugcaaccu gaacgacacc cgguacguga 2820

accgcuuccu gugccaguuc guggccgacc acauccugcu gaccggcaag ggcaagcggc 2880

ggguguucgc cuccaacggc cagaucacca accugcugcg gggcuucugg ggccugcgga 2940

aggugcgggc cgagaacgac cggcaccacg cccuggacgc cgugguggug gccugcucca 3000

ccguggccau gcagcagaag aucacccggu ucgugcggua caaggagaug aacgccuucg 3060

acggcaagac caucgacaag gagaccggca aggugcugca ccagaagacc cacuuccccc 3120

agcccuggga guucuucgcc caggagguga ugauccgggu guucggcaag cccgacggca 3180

agcccgaguu cgaggaggcc gacacccccg agaagcugcg gacccugcug gccgagaagc 3240

uguccucccg gcccgaggcc gugcacgagu acgugacccc ccuguucgug ucccgggccc 3300

ccaaccggaa gauguccggc gcccacaagg acacccugcg guccgccaag cgguucguga 3360

agcacaacga gaagaucucc gugaagcggg uguggcugac cgagaucaag cuggccgacc 3420

uggagaacau ggugaacuac aagaacggcc gggagaucga gcuguacgag gcccugaagg 3480

cccggcugga ggccuacggc ggcaacgcca agcaggccuu cgaccccaag gacaaccccu 3540

ucuacaagaa gggcggccag cuggugaagg ccgugcgggu ggagaagacc caggaguccg 3600

gcgugcugcu gaacaagaag aacgccuaca ccaucgccga caacggcgac auggugcggg 3660

uggacguguu cugcaaggug gacaagaagg gcaagaacca guacuucauc gugcccaucu 3720

acgccuggca gguggccgag aacauccugc ccgacaucga cugcaagggc uaccggaucg 3780

acgacuccua caccuucugc uucucccugc acaaguacga ccugaucgcc uuccagaagg 3840

acgagaaguc caagguggag uucgccuacu acaucaacug cgacuccucc aacggccggu 3900

ucuaccuggc cuggcacgac aagggcucca aggagcagca guuccggauc uccacccaga 3960

accuggugcu gauccagaag uaccagguga acgagcuggg caaggagauc cggcccugcc 4020

ggcugaagaa gcggcccccc gugcgguagu gacuagcacc agccucaaga acacccgaau 4080

ggagucucua agcuacauaa uaccaacuua cacuuuacaa aauguugucc cccaaaaugu 4140

agccauucgu aucugcuccu aauaaaaaga aaguuucuuc acauucucuc gagaaaaaaa 4200

aaaaauggaa aaaaaaaaaa cggaaaaaaa aaaagguaaa aaaaaaaaau auaaaaaaaa 4260

aaacauaaaa aaaaaaaacg aaaaaaaaaa aacguaaaaa aaaaaaacuc aaaaaaaaaa 4320

agauaaaaaa aaaaaaccua aaaaaaaaaa auguaaaaaa aaaaaaggga aaaaaaaaaa 4380

cgcaaaaaaa aaaaacacaa aaaaaaaaaa ugcaaaaaaa aaaaaucgaa aaaaaaaaaa 4440

ucuaaaaaaa aaaaacgaaa aaaaaaaaac ccaaaaaaaa aaaagacaaa aaaaaaaaau 4500

agaaaaaaaa aaaguuaaaa aaaaaaaacu gaaaaaaaaa aaauuuaaaa aaaaaaaauc 4560

uag 4563

<210> 316
<211> 944
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 316
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccaugacc aaccuguccg 60

acaucaucga gaaggagacc ggcaagcagc uggugaucca ggaguccauc cugaugcugc 120

ccgaggaggu ggaggaggug aucggcaaca agcccgaguc cgacauccug gugcacaccg 180

ccuacgacga guccaccgac gagaacguga ugcugcugac cuccgacgcc cccgaguaca 240

agcccugggc ccuggugauc caggacucca acggcgagaa caagaucaag augcuguccg 300

gcggcuccaa gcggaccgcc gacggcuccg aguucgaguc ccccaagaag aagcggaagg 360

uggaguccga guccgccacc cccgaguccg uguccggcug gcggcuguuc aagaagaucu 420

ccugauagcu agcaccagcc ucaagaacac ccgaauggag ucucuaagcu acauaauacc 480

aacuuacacu uuacaaaaug uuguccccca aaauguagcc auucguaucu gcuccuaaua 540

aaaagaaagu uucuucacau ucucucgaga aaaaaaaaaa auggaaaaaa aaaaaacgga 600

aaaaaaaaaa agguaaaaaa aaaaaauaua aaaaaaaaaa acauaaaaaa aaaaaacgaa 660

aaaaaaaaaa cguaaaaaaa aaaaacucaa aaaaaaaaaa gauaaaaaaa aaaaaccuaa 720

aaaaaaaaaa uguaaaaaaa aaaaagggaa aaaaaaaaaa cgcaaaaaaa aaaaacacaa 780

aaaaaaaaaa ugcaaaaaaa aaaaaucgaa aaaaaaaaaa ucuaaaaaaa aaaaacgaaa 840

aaaaaaaaac ccaaaaaaaa aaaagacaaa aaaaaaaaau agaaaaaaaa aaaaguuaaa 900

aaaaaaaaac ugaaaaaaaa aaaauuuaaa aaaaaaaaau cuag 944

<210> 317

<400> 317
000

<210> 318

<400> 318
000

<210> 319

<400> 319
000

<210> 320
<211> 1120
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 320

Met Asp Gly Ser Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val Glu
1 5 10 15

Asp Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
20 25 30

Lys Gly Gly Ser Gly Gly Gly Ala Ala Phe Lys Pro Asn Pro Ile Asn
35 40 45

Tyr Ile Leu Gly Leu Asp Ile Gly Ile Ala Ser Val Gly Trp Ala Met
50 55 60

Val Glu Ile Asp Glu Glu Glu Asn Pro Ile Arg Leu Ile Asp Leu Gly
65 70 75 80

Val Arg Val Phe Glu Arg Ala Glu Val Pro Lys Thr Gly Asp Ser Leu
85 90 95

Ala Met Ala Arg Arg Leu Ala Arg Ser Val Arg Arg Leu Thr Arg Arg
100 105 110

Arg Ala His Arg Leu Leu Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly
115 120 125

Val Leu Gln Ala Ala Asp Phe Asp Glu Asn Gly Leu Ile Lys Ser Leu
130 135 140

Pro Asn Thr Pro Trp Gln Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu

```
145                 150                 155                 160

Thr Pro Leu Glu Trp Ser Ala Val Leu Leu His Leu Ile Lys His Arg
                165                 170                 175

Gly Tyr Leu Ser Gln Arg Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu
            180                 185                 190

Leu Gly Ala Leu Leu Lys Gly Val Ala Asn Asn Ala His Ala Leu Gln
        195                 200                 205

Thr Gly Asp Phe Arg Thr Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu
    210                 215                 220

Lys Glu Ser Gly His Ile Arg Asn Gln Arg Gly Asp Tyr Ser His Thr
225                 230                 235                 240

Phe Ser Arg Lys Asp Leu Gln Ala Glu Leu Ile Leu Leu Phe Glu Lys
                245                 250                 255

Gln Lys Glu Phe Gly Asn Pro His Val Ser Gly Gly Leu Lys Glu Gly
            260                 265                 270

Ile Glu Thr Leu Leu Met Thr Gln Arg Pro Ala Leu Ser Gly Asp Ala
        275                 280                 285

Val Gln Lys Met Leu Gly His Cys Thr Phe Glu Pro Ala Glu Pro Lys
    290                 295                 300

Ala Ala Lys Asn Thr Tyr Thr Ala Glu Arg Phe Ile Trp Leu Thr Lys
305                 310                 315                 320

Leu Asn Asn Leu Arg Ile Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr
                325                 330                 335

Asp Thr Glu Arg Ala Thr Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys
            340                 345                 350
```

```
Leu Thr Tyr Ala Gln Ala Arg Lys Leu Leu Gly Leu Glu Asp Thr Ala
        355                 360                 365

Phe Phe Lys Gly Leu Arg Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr
    370                 375                 380

Leu Met Glu Met Lys Ala Tyr His Ala Ile Ser Arg Ala Leu Glu Lys
385                 390                 395                 400

Glu Gly Leu Lys Asp Lys Lys Ser Pro Leu Asn Leu Ser Ser Glu Leu
                405                 410                 415

Gln Asp Glu Ile Gly Thr Ala Phe Ser Leu Phe Lys Thr Asp Glu Asp
            420                 425                 430

Ile Thr Gly Arg Leu Lys Asp Arg Val Gln Pro Glu Ile Leu Glu Ala
        435                 440                 445

Leu Leu Lys His Ile Ser Phe Asp Lys Phe Val Gln Ile Ser Leu Lys
    450                 455                 460

Ala Leu Arg Arg Ile Val Pro Leu Met Glu Gln Gly Lys Arg Tyr Asp
465                 470                 475                 480

Glu Ala Cys Ala Glu Ile Tyr Gly Asp His Tyr Gly Lys Lys Asn Thr
                485                 490                 495

Glu Glu Lys Ile Tyr Leu Pro Pro Ile Pro Ala Asp Glu Ile Arg Asn
            500                 505                 510

Pro Val Val Leu Arg Ala Leu Ser Gln Ala Arg Lys Val Ile Asn Gly
        515                 520                 525

Val Val Arg Arg Tyr Gly Ser Pro Ala Arg Ile His Ile Glu Thr Ala
    530                 535                 540

Arg Glu Val Gly Lys Ser Phe Lys Asp Arg Lys Glu Ile Glu Lys Arg
```

```
545                 550                 555                 560

Gln Glu Glu Asn Arg Lys Asp Arg Glu Lys Ala Ala Ala Lys Phe Arg
                565                 570                 575

Glu Tyr Phe Pro Asn Phe Val Gly Glu Pro Lys Ser Lys Asp Ile Leu
            580                 585                 590

Lys Leu Arg Leu Tyr Glu Gln Gln His Gly Lys Cys Leu Tyr Ser Gly
        595                 600                 605

Lys Glu Ile Asn Leu Val Arg Leu Asn Glu Lys Gly Tyr Val Glu Ile
    610                 615                 620

Asp His Ala Leu Pro Phe Ser Arg Thr Trp Asp Asp Ser Phe Asn Asn
625                 630                 635                 640

Lys Val Leu Val Leu Gly Ser Glu Asn Gln Asn Lys Gly Asn Gln Thr
                645                 650                 655

Pro Tyr Glu Tyr Phe Asn Gly Lys Asp Asn Ser Arg Glu Trp Gln Glu
            660                 665                 670

Phe Lys Ala Arg Val Glu Thr Ser Arg Phe Pro Arg Ser Lys Lys Gln
        675                 680                 685

Arg Ile Leu Leu Gln Lys Phe Asp Glu Asp Gly Phe Lys Glu Cys Asn
    690                 695                 700

Leu Asn Asp Thr Arg Tyr Val Asn Arg Phe Leu Cys Gln Phe Val Ala
705                 710                 715                 720

Asp His Ile Leu Leu Thr Gly Lys Gly Lys Arg Arg Val Phe Ala Ser
                725                 730                 735

Asn Gly Gln Ile Thr Asn Leu Leu Arg Gly Phe Trp Gly Leu Arg Lys
            740                 745                 750
```

```
Val Arg Ala Glu Asn Asp Arg His His Ala Leu Asp Ala Val Val Val
        755                 760                 765

Ala Cys Ser Thr Val Ala Met Gln Gln Lys Ile Thr Arg Phe Val Arg
    770                 775                 780

Tyr Lys Glu Met Asn Ala Phe Asp Gly Lys Thr Ile Asp Lys Glu Thr
785                 790                 795                 800

Gly Lys Val Leu His Gln Lys Thr His Phe Pro Gln Pro Trp Glu Phe
                805                 810                 815

Phe Ala Gln Glu Val Met Ile Arg Val Phe Gly Lys Pro Asp Gly Lys
            820                 825                 830

Pro Glu Phe Glu Glu Ala Asp Thr Pro Glu Lys Leu Arg Thr Leu Leu
        835                 840                 845

Ala Glu Lys Leu Ser Ser Arg Pro Glu Ala Val His Glu Tyr Val Thr
    850                 855                 860

Pro Leu Phe Val Ser Arg Ala Pro Asn Arg Lys Met Ser Gly Ala His
865                 870                 875                 880

Lys Asp Thr Leu Arg Ser Ala Lys Arg Phe Val Lys His Asn Glu Lys
                885                 890                 895

Ile Ser Val Lys Arg Val Trp Leu Thr Glu Ile Lys Leu Ala Asp Leu
            900                 905                 910

Glu Asn Met Val Asn Tyr Lys Asn Gly Arg Glu Ile Glu Leu Tyr Glu
        915                 920                 925

Ala Leu Lys Ala Arg Leu Glu Ala Tyr Gly Gly Asn Ala Lys Gln Ala
    930                 935                 940

Phe Asp Pro Lys Asp Asn Pro Phe Tyr Lys Lys Gly Gly Gln Leu Val
```

945 950 955 960

Lys Ala Val Arg Val Glu Lys Thr Gln Glu Ser Gly Val Leu Leu Asn
965 970 975

Lys Lys Asn Ala Tyr Thr Ile Ala Asp Asn Gly Asp Met Val Arg Val
980 985 990

Asp Val Phe Cys Lys Val Asp Lys Lys Gly Lys Asn Gln Tyr Phe Ile
995 1000 1005

Val Pro Ile Tyr Ala Trp Gln Val Ala Glu Asn Ile Leu Pro Asp
1010 1015 1020

Ile Asp Cys Lys Gly Tyr Arg Ile Asp Asp Ser Tyr Thr Phe Cys
1025 1030 1035

Phe Ser Leu His Lys Tyr Asp Leu Ile Ala Phe Gln Lys Asp Glu
1040 1045 1050

Lys Ser Lys Val Glu Phe Ala Tyr Tyr Ile Asn Cys Asp Ser Ser
1055 1060 1065

Asn Gly Arg Phe Tyr Leu Ala Trp His Asp Lys Gly Ser Lys Glu
1070 1075 1080

Gln Gln Phe Arg Ile Ser Thr Gln Asn Leu Val Leu Ile Gln Lys
1085 1090 1095

Tyr Gln Val Asn Glu Leu Gly Lys Glu Ile Arg Pro Cys Arg Leu
1100 1105 1110

Lys Lys Arg Pro Pro Val Arg
1115 1120

<210> 321
<211> 1593
<212> PRT

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 321

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15


Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30


Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45


Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60


Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80


Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95


Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110


Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125


Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140


Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160


Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175
```

```
Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Thr Lys Asp Ser Thr Lys Asp Ile
        195                 200                 205

Pro Glu Thr Pro Ser Lys Asp Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
```

```
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575
```

```
Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
```

```
    770                 775                 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
            820                 825                 830

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
                965                 970                 975
```

```
Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
```

```
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350
```

```
Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500

Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
```

1535 1540 1545

Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His
1550 1555 1560

Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln
1565 1570 1575

Leu Gly Gly Asp Gly Gly Gly Ser Pro Lys Lys Lys Arg Lys Val
1580 1585 1590

<210> 322
<211> 1593
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 322

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
20 25 30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
35 40 45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
50 55 60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65 70 75 80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
85 90 95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala

```
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Arg Asp Val Arg Gln Pro Glu Val
        195                 200                 205

Lys Glu Glu Lys Pro Glu Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300
```

```
Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
```

```
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575

Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700
```

```
Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
    770                 775                 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
            820                 825                 830

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
```

```
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
                965                 970                 975

Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095
```

```
Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
```

```
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470
```

```
Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485


Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500


Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515


Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530


Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1535                1540                1545


Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560


Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575


Leu Gly  Gly Asp Gly Gly Gly  Ser Pro Lys Lys Lys  Arg Lys Val
    1580                1585                1590


<210>  323
<211>  1593
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  323

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                  10                  15


Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30
```

```
Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Gly Lys Ser Ser Gly Ser Gly Ser
        195                 200                 205

Glu Ser Lys Ser Thr Ala Gly Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
```

```
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430
```

```
Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575

Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
```

625 630 635 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
645 650 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
660 665 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
675 680 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
690 695 700

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705 710 715 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
725 730 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
740 745 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
755 760 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
770 775 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785 790 795 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
805 810 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
820 825 830

```
Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
                965                 970                 975

Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
```

```
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215
```

```
Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
```

```
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500

Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1535                1540                1545

Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560

Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575

Leu Gly  Gly Asp Gly Gly Gly  Ser Pro Lys Lys Lys  Arg Lys Val
    1580                1585                1590
```

<210> 324
<211> 1593
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 324

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15


Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30


Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45


Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60


Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80


Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95


Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110


Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125


Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140


Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160
```

```
Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Thr Pro Gly Ser Pro Ala Gly Ser Pro
        195                 200                 205

Thr Ser Thr Glu Glu Gly Thr Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
```

```
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560
```

```
Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575

Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
```

```
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
    770                 775                 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
            820                 825                 830

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
                885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960
```

```
Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
               965                 970                 975

Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080

Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
```

```
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335
```

```
Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455

Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470

Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485

Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500

Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515

Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
```

```
    1520                1525                1530

Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1535                1540                1545

Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560

Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575

Leu Gly  Gly Asp Gly Gly Gly  Ser Pro Lys Lys Lys  Arg Lys Val
    1580                1585                1590
```

<210> 325
<211> 1593
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 325

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
```

```
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Ser Glu Pro Ala Thr Ser Gly Ser
        195                 200                 205

Glu Thr Pro Gly Thr Ser Thr Asp Lys Lys Tyr Ser Ile Gly Leu Ala
    210                 215                 220

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
225                 230                 235                 240

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
                245                 250                 255

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
            260                 265                 270

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
        275                 280                 285
```

```
Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
    290                 295                 300

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
305                 310                 315                 320

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
                325                 330                 335

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
            340                 345                 350

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
        355                 360                 365

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
    370                 375                 380

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
385                 390                 395                 400

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
                405                 410                 415

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
            420                 425                 430

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
        435                 440                 445

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
    450                 455                 460

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
465                 470                 475                 480

Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile
```

```
                485                 490                 495

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
            500                 505                 510

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
        515                 520                 525

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
    530                 535                 540

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
545                 550                 555                 560

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
                565                 570                 575

Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
            580                 585                 590

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
        595                 600                 605

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
    610                 615                 620

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
625                 630                 635                 640

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
                645                 650                 655

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
            660                 665                 670

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
        675                 680                 685
```

```
Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
    690                 695                 700

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
705                 710                 715                 720

Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
                725                 730                 735

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
            740                 745                 750

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
        755                 760                 765

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
    770                 775                 780

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
785                 790                 795                 800

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
                805                 810                 815

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
            820                 825                 830

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
        835                 840                 845

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
    850                 855                 860

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
865                 870                 875                 880

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
```

```
               885                 890                 895

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
            900                 905                 910

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
        915                 920                 925

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
    930                 935                 940

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
945                 950                 955                 960

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
               965                 970                 975

Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
            980                 985                 990

Glu Arg Met Lys Arg Ile Glu Glu  Gly Ile Lys Glu Leu  Gly Ser Gln
        995                 1000                1005

Ile Leu  Lys Glu His Pro Val  Glu Asn Thr Gln Leu  Gln Asn Glu
    1010                1015                1020

Lys Leu  Tyr Leu Tyr Tyr Leu  Gln Asn Gly Arg Asp  Met Tyr Val
    1025                1030                1035

Asp Gln  Glu Leu Asp Ile Asn  Arg Leu Ser Asp Tyr  Asp Val Asp
    1040                1045                1050

His Ile  Val Pro Gln Ser Phe  Leu Lys Asp Asp Ser  Ile Asp Asn
    1055                1060                1065

Lys Val  Leu Thr Arg Ser Asp  Lys Asn Arg Gly Lys  Ser Asp Asn
    1070                1075                1080
```

```
Val Pro  Ser Glu Glu Val Val  Lys Lys Met Lys Asn  Tyr Trp Arg
    1085                1090                1095

Gln Leu  Leu Asn Ala Lys Leu  Ile Thr Gln Arg Lys  Phe Asp Asn
    1100                1105                1110

Leu Thr  Lys Ala Glu Arg Gly  Gly Leu Ser Glu Leu  Asp Lys Ala
    1115                1120                1125

Gly Phe  Ile Lys Arg Gln Leu  Val Glu Thr Arg Gln  Ile Thr Lys
    1130                1135                1140

His Val  Ala Gln Ile Leu Asp  Ser Arg Met Asn Thr  Lys Tyr Asp
    1145                1150                1155

Glu Asn  Asp Lys Leu Ile Arg  Glu Val Lys Val Ile  Thr Leu Lys
    1160                1165                1170

Ser Lys  Leu Val Ser Asp Phe  Arg Lys Asp Phe Gln  Phe Tyr Lys
    1175                1180                1185

Val Arg  Glu Ile Asn Asn Tyr  His His Ala His Asp  Ala Tyr Leu
    1190                1195                1200

Asn Ala  Val Val Gly Thr Ala  Leu Ile Lys Lys Tyr  Pro Lys Leu
    1205                1210                1215

Glu Ser  Glu Phe Val Tyr Gly  Asp Tyr Lys Val Tyr  Asp Val Arg
    1220                1225                1230

Lys Met  Ile Ala Lys Ser Glu  Gln Glu Ile Gly Lys  Ala Thr Ala
    1235                1240                1245

Lys Tyr  Phe Phe Tyr Ser Asn  Ile Met Asn Phe Phe  Lys Thr Glu
    1250                1255                1260

Ile Thr  Leu Ala Asn Gly Glu  Ile Arg Lys Arg Pro  Leu Ile Glu
```

```
    1265                1270                1275

Thr Asn  Gly Glu Thr Gly Glu  Ile Val Trp Asp Lys  Gly Arg Asp
    1280                1285                1290

Phe Ala  Thr Val Arg Lys Val  Leu Ser Met Pro Gln  Val Asn Ile
    1295                1300                1305

Val Lys  Lys Thr Glu Val Gln  Thr Gly Gly Phe Ser  Lys Glu Ser
    1310                1315                1320

Ile Leu  Pro Lys Arg Asn Ser  Asp Lys Leu Ile Ala  Arg Lys Lys
    1325                1330                1335

Asp Trp  Asp Pro Lys Lys Tyr  Gly Gly Phe Asp Ser  Pro Thr Val
    1340                1345                1350

Ala Tyr  Ser Val Leu Val Val  Ala Lys Val Glu Lys  Gly Lys Ser
    1355                1360                1365

Lys Lys  Leu Lys Ser Val Lys  Glu Leu Leu Gly Ile  Thr Ile Met
    1370                1375                1380

Glu Arg  Ser Ser Phe Glu Lys  Asn Pro Ile Asp Phe  Leu Glu Ala
    1385                1390                1395

Lys Gly  Tyr Lys Glu Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro
    1400                1405                1410

Lys Tyr  Ser Leu Phe Glu Leu  Glu Asn Gly Arg Lys  Arg Met Leu
    1415                1420                1425

Ala Ser  Ala Gly Glu Leu Gln  Lys Gly Asn Glu Leu  Ala Leu Pro
    1430                1435                1440

Ser Lys  Tyr Val Asn Phe Leu  Tyr Leu Ala Ser His  Tyr Glu Lys
    1445                1450                1455
```

```
Leu Lys  Gly Ser Pro Glu Asp  Asn Glu Gln Lys Gln  Leu Phe Val
    1460                1465                1470


Glu Gln  His Lys His Tyr Leu  Asp Glu Ile Ile Glu  Gln Ile Ser
    1475                1480                1485


Glu Phe  Ser Lys Arg Val Ile  Leu Ala Asp Ala Asn  Leu Asp Lys
    1490                1495                1500


Val Leu  Ser Ala Tyr Asn Lys  His Arg Asp Lys Pro  Ile Arg Glu
    1505                1510                1515


Gln Ala  Glu Asn Ile Ile His  Leu Phe Thr Leu Thr  Asn Leu Gly
    1520                1525                1530


Ala Pro  Ala Ala Phe Lys Tyr  Phe Asp Thr Thr Ile  Asp Arg Lys
    1535                1540                1545


Arg Tyr  Thr Ser Thr Lys Glu  Val Leu Asp Ala Thr  Leu Ile His
    1550                1555                1560


Gln Ser  Ile Thr Gly Leu Tyr  Glu Thr Arg Ile Asp  Leu Ser Gln
    1565                1570                1575


Leu Gly  Gly Asp Gly Gly Gly  Ser Pro Lys Lys Lys  Arg Lys Val
    1580                1585                1590


<210>  326
<211>  1592
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  326

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                  10                  15
```

```
Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Ala Ala Ala Lys Glu Ala Ala Ala
        195                 200                 205

Lys Glu Ala Ala Ala Lys Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile
```

```
    210                 215                 220

Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val
225                 230                 235                 240

Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile
                245                 250                 255

Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala
            260                 265                 270

Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg
        275                 280                 285

Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala
    290                 295                 300

Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val
305                 310                 315                 320

Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val
                325                 330                 335

Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg
            340                 345                 350

Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr
        355                 360                 365

Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu
    370                 375                 380

Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln
385                 390                 395                 400

Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala
                405                 410                 415
```

```
Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser
            420                 425                 430

Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn
        435                 440                 445

Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn
    450                 455                 460

Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser
465                 470                 475                 480

Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly
                485                 490                 495

Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala
            500                 505                 510

Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala
        515                 520                 525

Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp
    530                 535                 540

Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr
545                 550                 555                 560

Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile
                565                 570                 575

Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile
            580                 585                 590

Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg
        595                 600                 605

Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro
```

```
    610                 615                 620

His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu
625                 630                 635                 640

Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile
                645                 650                 655

Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn
            660                 665                 670

Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro
        675                 680                 685

Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe
    690                 695                 700

Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val
705                 710                 715                 720

Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu
                725                 730                 735

Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe
            740                 745                 750

Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr
        755                 760                 765

Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys
    770                 775                 780

Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe
785                 790                 795                 800

Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp
                805                 810                 815
```

```
Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile
            820                 825                 830

Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg
        835                 840                 845

Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu
    850                 855                 860

Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile
865                 870                 875                 880

Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu
                885                 890                 895

Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp
            900                 905                 910

Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly
        915                 920                 925

Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro
    930                 935                 940

Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu
945                 950                 955                 960

Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met
                965                 970                 975

Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu
            980                 985                 990

Arg Met Lys Arg Ile Glu Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile
        995                 1000                1005

Leu Lys  Glu His Pro Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys
```

```
    1010                1015                1020

Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp
    1025                1030                1035

Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His
    1040                1045                1050

Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys
    1055                1060                1065

Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val
    1070                1075                1080

Pro Ser  Glu Glu Val Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln
    1085                1090                1095

Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu
    1100                1105                1110

Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly
    1115                1120                1125

Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg Gln Ile  Thr Lys His
    1130                1135                1140

Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu
    1145                1150                1155

Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser
    1160                1165                1170

Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val
    1175                1180                1185

Arg Glu  Ile Asn Asn Tyr His  His Ala His Asp Ala  Tyr Leu Asn
    1190                1195                1200
```

```
Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu
    1205                1210                1215

Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys
    1220                1225                1230

Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys
    1235                1240                1245

Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile
    1250                1255                1260

Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr
    1265                1270                1275

Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe
    1280                1285                1290

Ala Thr  Val Arg Lys Val Leu  Ser Met Pro Gln Val  Asn Ile Val
    1295                1300                1305

Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile
    1310                1315                1320

Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp
    1325                1330                1335

Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala
    1340                1345                1350

Tyr Ser  Val Leu Val Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys
    1355                1360                1365

Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu
    1370                1375                1380

Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys
```

```
    1385                1390                1395

Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys
    1400                1405                1410

Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala
    1415                1420                1425

Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser
    1430                1435                1440

Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu
    1445                1450                1455

Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu
    1460                1465                1470

Gln His  Lys His Tyr Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu
    1475                1480                1485

Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val
    1490                1495                1500

Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln
    1505                1510                1515

Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala
    1520                1525                1530

Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg
    1535                1540                1545

Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala Thr Leu  Ile His Gln
    1550                1555                1560

Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu
    1565                1570                1575
```

```
Gly Gly  Asp Gly Gly Gly Ser  Pro Lys Lys Lys Arg  Lys Val
    1580                1585                1590
```

<210> 327
<211> 1679
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 327

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140
```

```
Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Gly Gly Gly Ser Glu Ala Ala Ala
        195                 200                 205

Lys Glu Ala Ala Ala Lys Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile
    210                 215                 220

Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val
225                 230                 235                 240

Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile
                245                 250                 255

Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala
            260                 265                 270

Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg
        275                 280                 285

Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala
    290                 295                 300

Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val
305                 310                 315                 320

Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val
                325                 330                 335

Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg
```

```
            340                 345                 350

Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr
        355                 360                 365

Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu
    370                 375                 380

Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln
385                 390                 395                 400

Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala
                405                 410                 415

Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser
            420                 425                 430

Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn
        435                 440                 445

Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn
    450                 455                 460

Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser
465                 470                 475                 480

Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly
                485                 490                 495

Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala
            500                 505                 510

Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala
        515                 520                 525

Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp
    530                 535                 540
```

```
Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr
545                 550                 555                 560

Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile
                565                 570                 575

Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile
            580                 585                 590

Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg
        595                 600                 605

Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro
    610                 615                 620

His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu
625                 630                 635                 640

Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile
                645                 650                 655

Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn
            660                 665                 670

Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro
        675                 680                 685

Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe
    690                 695                 700

Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val
705                 710                 715                 720

Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu
                725                 730                 735

Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe
```

```
            740                 745                 750

Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr
        755                 760                 765

Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys
    770                 775                 780

Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe
785                 790                 795                 800

Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp
                805                 810                 815

Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile
            820                 825                 830

Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg
        835                 840                 845

Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu
    850                 855                 860

Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile
865                 870                 875                 880

Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu
                885                 890                 895

Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp
            900                 905                 910

Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly
        915                 920                 925

Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro
    930                 935                 940
```

```
Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu
945                 950                 955                 960


Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met
                965                 970                 975


Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu
            980                 985                 990


Arg Met Lys Arg Ile Glu Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile
        995                 1000                1005


Leu Lys  Glu His Pro Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys
    1010                1015                1020


Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp
    1025                1030                1035


Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His
    1040                1045                1050


Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys
    1055                1060                1065


Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val
    1070                1075                1080


Pro Ser  Glu Glu Val Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln
    1085                1090                1095


Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu
    1100                1105                1110


Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly
    1115                1120                1125


Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg Gln Ile  Thr Lys His
```

```
    1130                1135                1140

Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu
    1145                1150                1155

Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser
    1160                1165                1170

Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val
    1175                1180                1185

Arg Glu  Ile Asn Asn Tyr His  His Ala His Asp Ala  Tyr Leu Asn
    1190                1195                1200

Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu
    1205                1210                1215

Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys
    1220                1225                1230

Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys
    1235                1240                1245

Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile
    1250                1255                1260

Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr
    1265                1270                1275

Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe
    1280                1285                1290

Ala Thr  Val Arg Lys Val Leu  Ser Met Pro Gln Val  Asn Ile Val
    1295                1300                1305

Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile
    1310                1315                1320
```

```
Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp
    1325                1330                1335

Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala
    1340                1345                1350

Tyr Ser  Val Leu Val Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys
    1355                1360                1365

Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu
    1370                1375                1380

Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys
    1385                1390                1395

Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys
    1400                1405                1410

Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala
    1415                1420                1425

Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser
    1430                1435                1440

Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu
    1445                1450                1455

Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu
    1460                1465                1470

Gln His  Lys His Tyr Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu
    1475                1480                1485

Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val
    1490                1495                1500

Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln
```

```
    1505                1510                1515

Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala
    1520                1525                1530

Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg
    1535                1540                1545

Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala Thr Leu  Ile His Gln
    1550                1555                1560

Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu
    1565                1570                1575

Gly Gly  Asp Ser Gly Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu
    1580                1585                1590

Lys Glu  Thr Gly Lys Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met
    1595                1600                1605

Leu Pro  Glu Glu Val Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser
    1610                1615                1620

Asp Ile  Leu Val His Thr Ala  Tyr Asp Glu Ser Thr  Asp Glu Asn
    1625                1630                1635

Val Met  Leu Leu Thr Ser Asp  Ala Pro Glu Tyr Lys  Pro Trp Ala
    1640                1645                1650

Leu Val  Ile Gln Asp Ser Asn  Gly Glu Asn Lys Ile  Lys Met Leu
    1655                1660                1665

Ser Gly  Gly Ser Pro Lys Lys  Lys Arg Lys Val
    1670                1675
```

<210> 328
<211> 1679
<212> PRT

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 328

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15


Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30


Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45


Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60


Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80


Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95


Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110


Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125


Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140


Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160


Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175
```

```
Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Ala Ala Ala Lys Gly Gly Gly Gly
        195                 200                 205

Ser Gly Gly Gly Gly Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile
    210                 215                 220

Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val
225                 230                 235                 240

Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile
                245                 250                 255

Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala
            260                 265                 270

Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg
        275                 280                 285

Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala
    290                 295                 300

Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val
305                 310                 315                 320

Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val
                325                 330                 335

Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg
            340                 345                 350

Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr
        355                 360                 365

Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu
```

```
    370                 375                 380

Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln
385                 390                 395                 400

Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala
                405                 410                 415

Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser
            420                 425                 430

Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn
        435                 440                 445

Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn
    450                 455                 460

Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser
465                 470                 475                 480

Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly
                485                 490                 495

Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala
            500                 505                 510

Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala
        515                 520                 525

Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp
    530                 535                 540

Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr
545                 550                 555                 560

Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile
                565                 570                 575
```

```
Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile
            580                 585                 590

Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg
        595                 600                 605

Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro
    610                 615                 620

His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu
625                 630                 635                 640

Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile
                645                 650                 655

Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn
            660                 665                 670

Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro
        675                 680                 685

Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe
    690                 695                 700

Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val
705                 710                 715                 720

Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu
                725                 730                 735

Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe
            740                 745                 750

Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr
        755                 760                 765

Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys
```

```
    770                 775                 780

Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe
785                 790                 795                 800

Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp
                805                 810                 815

Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile
            820                 825                 830

Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg
        835                 840                 845

Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu
    850                 855                 860

Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile
865                 870                 875                 880

Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu
                885                 890                 895

Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp
            900                 905                 910

Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly
        915                 920                 925

Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro
    930                 935                 940

Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu
945                 950                 955                 960

Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met
                965                 970                 975
```

```
Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu
            980                 985                 990


Arg Met Lys Arg Ile Glu Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile
        995                 1000                1005


Leu Lys  Glu His Pro Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys
    1010                1015                1020


Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp
    1025                1030                1035


Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His
    1040                1045                1050


Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys
    1055                1060                1065


Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val
    1070                1075                1080


Pro Ser  Glu Glu Val Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln
    1085                1090                1095


Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu
    1100                1105                1110


Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly
    1115                1120                1125


Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg Gln Ile  Thr Lys His
    1130                1135                1140


Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu
    1145                1150                1155


Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser
```

```
    1160                1165                1170

Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val
    1175                1180                1185

Arg Glu  Ile Asn Asn Tyr His  His Ala His Asp Ala  Tyr Leu Asn
    1190                1195                1200

Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu
    1205                1210                1215

Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys
    1220                1225                1230

Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys
    1235                1240                1245

Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile
    1250                1255                1260

Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr
    1265                1270                1275

Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe
    1280                1285                1290

Ala Thr  Val Arg Lys Val Leu  Ser Met Pro Gln Val  Asn Ile Val
    1295                1300                1305

Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile
    1310                1315                1320

Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp
    1325                1330                1335

Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala
    1340                1345                1350
```

```
Tyr Ser  Val Leu Val Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys
    1355                1360                1365

Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu
    1370                1375                1380

Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys
    1385                1390                1395

Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys
    1400                1405                1410

Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala
    1415                1420                1425

Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser
    1430                1435                1440

Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu
    1445                1450                1455

Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu
    1460                1465                1470

Gln His  Lys His Tyr Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu
    1475                1480                1485

Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val
    1490                1495                1500

Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln
    1505                1510                1515

Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala
    1520                1525                1530

Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg
```

```
    1535                1540                1545

Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala Thr Leu  Ile His Gln
    1550                1555                1560

Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu
    1565                1570                1575

Gly Gly  Asp Ser Gly Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu
    1580                1585                1590

Lys Glu  Thr Gly Lys Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met
    1595                1600                1605

Leu Pro  Glu Glu Val Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser
    1610                1615                1620

Asp Ile  Leu Val His Thr Ala  Tyr Asp Glu Ser Thr  Asp Glu Asn
    1625                1630                1635

Val Met  Leu Leu Thr Ser Asp  Ala Pro Glu Tyr Lys  Pro Trp Ala
    1640                1645                1650

Leu Val  Ile Gln Asp Ser Asn  Gly Glu Asn Lys Ile  Lys Met Leu
    1655                1660                1665

Ser Gly  Gly Ser Pro Lys Lys  Lys Arg Lys Val
    1670                1675


<210>  329
<211>  1679
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  329

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
```

```
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Ala Ala Ala Lys Glu Ala Ala Ala
        195                 200                 205
```

```
Lys Glu Ala Ala Ala Lys Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile
    210                 215                 220

Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val
225                 230                 235                 240

Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile
                245                 250                 255

Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala
            260                 265                 270

Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg
        275                 280                 285

Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala
    290                 295                 300

Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val
305                 310                 315                 320

Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val
                325                 330                 335

Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg
            340                 345                 350

Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr
        355                 360                 365

Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu
    370                 375                 380

Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln
385                 390                 395                 400

Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala
```

```
                405                 410                 415

Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser
            420                 425                 430

Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn
        435                 440                 445

Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn
    450                 455                 460

Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser
465                 470                 475                 480

Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly
                485                 490                 495

Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala
            500                 505                 510

Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala
        515                 520                 525

Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp
    530                 535                 540

Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr
545                 550                 555                 560

Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile
                565                 570                 575

Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile
            580                 585                 590

Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg
        595                 600                 605
```

```
Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro
    610                 615                 620

His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu
625                 630                 635                 640

Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile
                645                 650                 655

Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn
            660                 665                 670

Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro
        675                 680                 685

Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe
    690                 695                 700

Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val
705                 710                 715                 720

Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu
                725                 730                 735

Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe
            740                 745                 750

Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr
        755                 760                 765

Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys
    770                 775                 780

Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe
785                 790                 795                 800

Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp
```

```
                805                 810                 815

Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile
            820                 825                 830

Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg
        835                 840                 845

Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu
    850                 855                 860

Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile
865                 870                 875                 880

Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu
                885                 890                 895

Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp
            900                 905                 910

Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly
        915                 920                 925

Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro
    930                 935                 940

Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu
945                 950                 955                 960

Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met
                965                 970                 975

Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu
            980                 985                 990

Arg Met Lys Arg Ile Glu Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile
        995                 1000                1005
```

```
Leu Lys  Glu His Pro Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys
    1010                1015                1020

Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp
    1025                1030                1035

Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His
    1040                1045                1050

Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys
    1055                1060                1065

Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val
    1070                1075                1080

Pro Ser  Glu Glu Val Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln
    1085                1090                1095

Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu
    1100                1105                1110

Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly
    1115                1120                1125

Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg Gln Ile  Thr Lys His
    1130                1135                1140

Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu
    1145                1150                1155

Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser
    1160                1165                1170

Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val
    1175                1180                1185

Arg Glu  Ile Asn Asn Tyr His  His Ala His Asp Ala  Tyr Leu Asn
```

```
    1190                1195                1200

Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu
    1205                1210                1215

Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys
    1220                1225                1230

Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys
    1235                1240                1245

Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile
    1250                1255                1260

Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr
    1265                1270                1275

Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe
    1280                1285                1290

Ala Thr  Val Arg Lys Val Leu  Ser Met Pro Gln Val  Asn Ile Val
    1295                1300                1305

Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile
    1310                1315                1320

Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp
    1325                1330                1335

Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala
    1340                1345                1350

Tyr Ser  Val Leu Val Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys
    1355                1360                1365

Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu
    1370                1375                1380
```

```
Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys
    1385                1390                1395

Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys
    1400                1405                1410

Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala
    1415                1420                1425

Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser
    1430                1435                1440

Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu
    1445                1450                1455

Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu
    1460                1465                1470

Gln His  Lys His Tyr Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu
    1475                1480                1485

Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val
    1490                1495                1500

Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln
    1505                1510                1515

Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala
    1520                1525                1530

Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg
    1535                1540                1545

Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala Thr Leu  Ile His Gln
    1550                1555                1560

Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu
```

```
    1565                1570                1575

Gly Gly  Asp Ser Gly Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu
    1580                1585                1590

Lys Glu  Thr Gly Lys Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met
    1595                1600                1605

Leu Pro  Glu Glu Val Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser
    1610                1615                1620

Asp Ile  Leu Val His Thr Ala  Tyr Asp Glu Ser Thr  Asp Glu Asn
    1625                1630                1635

Val Met  Leu Leu Thr Ser Asp  Ala Pro Glu Tyr Lys  Pro Trp Ala
    1640                1645                1650

Leu Val  Ile Gln Asp Ser Asn  Gly Glu Asn Lys Ile  Lys Met Leu
    1655                1660                1665

Ser Gly  Gly Ser Pro Lys Lys  Lys Arg Lys Val
    1670                1675
```

<210> 330
<211> 1684
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 330

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
```

```
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Gly Gly Gly Ser Gly Gly Gly Gly
        195                 200                 205

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Lys Lys Tyr Ser
    210                 215                 220

Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr
225                 230                 235                 240
```

```
Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr
                245                 250                 255

Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp
            260                 265                 270

Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg
        275                 280                 285

Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe
    290                 295                 300

Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu
305                 310                 315                 320

Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile
                325                 330                 335

Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr
            340                 345                 350

Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp
        355                 360                 365

Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly
    370                 375                 380

His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp
385                 390                 395                 400

Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu
                405                 410                 415

Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala
            420                 425                 430

Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro
```

```
        435                 440                 445

Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu
    450                 455                 460

Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala
465                 470                 475                 480

Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu
                485                 490                 495

Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys
            500                 505                 510

Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr
        515                 520                 525

Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp
    530                 535                 540

Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln
545                 550                 555                 560

Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly
                565                 570                 575

Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys
            580                 585                 590

Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu
        595                 600                 605

Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp
    610                 615                 620

Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile
625                 630                 635                 640
```

```
Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu
               645                 650                 655

Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro
            660                 665                 670

Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu
        675                 680                 685

Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala
    690                 695                 700

Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu
705                 710                 715                 720

Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe
               725                 730                 735

Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met
            740                 745                 750

Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp
        755                 760                 765

Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu
    770                 775                 780

Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly
785                 790                 795                 800

Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu
               805                 810                 815

Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp
            820                 825                 830

Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu
```

```
        835                 840                 845

Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys
    850                 855                 860

Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu
865                 870                 875                 880

Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr
                885                 890                 895

Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met
            900                 905                 910

Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys
        915                 920                 925

Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn
    930                 935                 940

Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys
945                 950                 955                 960

Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn
                965                 970                 975

Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln
            980                 985                 990

Lys Asn Ser Arg Glu Arg Met Lys  Arg Ile Glu Glu Gly  Ile Lys Glu
        995                 1000                1005

Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro Val Glu  Asn Thr Gln
    1010                1015                1020

Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg
    1025                1030                1035
```

```
Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp
    1040                1045                1050

Tyr Asp  Val Asp His Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp
    1055                1060                1065

Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly
    1070                1075                1080

Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val Val Lys  Lys Met Lys
    1085                1090                1095

Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg
    1100                1105                1110

Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu
    1115                1120                1125

Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg
    1130                1135                1140

Gln Ile  Thr Lys His Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn
    1145                1150                1155

Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val
    1160                1165                1170

Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe
    1175                1180                1185

Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn Tyr His  His Ala His
    1190                1195                1200

Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys
    1205                1210                1215

Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val
```

```
    1220                1225                1230

Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly
    1235                1240                1245

Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe
    1250                1255                1260

Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg
    1265                1270                1275

Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp
    1280                1285                1290

Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys Val Leu  Ser Met Pro
    1295                1300                1305

Gln Val  Asn Ile Val Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe
    1310                1315                1320

Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile
    1325                1330                1335

Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp
    1340                1345                1350

Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val Val Ala  Lys Val Glu
    1355                1360                1365

Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly
    1370                1375                1380

Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp
    1385                1390                1395

Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile
    1400                1405                1410
```

```
Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg
    1415                1420                1425

Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu
    1430                1435                1440

Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser
    1445                1450                1455

His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys
    1460                1465                1470

Gln Leu  Phe Val Glu Gln His  Lys His Tyr Leu Asp  Glu Ile Ile
    1475                1480                1485

Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala
    1490                1495                1500

Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys
    1505                1510                1515

Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu
    1520                1525                1530

Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr
    1535                1540                1545

Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala
    1550                1555                1560

Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile
    1565                1570                1575

Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly Gly Ser  Thr Asn Leu
    1580                1585                1590

Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys Gln Leu  Val Ile Gln
```

```
    1595                1600                1605

Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val Glu Glu  Val Ile Gly
    1610                1615                1620

Asn Lys  Pro Glu Ser Asp Ile  Leu Val His Thr Ala  Tyr Asp Glu
    1625                1630                1635

Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr Ser Asp  Ala Pro Glu
    1640                1645                1650

Tyr Lys  Pro Trp Ala Leu Val  Ile Gln Asp Ser Asn  Gly Glu Asn
    1655                1660                1665

Lys Ile  Lys Met Leu Ser Gly  Gly Ser Pro Lys Lys  Lys Arg Lys
    1670                1675                1680

Val
```

<210> 331
<211> 1684
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 331

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
```

```
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Gly Gly Gly Ser Gly Gly Gly Gly
        195                 200                 205

Ser Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Asp Lys Lys Tyr Ser
    210                 215                 220

Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr
225                 230                 235                 240

Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr
                245                 250                 255
```

```
Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp
            260                 265                 270

Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg
        275                 280                 285

Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe
    290                 295                 300

Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu
305                 310                 315                 320

Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile
                325                 330                 335

Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr
            340                 345                 350

Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp
        355                 360                 365

Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly
    370                 375                 380

His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp
385                 390                 395                 400

Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu
                405                 410                 415

Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala
            420                 425                 430

Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro
        435                 440                 445

Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu
```

```
    450                 455                 460

Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala
465                 470                 475                 480

Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu
                485                 490                 495

Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys
            500                 505                 510

Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr
        515                 520                 525

Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp
    530                 535                 540

Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln
545                 550                 555                 560

Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly
                565                 570                 575

Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys
            580                 585                 590

Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu
        595                 600                 605

Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp
    610                 615                 620

Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile
625                 630                 635                 640

Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu
                645                 650                 655
```

```
Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro
            660                 665                 670

Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu
        675                 680                 685

Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala
    690                 695                 700

Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu
705                 710                 715                 720

Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe
                725                 730                 735

Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met
            740                 745                 750

Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp
        755                 760                 765

Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu
    770                 775                 780

Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly
785                 790                 795                 800

Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu
                805                 810                 815

Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp
            820                 825                 830

Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu
        835                 840                 845

Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys
```

```
    850                 855                 860

Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu
865                 870                 875                 880

Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr
                885                 890                 895

Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met
            900                 905                 910

Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys
        915                 920                 925

Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn
    930                 935                 940

Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys
945                 950                 955                 960

Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn
                965                 970                 975

Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln
            980                 985                 990

Lys Asn Ser Arg Glu Arg Met Lys  Arg Ile Glu Glu Gly  Ile Lys Glu
        995                 1000                1005

Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro Val Glu  Asn Thr Gln
    1010                1015                1020

Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg
    1025                1030                1035

Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp
    1040                1045                1050
```

```
Tyr Asp  Val Asp His Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp
    1055                1060                1065

Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly
    1070                1075                1080

Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val Val Lys  Lys Met Lys
    1085                1090                1095

Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg
    1100                1105                1110

Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu
    1115                1120                1125

Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg
    1130                1135                1140

Gln Ile  Thr Lys His Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn
    1145                1150                1155

Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val
    1160                1165                1170

Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe
    1175                1180                1185

Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn Tyr His  His Ala His
    1190                1195                1200

Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys
    1205                1210                1215

Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val
    1220                1225                1230

Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly
```

```
    1235                1240                1245

Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe
    1250                1255                1260

Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg
    1265                1270                1275

Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp
    1280                1285                1290

Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys Val Leu  Ser Met Pro
    1295                1300                1305

Gln Val  Asn Ile Val Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe
    1310                1315                1320

Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile
    1325                1330                1335

Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp
    1340                1345                1350

Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val Val Ala  Lys Val Glu
    1355                1360                1365

Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly
    1370                1375                1380

Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp
    1385                1390                1395

Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile
    1400                1405                1410

Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg
    1415                1420                1425
```

```
Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu
    1430                1435                1440

Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser
    1445                1450                1455

His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys
    1460                1465                1470

Gln Leu  Phe Val Glu Gln His  Lys His Tyr Leu Asp  Glu Ile Ile
    1475                1480                1485

Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala
    1490                1495                1500

Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys
    1505                1510                1515

Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu
    1520                1525                1530

Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr
    1535                1540                1545

Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala
    1550                1555                1560

Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile
    1565                1570                1575

Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly Gly Ser  Thr Asn Leu
    1580                1585                1590

Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys Gln Leu  Val Ile Gln
    1595                1600                1605

Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val Glu Glu  Val Ile Gly
```

```
    1610                1615                1620

Asn Lys  Pro Glu Ser Asp Ile  Leu Val His Thr Ala  Tyr Asp Glu
    1625                1630                1635

Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr Ser Asp  Ala Pro Glu
    1640                1645                1650

Tyr Lys  Pro Trp Ala Leu Val  Ile Gln Asp Ser Asn  Gly Glu Asn
    1655                1660                1665

Lys Ile  Lys Met Leu Ser Gly  Gly Ser Pro Lys Lys  Lys Arg Lys
    1670                1675                1680

Val
```

<210> 332
<211> 1684
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 332

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
```

```
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Gly Gly Gly Ser Glu Ala Ala Ala
        195                 200                 205

Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Lys Lys Tyr Ser
    210                 215                 220

Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr
225                 230                 235                 240

Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr
                245                 250                 255

Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp
            260                 265                 270
```

```
Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg
        275                 280                 285

Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe
    290                 295                 300

Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu
305                 310                 315                 320

Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile
                325                 330                 335

Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr
            340                 345                 350

Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp
        355                 360                 365

Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly
    370                 375                 380

His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp
385                 390                 395                 400

Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu
                405                 410                 415

Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala
            420                 425                 430

Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro
        435                 440                 445

Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu
    450                 455                 460

Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala
```

```
465                 470                 475                 480

Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu
                485                 490                 495

Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys
            500                 505                 510

Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr
        515                 520                 525

Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp
    530                 535                 540

Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln
545                 550                 555                 560

Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly
                565                 570                 575

Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys
            580                 585                 590

Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu
        595                 600                 605

Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp
    610                 615                 620

Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile
625                 630                 635                 640

Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu
                645                 650                 655

Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro
            660                 665                 670
```

```
Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu
        675                 680                 685

Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala
    690                 695                 700

Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu
705                 710                 715                 720

Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe
                725                 730                 735

Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met
            740                 745                 750

Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp
        755                 760                 765

Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu
    770                 775                 780

Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly
785                 790                 795                 800

Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu
                805                 810                 815

Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp
            820                 825                 830

Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu
        835                 840                 845

Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys
    850                 855                 860

Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu
```

```
865                 870                 875                 880

Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr
                885                 890                 895

Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met
            900                 905                 910

Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys
        915                 920                 925

Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn
    930                 935                 940

Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys
945                 950                 955                 960

Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn
                965                 970                 975

Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln
            980                 985                 990

Lys Asn Ser Arg Glu Arg Met Lys  Arg Ile Glu Glu Gly  Ile Lys Glu
        995                 1000                1005

Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro Val Glu  Asn Thr Gln
    1010                1015                1020

Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg
    1025                1030                1035

Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp
    1040                1045                1050

Tyr Asp  Val Asp His Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp
    1055                1060                1065
```

```
Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly
    1070                1075                1080

Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val Val Lys  Lys Met Lys
    1085                1090                1095

Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg
    1100                1105                1110

Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu
    1115                1120                1125

Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg
    1130                1135                1140

Gln Ile  Thr Lys His Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn
    1145                1150                1155

Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val
    1160                1165                1170

Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe
    1175                1180                1185

Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn Tyr His  His Ala His
    1190                1195                1200

Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys
    1205                1210                1215

Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val
    1220                1225                1230

Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly
    1235                1240                1245

Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe
```

```
    1250                1255                1260

Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg
    1265                1270                1275

Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp
    1280                1285                1290

Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys Val Leu  Ser Met Pro
    1295                1300                1305

Gln Val  Asn Ile Val Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe
    1310                1315                1320

Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile
    1325                1330                1335

Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp
    1340                1345                1350

Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val Val Ala  Lys Val Glu
    1355                1360                1365

Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly
    1370                1375                1380

Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp
    1385                1390                1395

Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile
    1400                1405                1410

Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg
    1415                1420                1425

Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu
    1430                1435                1440
```

```
Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser
    1445                1450                1455

His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys
    1460                1465                1470

Gln Leu  Phe Val Glu Gln His  Lys His Tyr Leu Asp  Glu Ile Ile
    1475                1480                1485

Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala
    1490                1495                1500

Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys
    1505                1510                1515

Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu
    1520                1525                1530

Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr
    1535                1540                1545

Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala
    1550                1555                1560

Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile
    1565                1570                1575

Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly Gly Ser  Thr Asn Leu
    1580                1585                1590

Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys Gln Leu  Val Ile Gln
    1595                1600                1605

Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val Glu Glu  Val Ile Gly
    1610                1615                1620

Asn Lys  Pro Glu Ser Asp Ile  Leu Val His Thr Ala  Tyr Asp Glu
```

```
    1625                1630                1635

Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr Ser Asp  Ala Pro Glu
    1640                1645                1650

Tyr Lys  Pro Trp Ala Leu Val  Ile Gln Asp Ser Asn  Gly Glu Asn
    1655                1660                1665

Lys Ile  Lys Met Leu Ser Gly  Gly Ser Pro Lys Lys  Lys Arg Lys
    1670                1675                1680

Val
```

<210> 333
<211> 1684
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 333

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
```

```
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Ala Ala Ala Lys Glu Ala Ala Ala
        195                 200                 205

Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Asp Lys Lys Tyr Ser
    210                 215                 220

Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr
225                 230                 235                 240

Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr
                245                 250                 255

Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp
            260                 265                 270

Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg
        275                 280                 285
```

```
Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe
    290                 295                 300

Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu
305                 310                 315                 320

Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile
                325                 330                 335

Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr
            340                 345                 350

Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp
        355                 360                 365

Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly
    370                 375                 380

His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp
385                 390                 395                 400

Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu
                405                 410                 415

Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala
            420                 425                 430

Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro
        435                 440                 445

Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu
    450                 455                 460

Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala
465                 470                 475                 480

Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu
```

```
                485                 490                 495

Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys
            500                 505                 510

Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr
        515                 520                 525

Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp
    530                 535                 540

Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln
545                 550                 555                 560

Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly
                565                 570                 575

Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys
            580                 585                 590

Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu
        595                 600                 605

Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp
    610                 615                 620

Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile
625                 630                 635                 640

Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu
                645                 650                 655

Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro
            660                 665                 670

Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu
        675                 680                 685
```

```
Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala
    690                 695                 700

Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu
705                 710                 715                 720

Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe
                725                 730                 735

Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met
            740                 745                 750

Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp
        755                 760                 765

Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu
    770                 775                 780

Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly
785                 790                 795                 800

Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu
                805                 810                 815

Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp
            820                 825                 830

Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu
        835                 840                 845

Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys
    850                 855                 860

Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu
865                 870                 875                 880

Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr
```

```
               885                 890                 895

Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met
            900                 905                 910

Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys
        915                 920                 925

Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn
    930                 935                 940

Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys
945                 950                 955                 960

Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn
                965                 970                 975

Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln
            980                 985                 990

Lys Asn Ser Arg Glu Arg Met Lys  Arg Ile Glu Glu Gly  Ile Lys Glu
        995                 1000                1005

Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro Val Glu  Asn Thr Gln
    1010                1015                1020

Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr Leu Gln  Asn Gly Arg
    1025                1030                1035

Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile Asn Arg  Leu Ser Asp
    1040                1045                1050

Tyr Asp  Val Asp His Ile Val  Pro Gln Ser Phe Leu  Lys Asp Asp
    1055                1060                1065

Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser Asp Lys  Asn Arg Gly
    1070                1075                1080
```

```
Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val Val Lys  Lys Met Lys
    1085                1090                1095

Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys Leu Ile  Thr Gln Arg
    1100                1105                1110

Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg Gly Gly  Leu Ser Glu
    1115                1120                1125

Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln Leu Val  Glu Thr Arg
    1130                1135                1140

Gln Ile  Thr Lys His Val Ala  Gln Ile Leu Asp Ser  Arg Met Asn
    1145                1150                1155

Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile Arg Glu  Val Lys Val
    1160                1165                1170

Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp Phe Arg  Lys Asp Phe
    1175                1180                1185

Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn Tyr His  His Ala His
    1190                1195                1200

Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr Ala Leu  Ile Lys Lys
    1205                1210                1215

Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr Gly Asp  Tyr Lys Val
    1220                1225                1230

Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser Glu Gln  Glu Ile Gly
    1235                1240                1245

Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser Asn Ile  Met Asn Phe
    1250                1255                1260

Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly Glu Ile  Arg Lys Arg
```

```
    1265                1270                1275

Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly Glu Ile  Val Trp Asp
    1280                1285                1290

Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys Val Leu  Ser Met Pro
    1295                1300                1305

Gln Val  Asn Ile Val Lys Lys  Thr Glu Val Gln Thr  Gly Gly Phe
    1310                1315                1320

Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn Ser Asp  Lys Leu Ile
    1325                1330                1335

Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys Tyr Gly  Gly Phe Asp
    1340                1345                1350

Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val Val Ala  Lys Val Glu
    1355                1360                1365

Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val Lys Glu  Leu Leu Gly
    1370                1375                1380

Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu Lys Asn  Pro Ile Asp
    1385                1390                1395

Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val Lys Lys  Asp Leu Ile
    1400                1405                1410

Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu Leu Glu  Asn Gly Arg
    1415                1420                1425

Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu Gln Lys  Gly Asn Glu
    1430                1435                1440

Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe Leu Tyr  Leu Ala Ser
    1445                1450                1455
```

```
His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu Asp Asn  Glu Gln Lys
    1460                1465                1470

Gln Leu  Phe Val Glu Gln His  Lys His Tyr Leu Asp  Glu Ile Ile
    1475                1480                1485

Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val Ile Leu  Ala Asp Ala
    1490                1495                1500

Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn Lys His  Arg Asp Lys
    1505                1510                1515

Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile His Leu  Phe Thr Leu
    1520                1525                1530

Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys Tyr Phe  Asp Thr Thr
    1535                1540                1545

Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys Glu Val  Leu Asp Ala
    1550                1555                1560

Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu Tyr Glu  Thr Arg Ile
    1565                1570                1575

Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly Gly Ser  Thr Asn Leu
    1580                1585                1590

Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys Gln Leu  Val Ile Gln
    1595                1600                1605

Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val Glu Glu  Val Ile Gly
    1610                1615                1620

Asn Lys  Pro Glu Ser Asp Ile  Leu Val His Thr Ala  Tyr Asp Glu
    1625                1630                1635

Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr Ser Asp  Ala Pro Glu
```

1640 1645 1650

Tyr Lys Pro Trp Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn
1655 1660 1665

Lys Ile Lys Met Leu Ser Gly Gly Ser Pro Lys Lys Lys Arg Lys
1670 1675 1680

Val

<210> 334
<211> 1689
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 334

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
20 25 30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
35 40 45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
50 55 60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65 70 75 80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
85 90 95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala

```
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Gly Gly Gly Gly Ser Glu Ala Ala Ala
        195                 200                 205

Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys
    210                 215                 220

Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly
225                 230                 235                 240

Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys
                245                 250                 255

Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly
            260                 265                 270

Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys
        275                 280                 285

Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr
    290                 295                 300
```

```
Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe
305                 310                 315                 320

Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His
                325                 330                 335

Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His
            340                 345                 350

Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser
        355                 360                 365

Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met
    370                 375                 380

Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp
385                 390                 395                 400

Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn
                405                 410                 415

Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys
            420                 425                 430

Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu
        435                 440                 445

Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu
    450                 455                 460

Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp
465                 470                 475                 480

Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp
                485                 490                 495

Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu
```

```
            500                 505                 510

Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile
        515                 520                 525

Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
    530                 535                 540

Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala
545                 550                 555                 560

Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
                565                 570                 575

Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln
            580                 585                 590

Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly
        595                 600                 605

Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys
    610                 615                 620

Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly
625                 630                 635                 640

Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu
                645                 650                 655

Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro
            660                 665                 670

Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met
        675                 680                 685

Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val
    690                 695                 700
```

```
Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn
705                 710                 715                 720


Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu
                725                 730                 735


Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr
            740                 745                 750


Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys
        755                 760                 765


Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
    770                 775                 780


Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser
785                 790                 795                 800


Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr
                805                 810                 815


Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn
            820                 825                 830


Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu
        835                 840                 845


Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His
    850                 855                 860


Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr
865                 870                 875                 880


Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys
                885                 890                 895


Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala
```

```
            900                 905                 910

Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys
        915                 920                 925

Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His
    930                 935                 940

Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
945                 950                 955                 960

Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg
                965                 970                 975

His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr
            980                 985                 990

Thr Gln Lys Gly Gln Lys Asn Ser  Arg Glu Arg Met Lys  Arg Ile Glu
        995                 1000                1005

Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro
    1010                1015                1020

Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr
    1025                1030                1035

Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile
    1040                1045                1050

Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His Ile Val  Pro Gln Ser
    1055                1060                1065

Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser
    1070                1075                1080

Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val
    1085                1090                1095
```

```
Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys
    1100                1105                1110

Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg
    1115                1120                1125

Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln
    1130                1135                1140

Leu Val  Glu Thr Arg Gln Ile  Thr Lys His Val Ala  Gln Ile Leu
    1145                1150                1155

Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile
    1160                1165                1170

Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp
    1175                1180                1185

Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn
    1190                1195                1200

Tyr His  His Ala His Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr
    1205                1210                1215

Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr
    1220                1225                1230

Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser
    1235                1240                1245

Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser
    1250                1255                1260

Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly
    1265                1270                1275

Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly
```

```
    1280                1285                1290

Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys
    1295                1300                1305

Val Leu  Ser Met Pro Gln Val  Asn Ile Val Lys Lys  Thr Glu Val
    1310                1315                1320

Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn
    1325                1330                1335

Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys
    1340                1345                1350

Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val
    1355                1360                1365

Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val
    1370                1375                1380

Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu
    1385                1390                1395

Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val
    1400                1405                1410

Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu
    1415                1420                1425

Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu
    1430                1435                1440

Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe
    1445                1450                1455

Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu
    1460                1465                1470
```

```
Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu Gln His  Lys His Tyr
    1475                1480                1485

Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val
    1490                1495                1500

Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn
    1505                1510                1515

Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile
    1520                1525                1530

His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys
    1535                1540                1545

Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys
    1550                1555                1560

Glu Val  Leu Asp Ala Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu
    1565                1570                1575

Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly
    1580                1585                1590

Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys
    1595                1600                1605

Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val
    1610                1615                1620

Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser Asp Ile  Leu Val His
    1625                1630                1635

Thr Ala  Tyr Asp Glu Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr
    1640                1645                1650

Ser Asp  Ala Pro Glu Tyr Lys  Pro Trp Ala Leu Val  Ile Gln Asp
```

1655 1660 1665

Ser Asn Gly Glu Asn Lys Ile Lys Met Leu Ser Gly Gly Ser Pro
1670 1675 1680

Lys Lys Lys Arg Lys Val
1685

<210> 335
<211> 1689
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 335

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
20 25 30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
35 40 45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
50 55 60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65 70 75 80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
85 90 95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
100 105 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg

```
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Ala Ala Ala Lys Glu Ala Ala Ala
        195                 200                 205

Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
    210                 215                 220

Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly
225                 230                 235                 240

Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys
                245                 250                 255

Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly
            260                 265                 270

Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys
        275                 280                 285

Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr
    290                 295                 300

Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe
305                 310                 315                 320
```

```
Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His
               325                 330                 335

Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His
            340                 345                 350

Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser
        355                 360                 365

Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met
    370                 375                 380

Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp
385                 390                 395                 400

Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn
               405                 410                 415

Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys
            420                 425                 430

Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu
        435                 440                 445

Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu
    450                 455                 460

Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp
465                 470                 475                 480

Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp
               485                 490                 495

Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu
            500                 505                 510

Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile
```

```
        515                 520                 525

Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
    530                 535                 540

Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala
545                 550                 555                 560

Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
                565                 570                 575

Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln
            580                 585                 590

Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly
        595                 600                 605

Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys
    610                 615                 620

Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly
625                 630                 635                 640

Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu
                645                 650                 655

Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro
            660                 665                 670

Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met
        675                 680                 685

Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val
    690                 695                 700

Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn
705                 710                 715                 720
```

```
Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu
                725                 730                 735

Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr
            740                 745                 750

Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys
        755                 760                 765

Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
    770                 775                 780

Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser
785                 790                 795                 800

Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr
                805                 810                 815

Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn
            820                 825                 830

Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu
        835                 840                 845

Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His
    850                 855                 860

Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr
865                 870                 875                 880

Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys
                885                 890                 895

Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala
            900                 905                 910

Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys
```

```
        915                 920                 925

Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His
    930                 935                 940

Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
945                 950                 955                 960

Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg
                965                 970                 975

His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr
            980                 985                 990

Thr Gln Lys Gly Gln Lys Asn Ser  Arg Glu Arg Met Lys  Arg Ile Glu
        995                 1000                1005

Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro
    1010                1015                1020

Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr
    1025                1030                1035

Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile
    1040                1045                1050

Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His Ile Val  Pro Gln Ser
    1055                1060                1065

Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser
    1070                1075                1080

Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val
    1085                1090                1095

Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys
    1100                1105                1110
```

```
Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg
    1115                1120                1125

Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln
    1130                1135                1140

Leu Val  Glu Thr Arg Gln Ile  Thr Lys His Val Ala  Gln Ile Leu
    1145                1150                1155

Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile
    1160                1165                1170

Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp
    1175                1180                1185

Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn
    1190                1195                1200

Tyr His  His Ala His Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr
    1205                1210                1215

Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr
    1220                1225                1230

Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser
    1235                1240                1245

Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser
    1250                1255                1260

Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly
    1265                1270                1275

Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly
    1280                1285                1290

Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys
```

```
    1295                1300                1305

Val Leu  Ser Met Pro Gln Val  Asn Ile Val Lys Lys  Thr Glu Val
    1310                1315                1320

Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn
    1325                1330                1335

Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys
    1340                1345                1350

Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val
    1355                1360                1365

Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val
    1370                1375                1380

Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu
    1385                1390                1395

Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val
    1400                1405                1410

Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu
    1415                1420                1425

Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu
    1430                1435                1440

Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe
    1445                1450                1455

Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu
    1460                1465                1470

Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu Gln His  Lys His Tyr
    1475                1480                1485
```

```
Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val
    1490                1495                1500

Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn
    1505                1510                1515

Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile
    1520                1525                1530

His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys
    1535                1540                1545

Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys
    1550                1555                1560

Glu Val  Leu Asp Ala Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu
    1565                1570                1575

Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly
    1580                1585                1590

Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys
    1595                1600                1605

Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val
    1610                1615                1620

Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser Asp Ile  Leu Val His
    1625                1630                1635

Thr Ala  Tyr Asp Glu Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr
    1640                1645                1650

Ser Asp  Ala Pro Glu Tyr Lys  Pro Trp Ala Leu Val  Ile Gln Asp
    1655                1660                1665

Ser Asn  Gly Glu Asn Lys Ile  Lys Met Leu Ser Gly  Gly Ser Pro
```

1670 1675 1680

Lys Lys Lys Arg Lys Val
1685

<210> 336
<211> 1689
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 336

Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
20 25 30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
35 40 45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
50 55 60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65 70 75 80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
85 90 95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
100 105 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
115 120 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg

```
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Ala Ala Ala Lys Glu Ala Ala Ala
        195                 200                 205

Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Glu Ala Ala Ala Lys
    210                 215                 220

Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly
225                 230                 235                 240

Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys
                245                 250                 255

Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly
            260                 265                 270

Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys
        275                 280                 285

Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr
    290                 295                 300

Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe
305                 310                 315                 320

Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His
                325                 330                 335
```

```
Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His
            340                 345                 350

Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser
        355                 360                 365

Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met
    370                 375                 380

Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp
385                 390                 395                 400

Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn
                405                 410                 415

Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys
            420                 425                 430

Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu
        435                 440                 445

Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu
    450                 455                 460

Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp
465                 470                 475                 480

Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp
                485                 490                 495

Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu
            500                 505                 510

Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile
        515                 520                 525

Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
```

```
    530                 535                 540

Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala
545                 550                 555                 560

Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
                565                 570                 575

Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln
            580                 585                 590

Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly
        595                 600                 605

Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys
    610                 615                 620

Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly
625                 630                 635                 640

Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu
                645                 650                 655

Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro
            660                 665                 670

Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met
        675                 680                 685

Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val
    690                 695                 700

Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn
705                 710                 715                 720

Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu
                725                 730                 735
```

```
Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr
            740                 745                 750

Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys
        755                 760                 765

Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
    770                 775                 780

Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser
785                 790                 795                 800

Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr
                805                 810                 815

Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn
            820                 825                 830

Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu
        835                 840                 845

Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His
    850                 855                 860

Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr
865                 870                 875                 880

Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys
                885                 890                 895

Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala
            900                 905                 910

Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys
        915                 920                 925

Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His
```

```
    930                 935                 940

Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
945                 950                 955                 960

Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg
                965                 970                 975

His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr
            980                 985                 990

Thr Gln Lys Gly Gln Lys Asn Ser  Arg Glu Arg Met Lys  Arg Ile Glu
        995                 1000                1005

Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro
    1010                1015                1020

Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr
    1025                1030                1035

Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile
    1040                1045                1050

Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His Ile Val  Pro Gln Ser
    1055                1060                1065

Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser
    1070                1075                1080

Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val
    1085                1090                1095

Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys
    1100                1105                1110

Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg
    1115                1120                1125
```

```
Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln
    1130                1135                1140

Leu Val  Glu Thr Arg Gln Ile  Thr Lys His Val Ala  Gln Ile Leu
    1145                1150                1155

Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile
    1160                1165                1170

Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp
    1175                1180                1185

Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn
    1190                1195                1200

Tyr His  His Ala His Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr
    1205                1210                1215

Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr
    1220                1225                1230

Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser
    1235                1240                1245

Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser
    1250                1255                1260

Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly
    1265                1270                1275

Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly
    1280                1285                1290

Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys
    1295                1300                1305

Val Leu  Ser Met Pro Gln Val  Asn Ile Val Lys Lys  Thr Glu Val
```

```
    1310                1315                1320

Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn
    1325                1330                1335

Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys
    1340                1345                1350

Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val
    1355                1360                1365

Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val
    1370                1375                1380

Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu
    1385                1390                1395

Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val
    1400                1405                1410

Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu
    1415                1420                1425

Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu
    1430                1435                1440

Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe
    1445                1450                1455

Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu
    1460                1465                1470

Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu Gln His  Lys His Tyr
    1475                1480                1485

Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val
    1490                1495                1500
```

```
Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn
    1505                1510                1515

Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile
    1520                1525                1530

His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys
    1535                1540                1545

Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys
    1550                1555                1560

Glu Val  Leu Asp Ala Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu
    1565                1570                1575

Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly
    1580                1585                1590

Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys
    1595                1600                1605

Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val
    1610                1615                1620

Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser Asp Ile  Leu Val His
    1625                1630                1635

Thr Ala  Tyr Asp Glu Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr
    1640                1645                1650

Ser Asp  Ala Pro Glu Tyr Lys  Pro Trp Ala Leu Val  Ile Gln Asp
    1655                1660                1665

Ser Asn  Gly Glu Asn Lys Ile  Lys Met Leu Ser Gly  Gly Ser Pro
    1670                1675                1680

Lys Lys  Lys Arg Lys Val
```

1685

<210> 337
<211> 1689
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 337

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                  10                  15


Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30


Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45


Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60


Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80


Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95


Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110


Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125


Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140


Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
```

```
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn Glu Ala Ala Ala Lys Glu Ala Ala Ala
        195                 200                 205

Lys Glu Ala Ala Ala Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
    210                 215                 220

Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly
225                 230                 235                 240

Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys
                245                 250                 255

Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly
            260                 265                 270

Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys
        275                 280                 285

Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr
    290                 295                 300

Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe
305                 310                 315                 320

Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His
                325                 330                 335

Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His
            340                 345                 350
```

```
Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser
        355                 360                 365

Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met
    370                 375                 380

Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp
385                 390                 395                 400

Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn
                405                 410                 415

Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys
            420                 425                 430

Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu
        435                 440                 445

Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu
    450                 455                 460

Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp
465                 470                 475                 480

Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp
                485                 490                 495

Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu
            500                 505                 510

Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile
        515                 520                 525

Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
    530                 535                 540

Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala
```

```
545                 550                 555                 560

Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
                565                 570                 575

Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln
            580                 585                 590

Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly
        595                 600                 605

Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys
    610                 615                 620

Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly
625                 630                 635                 640

Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu
                645                 650                 655

Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro
            660                 665                 670

Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met
        675                 680                 685

Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val
    690                 695                 700

Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn
705                 710                 715                 720

Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu
                725                 730                 735

Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr
            740                 745                 750
```

```
Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys
        755                 760                 765

Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
    770                 775                 780

Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser
785                 790                 795                 800

Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr
                805                 810                 815

Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn
            820                 825                 830

Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu
        835                 840                 845

Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His
    850                 855                 860

Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr
865                 870                 875                 880

Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys
                885                 890                 895

Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala
            900                 905                 910

Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys
        915                 920                 925

Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His
    930                 935                 940

Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
```

```
945                 950                 955                 960

Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg
               965                 970                 975

His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr
           980                 985                 990

Thr Gln Lys Gly Gln Lys Asn Ser  Arg Glu Arg Met Lys  Arg Ile Glu
        995                 1000                1005

Glu Gly  Ile Lys Glu Leu Gly  Ser Gln Ile Leu Lys  Glu His Pro
    1010                1015                1020

Val Glu  Asn Thr Gln Leu Gln  Asn Glu Lys Leu Tyr  Leu Tyr Tyr
    1025                1030                1035

Leu Gln  Asn Gly Arg Asp Met  Tyr Val Asp Gln Glu  Leu Asp Ile
    1040                1045                1050

Asn Arg  Leu Ser Asp Tyr Asp  Val Asp His Ile Val  Pro Gln Ser
    1055                1060                1065

Phe Leu  Lys Asp Asp Ser Ile  Asp Asn Lys Val Leu  Thr Arg Ser
    1070                1075                1080

Asp Lys  Asn Arg Gly Lys Ser  Asp Asn Val Pro Ser  Glu Glu Val
    1085                1090                1095

Val Lys  Lys Met Lys Asn Tyr  Trp Arg Gln Leu Leu  Asn Ala Lys
    1100                1105                1110

Leu Ile  Thr Gln Arg Lys Phe  Asp Asn Leu Thr Lys  Ala Glu Arg
    1115                1120                1125

Gly Gly  Leu Ser Glu Leu Asp  Lys Ala Gly Phe Ile  Lys Arg Gln
    1130                1135                1140
```

```
Leu Val  Glu Thr Arg Gln Ile  Thr Lys His Val Ala  Gln Ile Leu
    1145                1150                1155

Asp Ser  Arg Met Asn Thr Lys  Tyr Asp Glu Asn Asp  Lys Leu Ile
    1160                1165                1170

Arg Glu  Val Lys Val Ile Thr  Leu Lys Ser Lys Leu  Val Ser Asp
    1175                1180                1185

Phe Arg  Lys Asp Phe Gln Phe  Tyr Lys Val Arg Glu  Ile Asn Asn
    1190                1195                1200

Tyr His  His Ala His Asp Ala  Tyr Leu Asn Ala Val  Val Gly Thr
    1205                1210                1215

Ala Leu  Ile Lys Lys Tyr Pro  Lys Leu Glu Ser Glu  Phe Val Tyr
    1220                1225                1230

Gly Asp  Tyr Lys Val Tyr Asp  Val Arg Lys Met Ile  Ala Lys Ser
    1235                1240                1245

Glu Gln  Glu Ile Gly Lys Ala  Thr Ala Lys Tyr Phe  Phe Tyr Ser
    1250                1255                1260

Asn Ile  Met Asn Phe Phe Lys  Thr Glu Ile Thr Leu  Ala Asn Gly
    1265                1270                1275

Glu Ile  Arg Lys Arg Pro Leu  Ile Glu Thr Asn Gly  Glu Thr Gly
    1280                1285                1290

Glu Ile  Val Trp Asp Lys Gly  Arg Asp Phe Ala Thr  Val Arg Lys
    1295                1300                1305

Val Leu  Ser Met Pro Gln Val  Asn Ile Val Lys Lys  Thr Glu Val
    1310                1315                1320

Gln Thr  Gly Gly Phe Ser Lys  Glu Ser Ile Leu Pro  Lys Arg Asn
```

```
    1325                1330                1335

Ser Asp  Lys Leu Ile Ala Arg  Lys Lys Asp Trp Asp  Pro Lys Lys
    1340                1345                1350

Tyr Gly  Gly Phe Asp Ser Pro  Thr Val Ala Tyr Ser  Val Leu Val
    1355                1360                1365

Val Ala  Lys Val Glu Lys Gly  Lys Ser Lys Lys Leu  Lys Ser Val
    1370                1375                1380

Lys Glu  Leu Leu Gly Ile Thr  Ile Met Glu Arg Ser  Ser Phe Glu
    1385                1390                1395

Lys Asn  Pro Ile Asp Phe Leu  Glu Ala Lys Gly Tyr  Lys Glu Val
    1400                1405                1410

Lys Lys  Asp Leu Ile Ile Lys  Leu Pro Lys Tyr Ser  Leu Phe Glu
    1415                1420                1425

Leu Glu  Asn Gly Arg Lys Arg  Met Leu Ala Ser Ala  Gly Glu Leu
    1430                1435                1440

Gln Lys  Gly Asn Glu Leu Ala  Leu Pro Ser Lys Tyr  Val Asn Phe
    1445                1450                1455

Leu Tyr  Leu Ala Ser His Tyr  Glu Lys Leu Lys Gly  Ser Pro Glu
    1460                1465                1470

Asp Asn  Glu Gln Lys Gln Leu  Phe Val Glu Gln His  Lys His Tyr
    1475                1480                1485

Leu Asp  Glu Ile Ile Glu Gln  Ile Ser Glu Phe Ser  Lys Arg Val
    1490                1495                1500

Ile Leu  Ala Asp Ala Asn Leu  Asp Lys Val Leu Ser  Ala Tyr Asn
    1505                1510                1515
```

```
Lys His  Arg Asp Lys Pro Ile  Arg Glu Gln Ala Glu  Asn Ile Ile
    1520                1525                1530

His Leu  Phe Thr Leu Thr Asn  Leu Gly Ala Pro Ala  Ala Phe Lys
    1535                1540                1545

Tyr Phe  Asp Thr Thr Ile Asp  Arg Lys Arg Tyr Thr  Ser Thr Lys
    1550                1555                1560

Glu Val  Leu Asp Ala Thr Leu  Ile His Gln Ser Ile  Thr Gly Leu
    1565                1570                1575

Tyr Glu  Thr Arg Ile Asp Leu  Ser Gln Leu Gly Gly  Asp Ser Gly
    1580                1585                1590

Gly Ser  Thr Asn Leu Ser Asp  Ile Ile Glu Lys Glu  Thr Gly Lys
    1595                1600                1605

Gln Leu  Val Ile Gln Glu Ser  Ile Leu Met Leu Pro  Glu Glu Val
    1610                1615                1620

Glu Glu  Val Ile Gly Asn Lys  Pro Glu Ser Asp Ile  Leu Val His
    1625                1630                1635

Thr Ala  Tyr Asp Glu Ser Thr  Asp Glu Asn Val Met  Leu Leu Thr
    1640                1645                1650

Ser Asp  Ala Pro Glu Tyr Lys  Pro Trp Ala Leu Val  Ile Gln Asp
    1655                1660                1665

Ser Asn  Gly Glu Asn Lys Ile  Lys Met Leu Ser Gly  Gly Ser Pro
    1670                1675                1680

Lys Lys  Lys Arg Lys Val
    1685
```

<210> 338

<400> 338
000

<210> 339

<400> 339
000

<210> 340
<211> 3921
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 340
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggac ggcuccggcg 60

gcggcucccc caagaagaag cggaaggugg aggacaagcg gcccgccgcc accaagaagg 120

ccggccaggc caagaagaag aagggcggcu ccggcggcgg cgccgccuuc aagcccaacc 180

ccaucaacua cauccugggc cuggacaucg gcaucgccuc cgugggcugg gccauggugg 240

agaucgacga ggaggagaac cccauccggc ugaucgaccu gggcgugcgg guguucgagc 300

gggccgaggu gcccaagacc ggcgacuccc uggccauggc ccggcggcug gcccgguccg 360

ugcggcggcu gacccggcgg cgggcccacc ggcugcugcg ggcccggcgg cugcugaagc 420

gggagggcgu gcugcaggcc gccgacuucg acgagaacgg ccugaucaag ucccugccca 480

acaccccug gcagcugcgg gccgccgccc uggaccggaa gcugaccccc cuggaguggu 540

ccgccgugcu gcugcaccug aucaagcacc ggggcuaccu gucccagcgg aagaacgagg 600

gcgagaccgc cgacaaggag cugggcgccc ugcugaaggg cguggccaac aacgcccacg 660

cccugcagac cggcgacuuc cggacccccg ccgagcuggc ccugaacaag uucgagaagg 720

aguccggcca cauccggaac cagcggggcg acuacuccca caccuucucc cggaaggacc 780

ugcaggccga gcugauccug cuguucgaga agcagaagga guucggcaac ccccacgugu 840

ccggcggccu gaaggagggc aucgagaccc ugcugaugac ccagcggccc gcccuguccg 900

gcgacgccgu gcagaagaug cugggccacu gcaccuucga gcccgccgag cccaaggccg 960

ccaagaacac cuacaccgcc gagcgguuca ucuggcugac caagcugaac aaccugcgga 1020

uccuggagca gggcuccgag cggccccuga ccgacaccga gcgggccacc cugauggacg 1080

agcccuaccg gaaguccaag cugaccuacg cccaggcccg gaagcugcug ggccuggagg 1140

acaccgccuu cuucaagggc cugcgguacg gcaaggacaa cgccgaggcc uccacccuga 1200

uggagaugaa ggccuaccac gccaucuccc gggcccugga gaaggagggc cugaaggaca 1260

agaagucccc ccugaaccug uccuccgagc ugcaggacga gaucggcacc gccuucuccc 1320

uguucaagac cgacgaggac aucaccggcc ggcugaagga ccgggugcag cccgagaucc 1380

uggaggcccu gcugaagcac aucuccuucg acaaguucgu gcagaucucc cugaaggccc 1440

ugcggcggau cgugccccug auggagcagg gcaagcggua cgacgaggcc ugcgccgaga 1500

ucuacggcga ccacuacggc aagaagaaca ccgaggagaa gaucuaccug ccccccaucc 1560

ccgccgacga gauccggaac cccguggugc ugcgggcccu gucccaggcc cggaagguga 1620

ucaacggcgu ggugcggcgg uacggcuccc ccgcccggau ccacaucgag accgcccggg 1680

aggugggcaa guccuucaag gaccggaagg agaucgagaa gcggcaggag gagaaccgga 1740

aggaccggga gaaggccgcc gccaaguucc gggaguacuu ccccaacuuc gugggcgagc 1800

ccaaguccaa ggacauccug aagcugcggc uguacgagca gcagcacggc aagugccugu 1860

acuccggcaa ggagaucaac cuggugcggc ugaacgagaa gggcuacgug gagaucgacc 1920

acgcccugcc cuucucccgg accugggacg acuccuucaa caacaaggug cuggugcugg 1980

gcuccgagaa ccagaacaag ggcaaccaga cccccuacga guacuucaac ggcaaggaca 2040

acucccggga guggcaggag uucaaggccc ggguggagac cucccgguuc ccccggucca 2100

agaagcagcg gauccugcug cagaaguucg acgaggacgg cuucaaggag ugcaaccuga 2160

acgacacccg guacgugaac cgguuccugu gccaguucgu ggccgaccac auccugcuga 2220

ccggcaaggg caagcggcgg guguucgccu ccaacggcca gaucaccaac cugcugcggg 2280

gcuucugggg ccugcggaag gugcgggccg agaacgaccg gcaccacgcc cuggacgccg 2340

uggugguggc cugcuccacc guggccaugc agcagaagau cacccgguuc gugcgguaca 2400

aggagaugaa cgccuucgac ggcaagacca ucgacaagga gaccggcaag gugcugcacc 2460

agaagaccca cuucccccag cccugggagu ucuucgccca ggaggugaug auccgggugu 2520

ucggcaagcc cgacggcaag cccgaguucg aggaggccga cacccccgag aagcugcgga 2580

cccugcuggc cgagaagcug uccucccggc ccgaggccgu gcacgaguac gugacccccc 2640

uguucguguc ccgggccccc aaccggaaga uguccggcgc ccacaaggac acccugcggu 2700

ccgccaagcg guucgugaag cacaacgaga agaucuccgu gaagcgggug uggcugaccg 2760

agaucaagcu ggccgaccug gagaacaugg ugaacuacaa gaacggccgg gagaucgagc 2820

uguacgaggc ccugaaggcc cggcuggagg ccuacggcgg caacgccaag caggccuucg 2880

accccaagga caaccccuuc uacaagaagg gcggccagcu ggugaaggcc gugcgggugg 2940

agaagaccca ggaguccggc gugcugcuga acaagaagaa cgccuacacc aucgccgaca 3000

acggcgacau ggugcgggug gacguguucu gcaaggugga caagaagggc aagaaccagu 3060

acuucaucgu gcccaucuac gccuggcagg uggccgagaa cauccugccc gacaucgacu 3120

gcaagggcua ccggaucgac gacuccuaca ccuucugcuu cucccugcac aaguacgacc 3180

ugaucgccuu ccagaaggac gagaagucca agguggaguu cgccuacuac aucaacugcg 3240

acuccuccaa cggccgguuc uaccuggccu ggcacgacaa gggcuccaag gagcagcagu 3300

uccggaucuc cacccagaac cuggugcuga uccagaagua ccaggugaac gagcugggca 3360

aggagauccg gcccugccgg cugaagaagc ggccccccgu gcgguagcua gcaccagccu 3420

caagaacacc cgaauggagu cucuaagcua cauaauacca acuuacacuu uacaaaaugu 3480

uguccccaa aauguagcca uucguaucug cuccuaauaa aaagaaaguu ucuucacauu 3540

cucucgagaa aaaaaaaaaa uggaaaaaaa aaaaacggaa aaaaaaaaaa gguaaaaaaa 3600

aaaaauauaa aaaaaaaaaa cauaaaaaaa aaaaacgaaa aaaaaaaaac guaaaaaaaa 3660

aaaacucaaa aaaaaaaaga uaaaaaaaaa aaaccuaaaa aaaaaaaaug uaaaaaaaaa 3720

aaagggaaaa aaaaaaaacg caaaaaaaaa aaacacaaaa aaaaaaaaug caaaaaaaaa 3780

aaaucgaaaa aaaaaaaauc uaaaaaaaaa aaacgaaaaa aaaaaaaccc aaaaaaaaaa 3840

aagacaaaaa aaaaaaauag aaaaaaaaaa aguuaaaaaa aaaaaacuga aaaaaaaaaa 3900

auuuaaaaaa aaaaaaucua g 3921

<210> 341
<211> 5342
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 341
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cggcaccaag gacuccacca 660

aggacauccc cgagaccccc uccaaggacg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

uggugggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucgugugggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac ggcggcggcu 4800

cccccaagaa gaagcggaag gugugacuag caccagccuc aagaacaccc gaauggaguc 4860

ucuaagcuac auaauaccaa cuuacacuuu acaaaauguu gucccccaaa auguagccau 4920

ucguaucugc uccuaauaaa aagaaaguuu cuucacauuc ucucgagaaa aaaaaaaaau 4980

ggaaaaaaaa aaaacggaaa aaaaaaaaag guaaaaaaaa aaaauauaaa aaaaaaaaac 5040

auaaaaaaaa aaaacgaaaa aaaaaaaacg uaaaaaaaaa aaacucaaaa aaaaaaaaga 5100

uaaaaaaaaa aaaccuaaaa aaaaaaaaug uaaaaaaaaa aaagggaaaa aaaaaaaacg 5160

caaaaaaaaa aaacacaaaa aaaaaaaaug caaaaaaaaa aaaucgaaaa aaaaaaaauc 5220

uaaaaaaaaa aaacgaaaaa aaaaaaaccc aaaaaaaaaa aagacaaaaa aaaaaaauag 5280

aaaaaaaaaa aaguuaaaaa aaaaaaacug aaaaaaaaaa aauuuaaaaa aaaaaaaucu 5340

ag 5342

<210> 342
<211> 5342
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 342

gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cggccgggac gugcggcagc 660

ccgaggugaa ggaggagaag cccgaguccg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

uggugggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucguguggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac ggcggcggcu 4800

cccccaagaa gaagcggaag gugugacuag caccagccuc aagaacaccc gaauggaguc 4860

ucuaagcuac auaauaccaa cuuacacuuu acaaaauguu gucccccaaa auguagccau 4920

ucguaucugc uccuaauaaa aagaaaguuu cuucacauuc ucucgagaaa aaaaaaaaau 4980

ggaaaaaaaa aaaacggaaa aaaaaaaaag guaaaaaaaa aaaauauaaa aaaaaaaaac 5040

auaaaaaaaa aaaacgaaaa aaaaaaaacg uaaaaaaaaa aaacucaaaa aaaaaaaaga 5100

uaaaaaaaaa aaaccuaaaa aaaaaaaaug uaaaaaaaaa aaagggaaaa aaaaaaaacg 5160

caaaaaaaaa aaacacaaaa aaaaaaaaug caaaaaaaaa aaaucgaaaa aaaaaaaauc 5220

uaaaaaaaaa aaacgaaaaa aaaaaaaccc aaaaaaaaaa aagacaaaaa aaaaaaauag 5280

aaaaaaaaaa aaguuaaaaa aaaaaaacug aaaaaaaaaa aauuuaaaaa aaaaaaaucu 5340

ag 5342

<210> 343
<211> 5342
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 343
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cgagggcaag uccuccggcu 660

ccggcuccga guccaagucc accgccggcg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

ugguggggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucguguggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac ggcggcggcu 4800

```
cccccaagaa gaagcggaag gugugacuag caccagccuc aagaacaccc gaauggaguc     4860

ucuaagcuac auaauaccaa cuuacacuuu acaaaauguu gucccccaaa auguagccau     4920

ucguaucugc uccuaauaaa aagaaaguuu cuucacauuc ucucgagaaa aaaaaaaaau     4980

ggaaaaaaaa aaaacggaaa aaaaaaaaag guaaaaaaaa aaaauauaaa aaaaaaaaac     5040

auaaaaaaaa aaaacgaaaa aaaaaaaacg uaaaaaaaaa aaacucaaaa aaaaaaaaga     5100

uaaaaaaaaa aaaccuaaaa aaaaaaaaug uaaaaaaaaa aaagggaaaa aaaaaaaacg     5160

caaaaaaaaa aaacacaaaa aaaaaaaaug caaaaaaaaa aaaucgaaaa aaaaaaaauc     5220

uaaaaaaaaa aaacgaaaaa aaaaaaaccc aaaaaaaaaa aagacaaaaa aaaaaaauag     5280

aaaaaaaaaa aaguuaaaaa aaaaaaacug aaaaaaaaaa aauuuaaaaa aaaaaaaucu     5340

ag                                                                    5342
```

<210> 344
<211> 5342
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 344

```
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg       60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca      120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg      180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu      240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc      300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu      360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg      420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg      480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg      540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc      600
```

uguccggccg gcugcgggcc auccugcaga accagggcaa cacccccggc ucccccgccg 660

gcucccccac cuccaccgag gagggcaccg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

ugguggggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucguguggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac ggcggcggcu 4800

cccccaagaa gaagcggaag gugugacuag caccagccuc aagaacaccc gaauggaguc 4860

ucuaagcuac auaauaccaa cuuacacuuu acaaaauguu gucccccaaa auguagccau 4920

ucguaucugc uccuaauaaa aagaaaguuu cuucacauuc ucucgagaaa aaaaaaaaau 4980

ggaaaaaaaa aaaacggaaa aaaaaaaaag guaaaaaaaa aaaauauaaa aaaaaaaaac 5040

auaaaaaaaa aaaacgaaaa aaaaaaaacg uaaaaaaaaa aaacucaaaa aaaaaaaaga 5100

uaaaaaaaaa aaaccuaaaa aaaaaaaaug uaaaaaaaaa aaagggaaaa aaaaaaaacg 5160

caaaaaaaaa aaacacaaaa aaaaaaaaug caaaaaaaaa aaaucgaaaa aaaaaaaauc 5220

uaaaaaaaaa aaacgaaaaa aaaaaaaccc aaaaaaaaaa aagacaaaaa aaaaaaauag 5280

aaaaaaaaaa aaguuaaaaa aaaaaaacug aaaaaaaaaa aauuuaaaaa aaaaaaaucu 5340

ag 5342

<210> 345
<211> 5342
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 345
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cggcuccgag cccgccaccu 660

ccggcuccga gacccccggc accuccaccg acaagaagua cuccaucggc cuggccaucg 720

gcaccaacuc cgugggcugg gccgugauca ccgacgagua caaggugccc uccaagaagu 780

ucaaggugcu gggcaacacc gaccggcacu ccaucaagaa gaaccugauc ggcgcccugc 840

uguucgacuc cggcgagacc gccgaggcca cccggcugaa gcggaccgcc cggcggcggu 900

acacccggcg gaagaaccgg aucugcuacc ugcaggagau cuucuccaac gagauggcca 960

agguggacga cuccuucuuc caccggcugg aggaguccuu ccugguggag gaggacaaga 1020

agcacgagcg gcaccccauc uucggcaaca ucguggacga gguggccuac cacgagaagu 1080

accccaccau cuaccaccug cggaagaagc ugguggacuc caccgacaag gccgaccugc 1140

ggcugaucua ccuggcccug gcccacauga ucaaguuccg gggccacuuc cugaucgagg 1200

gcgaccugaa ccccgacaac uccgacgugg acaagcuguu cauccagcug gugcagaccu 1260

acaaccagcu guucgaggag aaccccauca acgccuccgg cguggacgcc aaggccaucc 1320

uguccgcccg gcuguccaag ucccggcggc uggagaaccu gaucgcccag cugcccggcg 1380

agaagaagaa cggccuguuc ggcaaccuga ucgcccuguc ccugggccug acccccaacu 1440

ucaaguccaa cuucgaccug gccgaggacg ccaagcugca gcuguccaag gacaccuacg 1500

acgacgaccu ggacaaccug cuggcccaga ucggcgacca guacgccgac cuguuccugg 1560

ccgccaagaa ccuguccgac gccauccugc uguccgacau ccugcgggug aacaccgaga 1620

ucaccaaggc cccccugucc gccuccauga ucaagcggua cgacgagcac caccaggacc 1680

ugacccugcu gaaggcccug gugcggcagc agcugcccga gaaguacaag gagaucuucu 1740

ucgaccaguc caagaacggc uacgccggcu acaucgacgg cggcgccucc caggaggagu 1800

ucuacaaguu caucaagccc auccuggaga agauggacgg caccgaggag cugcugguga 1860

agcugaaccg ggaggaccug cugcggaagc agcggaccuu cgacaacggc uccaucccccc 1920

accagaucca ccugggcgag cugcacgcca uccugcggcg gcaggaggac uucuaccccu 1980

uccugaagga caaccgggag aagaucgaga agauccugac cuuccggauc cccuacuacg 2040

ugggcccccu ggcccggggc aacucccggu ucgccuggau gacccggaag uccgaggaga 2100

ccaucacccc cuggaacuuc gaggaggugg uggacaaggg cgccuccgcc caguccuuca 2160

ucgagcggau gaccaacuuc gacaagaacc ugcccaacga gaaggugcug cccaagcacu 2220

cccugcugua cgaguacuuc accguguaca acgagcugac caaggugaag uacgugaccg 2280

agggcaugcg gaagcccgcc uuccuguccg gcgagcagaa gaaggccauc guggaccugc 2340

uguucaagac caaccggaag gugaccguga agcagcugaa ggaggacuac uucaagaaga 2400

ucgagugcuu cgacuccgug gagaucuccg gcguggagga ccgguucaac gccucccugg 2460

gcaccuacca cgaccugcug aagaucauca aggacaagga cuuccuggac aacgaggaga 2520

acgaggacau ccuggaggac aucgugcuga cccugacccu guucgaggac cgggagauga 2580

ucgaggagcg gcugaagacc uacgcccacc uguucgacga caaggugaug aagcagcuga 2640

agcggcggcg guacaccggc uggggccggc ugucccggaa gcugaucaac ggcauccggg 2700

acaagcaguc cggcaagacc auccuggacu uccugaaguc cgacggcuuc gccaaccgga 2760

acuucaugca gcugauccac gacgacuccc ugaccuucaa ggaggacauc cagaaggccc 2820

agguguccgg ccagggcgac ucccugcacg agcacaucgc caaccuggcc ggcucccccg 2880

ccaucaagaa gggcauccug cagaccguga agguggugga cgagcuggug aaggugaugg 2940

gccggcacaa gcccgagaac aucgugaucg agauggcccg ggagaaccag accacccaga 3000

agggccagaa gaacucccgg gagcggauga agcggaucga ggagggcauc aaggagcugg 3060

gcucccagau ccugaaggag caccccgugg agaacaccca gcugcagaac gagaagcugu 3120

accuguacua ccugcagaac ggccgggaca uguacgugga ccaggagcug gacaucaacc 3180

ggcuguccga cuacgacgug gaccacaucg ugccccaguc cuuccugaag gacgacucca 3240

ucgacaacaa ggugcugacc cgguccgaca agaaccgggg caaguccgac aacgugcccu 3300

ccgaggaggu ggugaagaag augaagaacu acuggcggca gcugcugaac gccaagcuga 3360

ucacccagcg gaaguucgac aaccugacca aggccgagcg gggcggccug uccgagcugg 3420

acaaggccgg cuucaucaag cggcagcugg uggagacccg gcagaucacc aagcacgugg 3480

cccagauccu ggacucccgg augaacacca aguacgacga gaacgacaag cugauccggg 3540

aggugaaggu gaucacccug aaguccaagc ugguguccga cuuccggaag gacuuccagu 3600

ucuacaaggu gcgggagauc aacaacuacc accacgccca cgacgccuac cugaacgccg 3660

uggugggcac cgcccugauc aagaaguacc ccaagcugga guccgaguuc guguacggcg 3720

acuacaaggu guacgacgug cggaagauga ucgccaaguc cgagcaggag aucggcaagg 3780

ccaccgccaa guacuucuuc uacuccaaca ucaugaacuu cuucaagacc gagaucaccc 3840

uggccaacgg cgagauccgg aagcggcccc ugaucgagac caacggcgag accggcgaga 3900

ucgugуggga caagggccgg gacuucgcca ccgugcggaa ggugcugucc augccccagg 3960

ugaacaucgu gaagaagacc gaggugcaga ccggcggcuu cuccaaggag uccauccugc 4020

ccaagcggaa cuccgacaag cugaucgccc ggaagaagga cugggacccc aagaaguacg 4080

gcggcuucga cucccccacc guggccuacu ccgugcuggu gguggccaag guggagaagg 4140

gcaaguccaa gaagcugaag uccgugaagg agcugcuggg caucaccauc auggagcggu 4200

ccuccuucga gaagaacccc aucgacuucc uggaggccaa gggcuacaag gaggugaaga 4260

aggaccugau caucaagcug cccaaguacu cccuguucga gcuggagaac ggccggaagc 4320

ggaugcuggc cuccgccggc gagcugcaga agggcaacga gcuggcccug cccuccaagu 4380

acgugaacuu ccuguaccug gccucccacu acgagaagcu gaagggcucc cccgaggaca 4440

acgagcagaa gcagcuguuc guggagcagc acaagcacua ccuggacgag aucaucgagc 4500

agaucuccga guucuccaag cgggugaucc uggccgacgc caaccuggac aaggugcugu 4560

ccgccuacaa caagcaccgg gacaagccca uccgggagca ggccgagaac aucauccacc 4620

uguucacccu gaccaaccug ggcgcccccg ccgccuucaa guacuucgac accaccaucg 4680

accggaagcg guacaccucc accaaggagg ugcuggacgc cacccugauc caccagucca 4740

ucaccggccu guacgagacc cggaucgacc ugucccagcu gggcggcgac ggcggcggcu 4800

cccccaagaa gaagcggaag gugugacuag caccagccuc aagaacaccc gaauggaguc 4860

ucuaagcuac auaauaccaa cuuacacuuu acaaaauguu gucccccaaa auguagccau 4920

ucguaucugc uccuaauaaa aagaaaguuu cuucacauuc ucucgagaaa aaaaaaaaau 4980

ggaaaaaaaa aaaacggaaa aaaaaaaaag guaaaaaaaa aaaauauaaa aaaaaaaaac 5040

auaaaaaaaa aaaacgaaaa aaaaaaaacg uaaaaaaaaa aaacucaaaa aaaaaaaaga 5100

uaaaaaaaaa aaaccuaaaa aaaaaaaaug uaaaaaaaaa aaagggaaaa aaaaaaaacg 5160

caaaaaaaaa aaacacaaaa aaaaaaaaug caaaaaaaaa aaaucgaaaa aaaaaaaauc 5220

uaaaaaaaaa aaacgaaaaa aaaaaaaccc aaaaaaaaaa aagacaaaaa aaaaaaauag 5280

aaaaaaaaaa aaguuaaaaa aaaaaaacug aaaaaaaaaa aauuuaaaaa aaaaaaaucu 5340

ag 5342

<210> 346
<211> 5339
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 346
gggaagcuca gaauaaacgc ucaacuuugg ccggaucugc caccauggag gccucccccg 60

ccuccggccc ccggcaccug auggaccccc acaucuucac cuccaacuuc aacaacggca 120

ucggccggca caagaccuac cugugcuacg agguggagcg gcuggacaac ggcaccuccg 180

ugaagaugga ccagcaccgg ggcuuccugc acaaccaggc caagaaccug cugugcggcu 240

ucuacggccg gcacgccgag cugcgguucc uggaccuggu gcccucccug cagcuggacc 300

ccgcccagau cuaccgggug accugguuca ucuccugguc ccccugcuuc uccuggggcu 360

gcgccggcga ggugcgggcc uuccugcagg agaacaccca cgugcggcug cggaucuucg 420

ccgcccggau cuacgacuac gacccccugu acaaggaggc ccugcagaug cugcgggacg 480

ccggcgccca gguguccauc augaccuacg acgaguucaa gcacugcugg gacaccuucg 540

uggaccacca gggcugcccc uuccagcccu gggacggccu ggacgagcac ucccaggccc 600

uguccggccg gcugcgggcc auccugcaga accagggcaa cgaggccgcc gccaaggagg 660

ccgccgccaa ggaggccgcc gccaaggaca agaaguacuc caucggccug gccaucggca 720

ccaacuccgu gggcugggcc gugaucaccg acgaguacaa ggugcccucc aagaaguuca 780

aggugcuggg caacaccgac cggcacucca ucaagaagaa ccugaucggc gcccugcugu 840

ucgacuccgg cgagaccgcc gaggccaccc ggcugaagcg gaccgcccgg cggcgguaca 900

cccggcggaa gaaccggauc ugcuaccugc aggagaucuu cuccaacgag auggccaagg 960

uggacgacuc cuucuuccac cggcuggagg aguccuuccu gguggaggag gacaagaagc 1020

acgagcggca ccccaucuuc ggcaacaucg uggacgaggu ggccuaccac gagaaguacc 1080

ccaccaucua ccaccugcgg aagaagcugg uggacuccac cgacaaggcc gaccugcggc 1140

ugaucuaccu ggcccuggcc cacaugauca aguuccgggg ccacuuccug aucgagggcg 1200

accugaaccc cgacaacucc gacguggaca agcuguucau ccagcuggug cagaccuaca 1260

accagcuguu cgaggagaac cccaucaacg ccuccggcgu ggacgccaag gccauccugu 1320

ccgcccggcu guccaagucc cggcggcugg agaaccugau cgcccagcug cccggcgaga 1380

agaagaacgg ccuguucggc aaccugaucg cccugucccu gggccugacc cccaacuuca 1440

aguccaacuu cgaccuggcc gaggacgcca agcugcagcu guccaaggac accuacgacg 1500

acgaccugga caaccugcug gcccagaucg gcgaccagua cgccgaccug uuccuggccg 1560

ccaagaaccu guccgacgcc auccugcugu ccgacauccu gcgggugaac accgagauca 1620

ccaaggcccc ccuguccgcc uccaugauca agcgguacga cgagcaccac caggaccuga 1680

cccugcugaa ggcccuggug cggcagcagc ugcccgagaa guacaaggag aucuucuucg 1740

accaguccaa gaacggcuac gccggcuaca ucgacggcgg cgccucccag gaggaguucu 1800

acaaguucau caagcccauc cuggagaaga uggacggcac cgaggagcug cuggugaagc 1860

ugaaccggga ggaccugcug cggaagcagc ggaccuucga caacggcucc aucccccacc 1920

agauccaccu gggcgagcug cacgccaucc ugcggcggca ggaggacuuc uaccccuucc 1980

ugaaggacaa ccgggagaag aucgagaaga uccugaccuu ccggaucccc uacuacgugg 2040

gcccccuggc ccggggcaac ucccgguucg ccuggaugac ccggaagucc gaggagacca 2100

ucacccccug gaacuucgag gagguggugg acaagggcgc cuccgcccag uccuucaucg 2160

agcggaugac caacuucgac aagaaccugc ccaacgagaa ggugcugccc aagcacuccc 2220

ugcuguacga guacuucacc guguacaacg agcugaccaa ggugaaguac gugaccgagg 2280

gcaugcggaa gcccgccuuc cuguccggcg agcagaagaa ggccaucgug gaccugcugu 2340

ucaagaccaa ccggaaggug accgugaagc agcugaagga ggacuacuuc aagaagaucg 2400

agugcuucga cuccguggag aucuccggcg uggaggaccg guucaacgcc ucccugggca 2460

ccuaccacga ccugcugaag aucaucaagg acaaggacuu ccuggacaac gaggagaacg 2520

aggacauccu ggaggacauc gugcugaccc ugacccuguu cgaggaccgg gagaugaucg 2580

aggagcggcu gaagaccuac gcccaccugu ucgacgacaa ggugaugaag cagcugaagc 2640

ggcggcggua caccggcugg ggccggcugu cccggaagcu gaucaacggc auccgggaca 2700

agcaguccgg caagaccauc cuggacuucc ugaaguccga cggcuucgcc aaccggaacu 2760

ucaugcagcu gauccacgac gacucccuga ccuucaagga ggacauccag aaggcccagg 2820

uguccggcca gggcgacucc cugcacgagc acaucgccaa ccuggccggc ucccccgcca 2880

ucaagaaggg cauccugcag accgugaagg ugguggacga gcuggugaag gugaugggcc 2940

ggcacaagcc cgagaacauc gugaucgaga uggcccggga gaaccagacc acccagaagg 3000

gccagaagaa cucccgggag cggaugaagc ggaucgagga gggcaucaag gagcugggcu 3060

cccagauccu gaaggagcac cccguggaga acacccagcu gcagaacgag aagcuguacc 3120

uguacuaccu gcagaacggc cgggacaugu acguggacca ggagcuggac aucaaccggc 3180

uguccgacua cgacguggac cacaucgugc cccaguccuu ccugaaggac gacuccaucg 3240

acaacaaggu gcugacccgg uccgacaaga accggggcaa guccgacaac gugcccuccg 3300

aggagguggu gaagaagaug aagaacuacu ggcggcagcu gcugaacgcc aagcugauca 3360

cccagcggaa guucgacaac cugaccaagg ccgagcgggg cggccugucc gagcuggaca 3420

aggccggcuu caucaagcgg cagcuggugg agacccggca gaucaccaag cacguggccc 3480

agauccugga cucccggaug aacaccaagu acgacgagaa cgacaagcug auccgggagg 3540

ugaaggugau cacccugaag uccaagcugg uguccgacuu ccggaaggac uuccaguucu 3600

acaaggugcg ggagaucaac aacuaccacc acgcccacga cgccuaccug aacgccgugg 3660

ugggcaccgc ccugaucaag aaguacccca agcuggaguc cgaguucgug uacggcgacu 3720

acaaggugua cgacgugcgg aagaugaucg ccaaguccga gcaggagauc ggcaaggcca 3780

ccgccaagua cuucuucuac uccaacauca ugaacuucuu caagaccgag aucacccugg 3840

ccaacggcga gauccggaag cggccccuga ucgagaccaa cggcgagacc ggcgagaucg 3900

ugugggacaa gggccgggac uucgccaccg ugcggaaggu gcuguccaug ccccagguga 3960

acaucgugaa gaagaccgag gugcagaccg gcggcuucuc caaggagucc auccugccca 4020

agcggaacuc cgacaagcug aucgcccgga agaaggacug ggaccccaag aaguacggcg 4080

gcuucgacuc ccccaccgug gccuacuccg ugcugguggu ggccaaggug gagaagggca 4140

aguccaagaa gcugaagucc gugaaggagc ugcugggcau caccaucaug gagcgguccu 4200

ccuucgagaa gaaccccauc gacuuccugg aggccaaggg cuacaaggag gugaagaagg 4260

accugaucau caagcugccc aaguacuccc uguucgagcu ggagaacggc cggaagcgga 4320

ugcuggccuc cgccggcgag cugcagaagg gcaacgagcu ggcccugccc uccaaguacg 4380

ugaacuuccu guaccuggcc ucccacuacg agaagcugaa gggcucccccc gaggacaacg 4440

agcagaagca gcuguucgug gagcagcaca agcacuaccu ggacgagauc aucgagcaga 4500

ucuccgaguu cuccaagcgg gugauccugg ccgacgccaa ccuggacaag gugcuguccg 4560

ccuacaacaa gcaccgggac aagcccaucc gggagcaggc cgagaacauc auccaccugu 4620

ucacccugac caaccugggc gcccccgccg ccuucaagua cuucgacacc accaucgacc 4680

ggaagcggua caccuccacc aaggaggugc uggacgccac ccugauccac caguccauca 4740

ccggccugua cgagacccgg aucgaccugu cccagcuggg cggcgacggc ggcggcuccc 4800

ccaagaagaa gcggaaggug ugacuagcac cagccucaag aacacccgaa uggagucucu 4860

aagcuacaua auaccaacuu acacuuuaca aaauguuguc ccccaaaaug uagccauucg 4920

uaucugcucc uaauaaaaag aaaguuucuu cacauucucu cgagaaaaaa aaaaaaugga 4980

aaaaaaaaaa acggaaaaaa aaaaaaggua aaaaaaaaaa auauaaaaaa aaaaaacaua 5040

aaaaaaaaaa acgaaaaaaa aaaaacguaa aaaaaaaaaa cucaaaaaaa aaaaagauaa 5100

aaaaaaaaaa ccuaaaaaaa aaaaauguaa aaaaaaaaaa gggaaaaaaa aaaaacgcaa 5160

aaaaaaaaaa cacaaaaaaa aaaaaugcaa aaaaaaaaaa ucgaaaaaaa aaaaaucuaa 5220

aaaaaaaaaa cgaaaaaaaa aaaacccaaa aaaaaaaaag acaaaaaaaa aaaauagaaa 5280

aaaaaaaaag uuaaaaaaaa aaaacugaaa aaaaaaaaau uuaaaaaaaa aaaaucuag 5339

<210> 347
<211> 5455
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 347
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgga ggaggaggaa 720

gcgaagcagc agcaaaggaa gcagcagcaa aggacaagaa guacagcauc ggacuggcca 780

ucggaacaaa cagcgucgga ugggcaguca ucacagacga auacaagguc ccgagcaaga 840

aguucaaggu ccugggaaac acagacagac acagcaucaa gaagaaccug aucggagcac 900

ugcuguucga cagcggagaa acagcagaag caacaagacu gaagagaaca gcaagaagaa 960

gauacacaag aagaaagaac agaaucugcu accugcagga aaucuucagc aacgaaaugg 1020

caaaggucga cgacagcuuc uuccacagac uggaagaaag cuuccugguc gaagaagaca 1080

agaagcacga aagacacccg aucuucggaa acaucgucga cgaagucgca uaccacgaaa 1140

aguacccgac aaucuaccac cugagaaaga agcuggucga cagcacagac aaggcagacc 1200

ugagacugau cuaccuggca cuggcacaca ugaucaaguu cagaggacac uuccugaucg 1260

aaggagaccu gaacccggac aacagcgacg ucgacaagcu guucauccag cugguccaga 1320

cauacaacca gcuguucgaa gaaaacccga ucaacgcaag cggagucgac gcaaaggcaa 1380

uccugagcgc aagacugagc aagagcagaa gacuggaaaa ccugaucgca cagcugccgg 1440

gagaaaagaa gaacggacug uucggaaacc ugaucgcacu gagccuggga cugacaccga 1500

acuucaagag caacuucgac cuggcagaag acgcaaagcu gcagcugagc aaggacacau 1560

acgacgacga ccuggacaac cugcuggcac agaucggaga ccaguacgca gaccuguucc 1620

uggcagcaaa gaaccugagc gacgcaaucc ugcugagcga cauccugaga gucaacacag 1680

aaaucacaaa ggcaccgcug agcgcaagca ugaucaagag auacgacgaa caccaccagg 1740

accugacacu gcugaaggca cuggucagac agcagcugcc ggaaaaguac aaggaaaucu 1800

ucuucgacca gagcaagaac ggauacgcag gauacaucga cggaggagca agccaggaag 1860

aauucuacaa guucaucaag ccgauccugg aaaagaugga cggaacagaa gaacugcugg 1920

ucaagcugaa cagagaagac cugcugagaa agcagagaac auucgacaac ggaagcaucc 1980

cgcaccagau ccaccuggga gaacugcacg caauccugag aagacaggaa gacuucuacc 2040

cguuccugaa ggacaacaga gaaaagaucg aaaagauccu gacauucaga aucccguacu 2100

acgucggacc gcuggcaaga ggaaacagca gauucgcaug gaugacaaga aagagcgaag 2160

aaacaaucac accguggaac uucgaagaag ucgucgacaa gggagcaagc gcacagagcu 2220

ucaucgaaag aaugacaaac uucgacaaga accugccgaa cgaaaagguc cugccgaagc 2280

acagccugcu guacgaauac uucacagucu acaacgaacu gacaaagguc aaguacguca 2340

cagaaggaau gagaaagccg gcauuccuga gcggagaaca gaagaaggca aucgucgacc 2400

ugcuguucaa gacaaacaga aaggucacag ucaagcagcu gaaggaagac uacuucaaga 2460

agaucgaaug cuucgacagc gucgaaauca gcggagucga agacagauuc aacgcaagcc 2520

ugggaacaua ccacgaccug cugaagauca ucaaggacaa ggacuuccug gacaacgaag 2580

aaaacgaaga cauccuggaa gacaucgucc ugacacugac acuguucgaa gacagagaaa 2640

ugaucgaaga aagacugaag acauacgcac accuguucga cgacaagguc augaagcagc 2700

ugaagagaag aagauacaca ggaugggggaa gacugagcag aaagcugauc aacggaauca 2760

gagacaagca gagcggaaag acaauccugg acuuccugaa gagcgacgga uucgcaaaca 2820

gaaacuucau gcagcugauc cacgacgaca gccugacauu caaggaagac auccagaagg 2880

cacaggucag cggacaggga gacagccugc acgaacacau cgcaaaccug gcaggaagcc 2940

cggcaaucaa gaagggaauc cugcagacag ucaaggucgu cgacgaacug gucaagguca 3000

ugggaagaca caagccggaa aacaucguca ucgaaauggc aagagaaaac cagacaacac 3060

agaagggaca gaagaacagc agagaaagaa ugaagagaau cgaagaagga aucaaggaac 3120

ugggaagcca gauccugaag gaacacccgg ucgaaaacac acagcugcag aacgaaaagc 3180

uguaccugua cuaccugcag aacggaagag acauguacgu cgaccaggaa cuggacauca 3240

acagacugag cgacuacgac gucgaccaca ucgucccgca gagcuuccug aaggacgaca 3300

gcaucgacaa caagguccug acaagaagcg acaagaacag aggaaagagc gacaacgucc 3360

cgagcgaaga agucgucaag aagaugaaga acuacuggag acagcugcug aacgcaaagc 3420

ugaucacaca gagaaaguuc gacaaccuga caaaggcaga gagaggagga cugagcgaac 3480

uggacaaggc aggauucauc aagagacagc uggucgaaac aagacagauc acaaagcacg 3540

ucgcacagau ccuggacagc agaaugaaca caaaguacga cgaaaacgac aagcugauca 3600

gagaagucaa ggucaucaca cugaagagca agcuggucag cgacuucaga aaggacuucc 3660

aguucuacaa ggucagagaa aucaacaacu accaccacgc acacgacgca uaccugaacg 3720

cagucgucgg aacagcacug aucaagaagu acccgaagcu ggaaagcgaa uucgucuacg 3780

gagacuacaa ggucuacgac gucagaaaga ugaucgcaaa gagcgaacag gaaaucggaa 3840

aggcaacagc aaaguacuuc uucuacagca acaucaugaa cuucuucaag acagaaauca 3900

cacuggcaaa cggagaaauc agaaagagac cgcugaucga aacaaacgga gaaacaggag 3960

aaaucgucug ggacaaggga agagacuucg caacagucag aaagguccug agcaugccgc 4020

aggucaacau cgucaagaag acagaagucc agacaggagg auucagcaag gaaagcaucc 4080

ugccgaagag aaacagcgac aagcugaucg caagaaagaa ggacugggac ccgaagaagu 4140

acggaggauu cgacagcccg acagucgcau acagcguccu ggucgucgca aaggucgaaa 4200

agggaaagag caagaagcug aagagcguca aggaacugcu gggaaucaca aucauggaaa 4260

gaagcagcuu cgaaaagaac ccgaucgacu uccuggaagc aaagggauac aaggaaguca 4320

agaaggaccu gaucaucaag cugccgaagu acagccuguu cgaacuggaa aacggaagaa 4380

agagaaugcu ggcaagcgca ggagaacugc agaagggaaa cgaacuggca cugccgagca 4440

aguacgucaa cuuccuguac cuggcaagcc acuacgaaaa gcugaaggga agcccggaag 4500

acaacgaaca gaagcagcug uucgucgaac agcacaagca cuaccuggac gaaaucaucg 4560

aacagaucag cgaauucagc aagagaguca uccuggcaga cgcaaaccug gacaaggucc 4620

ugagcgcaua caacaagcac agagacaagc cgaucagaga acaggcagaa aacaucaucc 4680

accuguucac acugacaaac cugggagcac cggcagcauu caaguacuuc gacacaacaa 4740

ucgacagaaa gagauacaca agcacaaagg aaguccugga cgcaacacug auccaccaga 4800

gcaucacagg acuguacgaa acaagaaucg aucugagcca gcugggagga gacagcggag 4860

gaagcacaaa ccugagcgac aucaucgaaa aggaaacagg aaagcagcug gucauccagg 4920

aaagcauccu gaugcugccg gaagaagucg aagaagucau cggaaacaag ccggaaagcg 4980

acauccuggu ccacacagca uacgacgaaa gcacagacga aaacgucaug cugcugacaa 5040

gcgacgcacc ggaauacaag ccgugggcac uggucaucca ggacagcaac ggagaaaaca 5100

agaucaagau gcugagcgga ggaagcccga agaagaagag aaaggucuaa uagucuagac 5160

aucacauuua aaagcaucuc agccuaccau gagaauaaga gaaagaaaau gaagaucaau 5220

agcuuauuca ucucuuuuuc uuuuucguug guguaaagcc aacacccugu cuaaaaaaca 5280

uaaauuucuu uaaucauuuu gccucuuuuc ucugugcuuc aauuaauaaa aaauggaaag 5340

aaccucgaga aaaaaaaaaa aaaaaaaaaa aaaaaaaaag cgaaaaaaaa aaaaaaaaaa 5400

aaaaaaaaaa aaccgaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaau 5455

<210> 348
<211> 5455
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 348
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag gccgccgcca 720

agggcggcgg cggcagcggc ggcggcggca gcgacaagaa guacagcauc ggacuggcca 780

ucggaacaaa cagcgucgga ugggcaguca ucacagacga auacaagguc ccgagcaaga 840

aguucaaggu ccugggaaac acagacagac acagcaucaa gaagaaccug aucggagcac 900

ugcuguucga cagcggagaa acagcagaag caacaagacu gaagagaaca gcaagaagaa 960

gauacacaag aagaaagaac agaaucugcu accugcagga aaucuucagc aacgaaaugg 1020

caaaggucga cgacagcuuc uuccacagac uggaagaaag cuuccugguc gaagaagaca 1080

agaagcacga aagacacccg aucuucggaa acaucgucga cgaagucgca uaccacgaaa 1140

aguacccgac aaucuaccac cugagaaaga agcuggucga cagcacagac aaggcagacc 1200

ugagacugau cuaccuggca cuggcacaca ugaucaaguu cagaggacac uuccugaucg 1260

aaggagaccu gaacccggac aacagcgacg ucgacaagcu guucauccag cugguccaga 1320

cauacaacca gcuguucgaa gaaaacccga ucaacgcaag cggagucgac gcaaaggcaa 1380

uccugagcgc aagacugagc aagagcagaa gacuggaaaa ccugaucgca cagcugccgg 1440

gagaaaagaa gaacggacug uucggaaacc ugaucgcacu gagccuggga cugacaccga 1500

acuucaagag caacuucgac cuggcagaag acgcaaagcu gcagcugagc aaggacacau 1560

acgacgacga ccuggacaac cugcuggcac agaucggaga ccaguacgca gaccuguucc 1620

uggcagcaaa gaaccugagc gacgcaaucc ugcugagcga cauccugaga gucaacacag 1680

aaaucacaaa ggcaccgcug agcgcaagca ugaucaagag auacgacgaa caccaccagg 1740

accugacacu gcugaaggca cuggucagac agcagcugcc ggaaaaguac aaggaaaucu 1800

ucuucgacca gagcaagaac ggauacgcag gauacaucga cggaggagca agccaggaag 1860

aauucuacaa guucaucaag ccgauccugg aaaagaugga cggaacagaa gaacugcugg 1920

ucaagcugaa cagagaagac cugcugagaa agcagagaac auucgacaac ggaagcaucc 1980

cgcaccagau ccaccuggga gaacugcacg caauccugag aagacaggaa gacuucuacc 2040

cguuccugaa ggacaacaga gaaaagaucg aaaagauccu gacauucaga aucccguacu 2100

acgucggacc gcuggcaaga ggaaacagca gauucgcaug gaugacaaga aagagcgaag 2160

aaacaaucac accguggaac uucgaagaag ucgucgacaa gggagcaagc gcacagagcu 2220

ucaucgaaag aaugacaaac uucgacaaga accugccgaa cgaaaagguc cugccgaagc 2280

acagccugcu guacgaauac uucacagucu acaacgaacu gacaaagguc aaguacguca 2340

cagaaggaau gagaaagccg gcauuccuga gcggagaaca gaagaaggca aucgucgacc 2400

ugcuguucaa gacaaacaga aaggucacag ucaagcagcu gaaggaagac uacuucaaga 2460

agaucgaaug cuucgacagc gucgaaauca gcggagucga agacagauuc aacgcaagcc 2520

ugggaacaua ccacgaccug cugaagauca ucaaggacaa ggacuuccug gacaacgaag 2580

aaaacgaaga cauccuggaa gacaucgucc ugacacugac acuguucgaa gacagagaaa 2640

ugaucgaaga aagacugaag acauacgcac accuguucga cgacaagguc augaagcagc 2700

ugaagagaag aagauacaca ggaugggggaa gacugagcag aaagcugauc aacggaauca 2760

gagacaagca gagcggaaag acaauccugg acuuccugaa gagcgacgga uucgcaaaca 2820

gaaacuucau gcagcugauc cacgacgaca gccugacauu caaggaagac auccagaagg 2880

cacaggucag cggacaggga gacagccugc acgaacacau cgcaaaccug gcaggaagcc 2940

cggcaaucaa gaagggaauc cugcagacag ucaaggucgu cgacgaacug gucaagguca 3000

ugggaagaca caagccggaa aacaucguca ucgaaauggc aagagaaaac cagacaacac 3060

agaagggaca gaagaacagc agagaaagaa ugaagagaau cgaagaagga aucaaggaac 3120

ugggaagcca gauccugaag gaacacccgg ucgaaaacac acagcugcag aacgaaaagc 3180

uguaccugua cuaccugcag aacggaagag acauguacgu cgaccaggaa cuggacauca 3240

acagacugag cgacuacgac gucgaccaca ucgucccgca gagcuuccug aaggacgaca 3300

gcaucgacaa caagguccug acaagaagcg acaagaacag aggaaagagc gacaacgucc 3360

cgagcgaaga agucgucaag aagaugaaga acuacuggag acagcugcug aacgcaaagc 3420

ugaucacaca gagaaaguuc gacaaccuga caaaggcaga gagaggagga cugagcgaac 3480

uggacaaggc aggauucauc aagagacagc uggucgaaac aagacagauc acaaagcacg 3540

ucgcacagau ccuggacagc agaaugaaca caaaguacga cgaaaacgac aagcugauca 3600

gagaagucaa ggucaucaca cugaagagca agcuggucag cgacuucaga aaggacuucc 3660

aguucuacaa ggucagagaa aucaacaacu accaccacgc acacgacgca uaccugaacg 3720

cagucgucgg aacagcacug aucaagaagu acccgaagcu ggaaagcgaa uucgucuacg 3780

gagacuacaa ggucuacgac gucagaaaga ugaucgcaaa gagcgaacag gaaaucggaa 3840

aggcaacagc aaaguacuuc uucuacagca acaucaugaa cuucuucaag acagaaauca 3900

cacuggcaaa cggagaaauc agaaagagac cgcugaucga aacaaacgga gaaacaggag 3960

aaaucgucug ggacaaggga agagacuucg caacagucag aaagguccug agcaugccgc 4020

aggucaacau cgucaagaag acagaagucc agacaggagg auucagcaag gaaagcaucc 4080

ugccgaagag aaacagcgac aagcugaucg caagaaagaa ggacugggac ccgaagaagu 4140

acggaggauu cgacagcccg acagucgcau acagcguccu ggucgucgca aaggucgaaa 4200

agggaaagag caagaagcug aagagcguca aggaacugcu gggaaucaca aucauggaaa 4260

gaagcagcuu cgaaaagaac ccgaucgacu uccuggaagc aaagggauac aaggaaguca 4320

agaaggaccu gaucaucaag cugccgaagu acagccuguu cgaacuggaa aacggaagaa 4380

agagaaugcu ggcaagcgca ggagaacugc agaagggaaa cgaacuggca cugccgagca 4440

aguacgucaa cuuccuguac cuggcaagcc acuacgaaaa gcugaaggga agcccggaag 4500

acaacgaaca gaagcagcug uucgucgaac agcacaagca cuaccuggac gaaaucaucg 4560

aacagaucag cgaauucagc aagagaguca uccuggcaga cgcaaaccug gacaaggucc 4620

ugagcgcaua caacaagcac agagacaagc cgaucagaga acaggcagaa aacaucaucc 4680

accuguucac acugacaaac cugggagcac cggcagcauu caaguacuuc gacacaacaa 4740

ucgacagaaa gagauacaca agcacaaagg aaguccugga cgcaacacug auccaccaga 4800

gcaucacagg acuguacgaa acaagaaucg aucugagcca gcugggagga gacagcggag 4860

gaagcacaaa ccugagcgac aucaucgaaa aggaaacagg aaagcagcug gucauccagg 4920

aaagcauccu gaugcugccg gaagaagucg aagaagucau cggaaacaag ccggaaagcg 4980

acauccuggu ccacacagca uacgacgaaa gcacagacga aaacgucaug cugcugacaa 5040

gcgacgcacc ggaauacaag ccgugggcac uggucaucca ggacagcaac ggagaaaaca 5100

agaucaagau gcugagcgga ggaagcccga agaagaagag aaaggucuaa uagucuagac 5160

aucacauuua aaagcaucuc agccuaccau gagaauaaga gaaagaaaau gaagaucaau 5220

agcuuauuca ucucuuuuuc uuuuucguug guguaaagcc aacacccugu cuaaaaaaca 5280

uaaauuucuu uaaucauuuu gccucuuuuc ucugugcuuc aauuaauaaa aaauggaaag 5340

aaccucgaga aaaaaaaaaa aaaaaaaaaa aaaaaaaaag cgaaaaaaaa aaaaaaaaaa 5400

aaaaaaaaaa aaccgaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaau 5455

<210> 349
<211> 5455
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 349
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag gccgccgcca 720

aggaggccgc cgccaaggag gccgccgcca aggacaagaa guacagcauc ggacuggcca 780

ucggaacaaa cagcgucgga ugggcaguca ucacagacga auacaagguc ccgagcaaga 840

aguucaaggu ccugggaaac acagacagac acagcaucaa gaagaaccug aucggagcac 900

ugcuguucga cagcggagaa acagcagaag caacaagacu gaagagaaca gcaagaagaa 960

gauacacaag aagaaagaac agaaucugcu accugcagga aaucuucagc aacgaaaugg 1020

caaaggucga cgacagcuuc uuccacagac uggaagaaag cuuccugguc gaagaagaca 1080

agaagcacga aagacacccg aucuucggaa acaucgucga cgaagucgca uaccacgaaa 1140

aguacccgac aaucuaccac cugagaaaga agcuggucga cagcacagac aaggcagacc 1200

ugagacugau cuaccuggca cuggcacaca ugaucaaguu cagaggacac uuccugaucg 1260

aaggagaccu gaacccggac aacagcgacg ucgacaagcu guucauccag cugguccaga 1320

cauacaacca gcuguucgaa gaaaacccga ucaacgcaag cggagucgac gcaaaggcaa 1380

uccugagcgc aagacugagc aagagcagaa gacuggaaaa ccugaucgca cagcugccgg 1440

gagaaaagaa gaacggacug uucggaaacc ugaucgcacu gagccuggga cugacaccga 1500

acuucaagag caacuucgac cuggcagaag acgcaaagcu gcagcugagc aaggacacau 1560

acgacgacga ccuggacaac cugcuggcac agaucggaga ccaguacgca gaccuguucc 1620

uggcagcaaa gaaccugagc gacgcaaucc ugcugagcga cauccugaga gucaacacag 1680

aaaucacaaa ggcaccgcug agcgcaagca ugaucaagag auacgacgaa caccaccagg 1740

accugacacu gcugaaggca cuggucagac agcagcugcc ggaaaaguac aaggaaaucu 1800

ucuucgacca gagcaagaac ggauacgcag gauacaucga cggaggagca agccaggaag 1860

aauucuacaa guucaucaag ccgauccugg aaaagaugga cggaacagaa gaacugcugg 1920

ucaagcugaa cagagaagac cugcugagaa agcagagaac auucgacaac ggaagcaucc 1980

cgcaccagau ccaccuggga gaacugcacg caauccugag aagacaggaa gacuucuacc 2040

cguuccugaa ggacaacaga gaaaagaucg aaaagauccu gacauucaga aucccguacu 2100

acgucggacc gcuggcaaga ggaaacagca gauucgcaug gaugacaaga aagagcgaag 2160

aaacaaucac accguggaac uucgaagaag ucgucgacaa gggagcaagc gcacagagcu 2220

ucaucgaaag aaugacaaac uucgacaaga accugccgaa cgaaaagguc cugccgaagc 2280

acagccugcu guacgaauac uucacagucu acaacgaacu gacaaagguc aaguacguca 2340

cagaaggaau gagaaagccg gcauuccuga gcggagaaca gaagaaggca aucgucgacc 2400

ugcuguucaa gacaaacaga aaggucacag ucaagcagcu gaaggaagac uacuucaaga 2460

agaucgaaug cuucgacagc gucgaaauca gcggagucga agacagauuc aacgcaagcc 2520

ugggaacaua ccacgaccug cugaagauca ucaaggacaa ggacuuccug gacaacgaag 2580

aaaacgaaga cauccuggaa gacaucgucc ugacacugac acuguucgaa gacagagaaa 2640

ugaucgaaga aagacugaag acauacgcac accuguucga cgacaagguc augaagcagc 2700

ugaagagaag aagauacaca ggaugggggaa gacugagcag aaagcugauc aacggaauca 2760

gagacaagca gagcggaaag acaauccugg acuuccugaa gagcgacgga uucgcaaaca 2820

gaaacuucau gcagcugauc cacgacgaca gccugacauu caaggaagac auccagaagg 2880

cacaggucag cggacaggga gacagccugc acgaacacau cgcaaaccug gcaggaagcc 2940

cggcaaucaa gaagggaauc cugcagacag ucaaggucgu cgacgaacug gucaagguca 3000

ugggaagaca caagccggaa aacaucguca ucgaaauggc aagagaaaac cagacaacac 3060

agaagggaca gaagaacagc agagaaagaa ugaagagaau cgaagaagga aucaaggaac 3120

ugggaagcca gauccugaag gaacacccgg ucgaaaacac acagcugcag aacgaaaagc 3180

uguaccugua cuaccugcag aacggaagag acauguacgu cgaccaggaa cuggacauca 3240

acagacugag cgacuacgac gucgaccaca ucgucccgca gagcuuccug aaggacgaca 3300

gcaucgacaa caagguccug acaagaagcg acaagaacag aggaaagagc gacaacgucc 3360

cgagcgaaga agucgucaag aagaugaaga acuacuggag acagcugcug aacgcaaagc 3420

ugaucacaca gagaaaguuc gacaaccuga caaaggcaga gagaggagga cugagcgaac 3480

uggacaaggc aggauucauc aagagacagc uggucgaaac aagacagauc acaaagcacg 3540

ucgcacagau ccuggacagc agaaugaaca caaaguacga cgaaaacgac aagcugauca 3600

gagaagucaa ggucaucaca cugaagagca agcuggucag cgacuucaga aaggacuucc 3660

aguucuacaa ggucagagaa aucaacaacu accaccacgc acacgacgca uaccugaacg 3720

cagucgucgg aacagcacug aucaagaagu acccgaagcu ggaaagcgaa uucgucuacg 3780

gagacuacaa ggucuacgac gucagaaaga ugaucgcaaa gagcgaacag gaaaucggaa 3840

aggcaacagc aaaguacuuc uucuacagca acaucaugaa cuucuucaag acagaaauca 3900

cacuggcaaa cggagaaauc agaaagagac cgcugaucga aacaaacgga gaaacaggag 3960

aaaucgucug ggacaaggga agagacuucg caacagucag aaagguccug agcaugccgc 4020

aggucaacau cgucaagaag acagaagucc agacaggagg auucagcaag gaaagcaucc 4080

ugccgaagag aaacagcgac aagcugaucg caagaaagaa ggacugggac ccgaagaagu 4140

acggaggauu cgacagcccg acagucgcau acagcguccu ggucgucgca aaggucgaaa 4200

agggaaagag caagaagcug aagagcguca aggaacugcu gggaaucaca aucauggaaa 4260

gaagcagcuu cgaaaagaac ccgaucgacu uccuggaagc aaagggauac aaggaaguca 4320

agaaggaccu gaucaucaag cugccgaagu acagccuguu cgaacuggaa aacggaagaa 4380

agagaaugcu ggcaagcgca ggagaacugc agaagggaaa cgaacuggca cugccgagca 4440

aguacgucaa cuuccuguac cuggcaagcc acuacgaaaa gcugaaggga agcccggaag 4500

acaacgaaca gaagcagcug uucgucgaac agcacaagca cuaccuggac gaaaucaucg 4560

aacagaucag cgaauucagc aagagaguca uccuggcaga cgcaaaccug gacaaggucc 4620

ugagcgcaua caacaagcac agagacaagc cgaucagaga acaggcagaa aacaucaucc 4680

accuguucac acugacaaac cugggagcac cggcagcauu caaguacuuc gacacaacaa 4740

ucgacagaaa gagauacaca agcacaaagg aaguccugga cgcaacacug auccaccaga 4800

gcaucacagg acuguacgaa acaagaaucg aucugagcca gcugggagga gacagcggag 4860

gaagcacaaa ccugagcgac aucaucgaaa aggaaacagg aaagcagcug gucauccagg 4920

aaagcauccu gaugcugccg gaagaagucg aagaagucau cggaaacaag ccggaaagcg 4980

acauccuggu ccacacagca uacgacgaaa gcacagacga aaacgucaug cugcugacaa 5040

gcgacgcacc ggaauacaag ccgugggcac uggucaucca ggacagcaac ggagaaaaca 5100

agaucaagau gcugagcgga ggaagcccga agaagaagag aaaggucuaa uagucuagac 5160

aucacauuua aaagcaucuc agccuaccau gagaauaaga gaaagaaaau gaagaucaau 5220

agcuuauuca ucucuuuuuc uuuuucguug guguaaagcc aacacccugu cuaaaaaaca 5280

uaaauuucuu uaaucauuuu gccucuuuuc ucugugcuuc aauuaauaaa aaauggaaag 5340

aaccucgaga aaaaaaaaaa aaaaaaaaaa aaaaaaaaag cgaaaaaaaa aaaaaaaaaa 5400

aaaaaaaaaa aaccgaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaau 5455

<210> 350
<211> 5470
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 350
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccugguccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgugggа cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacggc ggcggcggca 720

gcggcggcgg cggcagcggc ggcggcggca gcggcggcgg cggcagcgac aagaaguaca 780

gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca gacgaauaca 840

aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc aucaagaaga 900

accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca agacugaaga 960

gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug caggaaaucu 1020

ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa gaaagcuucc 1080

uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc gucgacgaag 1140

ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug gucgacagca 1200

cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc aaguucagag 1260

gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac aagcuguuca 1320

uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac gcaagcggag 1380

ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug gaaaaccuga 1440

ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc gcacugagcc 1500

ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca aagcugcagc 1560

ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc ggagaccagu 1620

acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug agcgacaucc 1680

ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc aagagauacg 1740

acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag cugccggaaa 1800

aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac aucgacggag 1860

gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag auggacggaa 1920

cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag agaacauucg 1980

acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc cugagaagac 2040

aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag auccugacau 2100

ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc gcauggauga 2160

caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc gacaagggag 2220

caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug ccgaacgaaa 2280

agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac gaacugacaa 2340

aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga gaacagaaga 2400

aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag cagcugaagg 2460

aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga gucgaagaca 2520

gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag gacaaggacu 2580

uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca cugacacugu 2640

ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug uucgacgaca 2700

aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug agcagaaagc 2760

ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc cugaagagcg 2820

acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug acauucaagg 2880

aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa cacaucgcaa 2940

accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag gucgucgacg 3000

aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa auggcaagag 3060

aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag agaaucgaag 3120

aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa aacacacagc 3180

ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug uacgucgacc 3240

aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc ccgcagagcu 3300

uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag aacagaggaa 3360

agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac uggagacagc 3420

ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag gcagagagag 3480

gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc gaaacaagac 3540

agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag uacgacgaaa 3600

acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug gucagcgacu 3660

ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac cacgcacacg 3720

acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg aagcuggaaa 3780

gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc gcaaagagcg 3840

aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc augaacuucu 3900

ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug aucgaaacaa 3960

acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca gucagaaagg 4020

uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca ggaggauuca 4080

gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga aagaaggacu 4140

gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc guccuggucg 4200

ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa cugcugggaa 4260

ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug gaagcaaagg 4320

gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc cuguucgaac 4380

uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag ggaaacgaac 4440

uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac gaaaagcuga 4500

agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac aagcacuacc 4560

uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug gcagacgcaa 4620

accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc agagaacagg 4680

cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca gcauucaagu 4740

acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc cuggacgcaa 4800

cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug agccagcugg 4860

gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa acaggaaagc 4920

agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa gucaucggaa 4980

acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca gacgaaaacg 5040

ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc auccaggaca 5100

gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag aagagaaagg 5160

ucuaauaguc uagacaucac auuuaaaagc aucucagccu accaugagaa uaagagaaag 5220

aaaaugaaga ucaauagcuu auucaucucu uuuucuuuuu cguuggugua aagccaacac 5280

ccugucuaaa aaacauaaau uucuuuaauc auuuugccuc uuuucucugu gcuucaauua 5340

auaaaaaaug gaaagaaccu cgagaaaaaa aaaaaaaaaa aaaaaaaaaa aaaagcgaaa 5400

aaaaaaaaaa aaaaaaaaaa aaaaaaaccg aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5460

aaaaaaaaau 5470

<210> 351
<211> 5470
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 351
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccugguccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacggc ggcggcggca 720

gcggaggagg aggaagcgaa gccgccgcca aagaggccgc cgccaaggac aagaaguaca 780

gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca gacgaauaca 840

aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc aucaagaaga 900

accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca agacugaaga 960

gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug caggaaaucu 1020

ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa gaaagcuucc 1080

uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc gucgacgaag 1140

ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug gucgacagca 1200

cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc aaguucagag 1260

gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac aagcuguuca 1320

uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac gcaagcggag 1380

ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug gaaaaccuga 1440

ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc gcacugagcc 1500

ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca aagcugcagc 1560

ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc ggagaccagu 1620

acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug agcgacaucc 1680

ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc aagagauacg 1740

acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag cugccggaaa 1800

aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac aucgacggag 1860

gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag auggacggaa 1920

cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag agaacauucg 1980

acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc cugagaagac 2040

aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag auccugacau 2100

ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc gcauggauga 2160

caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc gacaagggag 2220

caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug ccgaacgaaa 2280

agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac gaacugacaa 2340

aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga gaacagaaga 2400

aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag cagcugaagg 2460

aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga gucgaagaca 2520

gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag gacaaggacu 2580

uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca cugacacugu 2640

ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug uucgacgaca 2700

aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug agcagaaagc 2760

ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc cugaagagcg 2820

acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug acauucaagg 2880

aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa cacaucgcaa 2940

accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag gucgucgacg 3000

aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa auggcaagag 3060

aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag agaaucgaag 3120

aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa aacacacagc 3180

ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug uacgucgacc 3240

aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc ccgcagagcu 3300

uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag aacagaggaa 3360

agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac uggagacagc 3420

ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag gcagagagag 3480

gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc gaaacaagac 3540

agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag uacgacgaaa 3600

acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug gucagcgacu 3660

ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac cacgcacacg 3720

acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg aagcuggaaa 3780

gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc gcaaagagcg 3840

aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc augaacuucu 3900

ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug aucgaaacaa 3960

acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca gucagaaagg 4020

uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca ggaggauuca 4080

gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga aagaaggacu 4140

gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc guccuggucg 4200

ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa cugcugggaa 4260

ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug gaagcaaagg 4320

gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc cuguucgaac 4380

uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag ggaaacgaac 4440

uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac gaaaagcuga 4500

agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac aagcacuacc 4560

uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug gcagacgcaa 4620

accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc agagaacagg 4680

cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca gcauucaagu 4740

acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc cuggacgcaa 4800

cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug agccagcugg 4860

gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa acaggaaagc 4920

agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa gucaucggaa 4980

acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca gacgaaaacg 5040

ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc auccaggaca 5100

gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag aagagaaagg 5160

ucuaauaguc uagacaucac auuuaaaagc aucucagccu accaugagaa uaagagaaag 5220

aaaaugaaga ucaauagcuu auucaucucu uuuucuuuuu cguuggugua aagccaacac 5280

ccugucuaaa aaacauaaau uucuuuaauc auuuugccuc uuuucucugu gcuucaauua 5340

auaaaaaaug gaaagaaccu cgagaaaaaa aaaaaaaaaa aaaaaaaaaa aaaagcgaaa 5400

aaaaaaaaaa aaaaaaaaaa aaaaaaaccg aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5460

aaaaaaaaau 5470

<210> 352
<211> 5470
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 352
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgugggа cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacggc ggcggcggaa 720

gcgaggccgc ugccaaaggc ggaggaggaa gcggaggagg cggcagcgac aagaaguaca 780

gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca gacgaauaca 840

aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc aucaagaaga 900

accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca agacugaaga 960

gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug caggaaaucu 1020

ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa gaaagcuucc 1080

uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc gucgacgaag 1140

ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug gucgacagca 1200

cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc aaguucagag 1260

gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac aagcuguuca 1320

uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac gcaagcggag 1380

ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug gaaaaccuga 1440

ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc gcacugagcc 1500

ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca aagcugcagc 1560

ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc ggagaccagu 1620

acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug agcgacaucc 1680

ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc aagagauacg 1740

acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag cugccggaaa 1800

aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac aucgacggag 1860

gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag auggacggaa 1920

cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag agaacauucg 1980

acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc cugagaagac 2040

aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag auccugacau 2100

ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc gcauggauga 2160

caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc gacaagggag 2220

caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug ccgaacgaaa 2280

agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac gaacugacaa 2340

aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga gaacagaaga 2400

aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag cagcugaagg 2460

aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga gucgaagaca 2520

gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag gacaaggacu 2580

uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca cugacacugu 2640

ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug uucgacgaca 2700

aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug agcagaaagc 2760

ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc cugaagagcg 2820

acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug acauucaagg 2880

aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa cacaucgcaa 2940

accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag gucgucgacg 3000

aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa auggcaagag 3060

aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag agaaucgaag 3120

aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa aacacacagc 3180

ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug uacgucgacc 3240

aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc ccgcagagcu 3300

uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag aacagaggaa 3360

agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac uggagacagc 3420

ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag gcagagagag 3480

gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc gaaacaagac 3540

agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag uacgacgaaa 3600

acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug gucagcgacu 3660

ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac cacgcacacg 3720

acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg aagcuggaaa 3780

gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc gcaaagagcg 3840

aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc augaacuucu 3900

ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug aucgaaacaa 3960

acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca gucagaaagg 4020

uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca ggaggauuca 4080

gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga aagaaggacu 4140

gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc guccuggucg 4200

ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa cugcugggaa 4260

ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug gaagcaaagg 4320

gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc cuguucgaac 4380

uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag ggaaacgaac 4440

uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac gaaaagcuga 4500

agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac aagcacuacc 4560

uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug gcagacgcaa 4620

accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc agagaacagg 4680

cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca gcauucaagu 4740

acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc cuggacgcaa 4800

cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug agccagcugg 4860

gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa acaggaaagc 4920

agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa gucaucggaa 4980

acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca gacgaaaacg 5040

ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc auccaggaca 5100

gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag aagagaaagg 5160

ucuaauaguc uagacaucac auuuaaaagc aucucagccu accaugagaa uaagagaaag 5220

aaaaugaaga ucaauagcuu auucaucucu uuuucuuuuu cguuggugua aagccaacac 5280

ccugucuaaa aaacauaaau uucuuuaauc auuuugccuc uuuucucugu gcuucaauua 5340

auaaaaaaug gaaagaaccu cgagaaaaaa aaaaaaaaaa aaaaaaaaaa aaaagcgaaa 5400

aaaaaaaaaa aaaaaaaaaa aaaaaaaccg aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5460

aaaaaaaaau 5470

<210> 353
<211> 5470
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 353
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccugguccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgugggа cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag gccgccgcca 720

aggaggccgc cgccaaggag gccgccgcca aggaggccgc cgccaaggac aagaaguaca 780

gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca gacgaauaca 840

aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc aucaagaaga 900

accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca agacugaaga 960

gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug caggaaaucu 1020

ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa gaaagcuucc 1080

uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc gucgacgaag 1140

ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug gucgacagca 1200

cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc aaguucagag 1260

gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac aagcuguuca 1320

uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac gcaagcggag 1380

ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug gaaaaccuga 1440

ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc gcacugagcc 1500

ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca aagcugcagc 1560

ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc ggagaccagu 1620

acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug agcgacaucc 1680

ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc aagagauacg 1740

acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag cugccggaaa 1800

aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac aucgacggag 1860

gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag auggacggaa 1920

cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag agaacauucg 1980

acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc cugagaagac 2040

aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag auccugacau 2100

ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc gcauggauga 2160

caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc gacaagggag 2220

caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug ccgaacgaaa 2280

agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac gaacugacaa 2340

aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga gaacagaaga 2400

aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag cagcugaagg 2460

aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga gucgaagaca 2520

gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag gacaaggacu 2580

uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca cugacacugu 2640

ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug uucgacgaca 2700

aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug agcagaaagc 2760

ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc cugaagagcg 2820

acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug acauucaagg 2880

aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa cacaucgcaa 2940

accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag gucgucgacg 3000

aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa auggcaagag 3060

aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag agaaucgaag 3120

aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa aacacacagc 3180

ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug uacgucgacc 3240

aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc ccgcagagcu 3300

uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag aacagaggaa 3360

agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac uggagacagc 3420

ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag gcagagagag 3480

gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc gaaacaagac 3540

agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag uacgacgaaa 3600

acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug gucagcgacu 3660

ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac cacgcacacg 3720

acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg aagcuggaaa 3780

gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc gcaaagagcg 3840

aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc augaacuucu 3900

ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug aucgaaacaa 3960

acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca gucagaaagg 4020

uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca ggaggauuca 4080

gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga aagaaggacu 4140

gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc guccuggucg 4200

ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa cugcugggaa 4260

ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug gaagcaaagg 4320

gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc cuguucgaac 4380

uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag ggaaacgaac 4440

uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac gaaaagcuga 4500

agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac aagcacuacc 4560

uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug gcagacgcaa 4620

accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc agagaacagg 4680

cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca gcauucaagu 4740

acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc cuggacgcaa 4800

cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug agccagcugg 4860

gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa acaggaaagc 4920

agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa gucaucggaa 4980

acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca gacgaaaacg 5040

ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc auccaggaca 5100

gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag aagagaaagg 5160

ucuaauaguc uagacaucac auuuaaaagc aucucagccu accaugagaa uaagagaaag 5220

aaaaugaaga ucaauagcuu auucaucucu uuuucuuuuu cguuggugua aagccaacac 5280

ccugucuaaa aaacauaaau uucuuuaauc auuuugccuc uuuucucugu gcuucaauua 5340

auaaaaaaug gaaagaaccu cgagaaaaaa aaaaaaaaaa aaaaaaaaaa aaaagcgaaa 5400

aaaaaaaaaa aaaaaaaaaa aaaaaaaccg aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5460

aaaaaaaaau 5470

<210> 354
<211> 5485
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 354

gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccugguccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgga ggaggaggaa 720

gcgaagccgc cgccaaagag gccgccgcca agggaggcgg cggcagcgag gccgccgcca 780

aggacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga ugggcaguca 840

ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac acagacagac 900

acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa acagcagaag 960

caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac agaaucugcu 1020

accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc uuccacagac 1080

uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg aucuucggaa 1140

acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac cugagaaaga 1200

agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca cuggcacaca 1260

ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac aacagcgacg 1320

ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa gaaaacccga 1380

ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc aagagcagaa 1440

gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug uucggaaacc 1500

ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac cuggcagaag 1560

acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac cugcuggcac 1620

agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc gacgcaaucc 1680

ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug agcgcaagca 1740

ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca cuggucagac 1800

agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac ggauacgcag 1860

gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag ccgauccugg 1920

aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac cugcugagaa 1980

agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga gaacugcacg 2040

caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga gaaaagaucg 2100

aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga ggaaacagca 2160

gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac uucgaagaag 2220

ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac uucgacaaga 2280

accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac uucacagucu 2340

acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg gcauuccuga 2400

gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga aaggucacag 2460

ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc gucgaaauca 2520

gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug cugaagauca 2580

ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa gacaucgucc 2640

ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag acauacgcac 2700

accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca ggaugggggaa 2760

gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag acaauccugg 2820

acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc cacgacgaca 2880

gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga gacagccugc 2940

acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc cugcagacag 3000

ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa aacaucguca 3060

ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc agagaaagaa 3120

ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag gaacacccgg 3180

ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag aacggaagag 3240

acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac gucgaccaca 3300

ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug acaagaagcg 3360

acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag aagaugaaga 3420

acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc gacaaccuga 3480

caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc aagagacagc 3540

uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc agaaugaaca 3600

caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca cugaagagca 3660

agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa aucaacaacu 3720

accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug aucaagaagu 3780

acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac gucagaaaga 3840

ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc uucuacagca 3900

acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc agaaagagac 3960

cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga agagacuucg 4020

caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag acagaagucc 4080

agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac aagcugaucg 4140

caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg acagucgcau 4200

acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug aagagcguca 4260

aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac ccgaucgacu 4320

uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag cugccgaagu 4380

acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca ggagaacugc 4440

agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac cuggcaagcc 4500

acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug uucgucgaac 4560

agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc aagagaguca 4620

uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac agagacaagc 4680

cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac cugggagcac 4740

cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca agcacaaagg 4800

aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa acaagaaucg 4860

aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac aucaucgaaa 4920

aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg gaagaagucg 4980

aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca uacgacgaaa 5040

gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag ccgugggcac 5100

uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga ggaagcccga 5160

agaagaagag aaaggucuaa uagucuagac aucacauuua aaagcaucuc agccuaccau 5220

gagaauaaga gaaagaaaau gaagaucaau agcuuauuca ucucuuuuuc uuuuucguug 5280

guguaaagcc aacacccugu cuaaaaaaca uaaauuucuu uaaucauuuu gccucuuuuc 5340

ucugugcuuc aauuaauaaa aaauggaaag aaccucgaga aaaaaaaaaa aaaaaaaaaa 5400

aaaaaaaaag cgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaccgaaaaa aaaaaaaaaa 5460

aaaaaaaaaa aaaaaaaaaa aaaau 5485

<210> 355
<211> 5485
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 355
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag gccgccgcca 720

aggaggccgc cgccaagggc ggcggcggca gcggcggcgg cggcagcggc ggcggcggca 780

gcgacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga ugggcaguca 840

ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac acagacagac 900

acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa acagcagaag 960

caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac agaaucugcu 1020

accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc uuccacagac 1080

uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg aucuucggaa 1140

acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac cugagaaaga 1200

agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca cuggcacaca 1260

ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac aacagcgacg 1320

ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa gaaaacccga 1380

ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc aagagcagaa 1440

gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug uucggaaacc 1500

ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac cuggcagaag 1560

acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac cugcuggcac 1620

agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc gacgcaaucc 1680

ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug agcgcaagca 1740

ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca cuggucagac 1800

agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac ggauacgcag 1860

gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag ccgauccugg 1920

aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac cugcugagaa 1980

agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga gaacugcacg 2040

caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga gaaaagaucg 2100

aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga ggaaacagca 2160

gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac uucgaagaag 2220

ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac uucgacaaga 2280

accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac uucacagucu 2340

acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg gcauuccuga 2400

gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga aaggucacag 2460

ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc gucgaaauca 2520

gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug cugaagauca 2580

ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa gacaucgucc 2640

ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag acauacgcac 2700

accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca ggaugggggaa 2760

gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag acaauccugg 2820

acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc cacgacgaca 2880

gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga gacagccugc 2940

acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc cugcagacag 3000

ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa aacaucguca 3060

ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc agagaaagaa 3120

ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag gaacacccgg 3180

ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag aacggaagag 3240

acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac gucgaccaca 3300

ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug acaagaagcg 3360

acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag aagaugaaga 3420

acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc gacaaccuga 3480

caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc aagagacagc 3540

uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc agaaugaaca 3600

caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca cugaagagca 3660

agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa aucaacaacu 3720

accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug aucaagaagu 3780

acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac gucagaaaga 3840

ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc uucuacagca 3900

acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc agaaagagac 3960

cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga agagacuucg 4020

caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag acagaagucc 4080

agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac aagcugaucg 4140

caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg acagucgcau 4200

acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug aagagcguca 4260

aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac ccgaucgacu 4320

uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag cugccgaagu 4380

acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca ggagaacugc 4440

agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac cuggcaagcc 4500

acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug uucgucgaac 4560

agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc aagagaguca 4620

uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac agagacaagc 4680

cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac cugggagcac 4740

cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca agcacaaagg 4800

aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa acaagaaucg 4860

aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac aucaucgaaa 4920

aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg gaagaagucg 4980

aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca uacgacgaaa 5040

gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag ccgugggcac 5100

uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga ggaagcccga 5160

agaagaagag aaaggucuaa uagucuagac aucacauuua aaagcaucuc agccuaccau 5220

gagaauaaga gaaagaaaau gaagaucaau agcuuauuca ucucuuuuuc uuuuucguug 5280

guguaaagcc aacacccugu cuaaaaaaca uaaauuucuu uaaucauuuu gccucuuuuc 5340

ucugugcuuc aauuaauaaa aaauggaaag aaccucgaga aaaaaaaaaa aaaaaaaaaa 5400

aaaaaaaaag cgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaccgaaaaa aaaaaaaaaa 5460

aaaaaaaaaa aaaaaaaaaa aaaau 5485

<210> 356
<211> 5485
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 356
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccugguccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgugggа cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgaa gcagcagcaa 720

aggaagcagc agcaaaggga ggaggaggaa gcggaggagg aggaagcgaa gcagcagcaa 780

aggacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga ugggcaguca 840

ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac acagacagac 900

acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa acagcagaag 960

caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac agaaucugcu 1020

accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc uuccacagac 1080

uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg aucuucggaa 1140

acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac cugagaaaga 1200

agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca cuggcacaca 1260

ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac aacagcgacg 1320

ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa gaaaacccga 1380

ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc aagagcagaa 1440

gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug uucggaaacc 1500

ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac cuggcagaag 1560

acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac cugcuggcac 1620

agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc gacgcaaucc 1680

ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug agcgcaagca 1740

ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca cuggucagac 1800

agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac ggauacgcag 1860

gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag ccgauccugg 1920

aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac cugcugagaa 1980

agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga gaacugcacg 2040

caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga gaaaagaucg 2100

aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga ggaaacagca 2160

gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac uucgaagaag 2220

ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac uucgacaaga 2280

accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac uucacagucu 2340

acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg gcauuccuga 2400

gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga aaggucacag 2460

ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc gucgaaauca 2520

gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug cugaagauca 2580

ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa gacaucgucc 2640

ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag acauacgcac 2700

accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca ggaugggggaa 2760

gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag acaauccugg 2820

acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc cacgacgaca 2880

gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga gacagccugc 2940

acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc cugcagacag 3000

ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa aacaucguca 3060

ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc agagaaagaa 3120

ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag gaacacccgg 3180

ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag aacggaagag 3240

acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac gucgaccaca 3300

ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug acaagaagcg 3360

acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag aagaugaaga 3420

acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc gacaaccuga 3480

caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc aagagacagc 3540

uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc agaaugaaca 3600

caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca cugaagagca 3660

agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa aucaacaacu 3720

accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug aucaagaagu 3780

acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac gucagaaaga 3840

ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc uucuacagca 3900

acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc agaaagagac 3960

cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga agagacuucg 4020

caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag acagaagucc 4080

agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac aagcugaucg 4140

caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg acagucgcau 4200

acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug aagagcguca 4260

aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac ccgaucgacu 4320

uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag cugccgaagu 4380

acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca ggagaacugc 4440

agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac cuggcaagcc 4500

acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug uucgucgaac 4560

agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc aagagaguca 4620

uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac agagacaagc 4680

cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac cugggagcac 4740

cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca agcacaaagg 4800

aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa acaagaaucg 4860

aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac aucaucgaaa 4920

aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg gaagaagucg 4980

aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca uacgacgaaa 5040

gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag ccgugggcac 5100

uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga ggaagcccga 5160

agaagaagag aaaggucuaa uagucuagac aucacauuua aaagcaucuc agccuaccau 5220

gagaauaaga gaaagaaaau gaagaucaau agcuuauuca ucucuuuuuc uuuuucguug 5280

guguaaagcc aacacccugu cuaaaaaaca uaaauuucuu uaaucauuuu gccucuuuuc 5340

ucugugcuuc aauuaauaaa aaauggaaag aaccucgaga aaaaaaaaaa aaaaaaaaaa 5400

aaaaaaaaag cgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaccgaaaaa aaaaaaaaaa 5460

aaaaaaaaaa aaaaaaaaaa aaaau 5485

<210> 357
<211> 5485
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 357
gggagaccca agcuggcuag cguuuaaacu uaagcuuucc cgcagucggc guccagcggc 60

ucugcuuguu cgugugugug ucguugcagg ccuuauucgg auccgccacc auggaagcaa 120

gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc aacuucaaca 180

acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug gacaacggaa 240

caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag aaccugcugu 300

gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg agccugcagc 360

uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg ugcuucagcu 420

ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc agacugagaa 480

ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug cagaugcuga 540

gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac ugcugggaca 600

cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac gaacacagcc 660

aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgaa gcagcagcaa 720

aggaagcagc agcaaaggaa gcagcagcaa agggaggagg aggaagcgga ggaggaggaa 780

gcgacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga ugggcaguca 840

ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac acagacagac 900

acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa acagcagaag 960

caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac agaaucugcu 1020

accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc uuccacagac 1080

uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg aucuucggaa 1140

acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac cugagaaaga 1200

agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca cuggcacaca 1260

ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac aacagcgacg 1320

ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa gaaaacccga 1380

ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc aagagcagaa 1440

gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug uucggaaacc 1500

ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac cuggcagaag 1560

acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac cugcuggcac 1620

agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc gacgcaaucc 1680

ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug agcgcaagca 1740

ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca cuggucagac 1800

agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac ggauacgcag 1860

gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag ccgauccugg 1920

aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac cugcugagaa 1980

agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga gaacugcacg 2040

caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga gaaaagaucg 2100

aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga ggaaacagca 2160

gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac uucgaagaag 2220

ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac uucgacaaga 2280

accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac uucacagucu 2340

acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg gcauuccuga 2400

gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga aaggucacag 2460

ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc gucgaaauca 2520

gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug cugaagauca 2580

ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa gacaucgucc 2640

ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag acauacgcac 2700

accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca ggaugggggaa 2760

gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag acaauccugg 2820

acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc cacgacgaca 2880

gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga gacagccugc 2940

acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc cugcagacag 3000

ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa aacaucguca 3060

ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc agagaaagaa 3120

ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag gaacacccgg 3180

ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag aacggaagag 3240

acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac gucgaccaca 3300

ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug acaagaagcg 3360

acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag aagaugaaga 3420

acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc gacaaccuga 3480

caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc aagagacagc 3540

uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc agaaugaaca 3600

caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca cugaagagca 3660

agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa aucaacaacu 3720

accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug aucaagaagu 3780

acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac gucagaaaga 3840

ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc uucuacagca 3900

acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc agaaagagac 3960

cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga agagacuucg 4020

caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag acagaagucc 4080

agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac aagcugaucg 4140

caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg acagucgcau 4200

acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug aagagcguca 4260

aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac ccgaucgacu 4320

uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag cugccgaagu 4380

acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca ggagaacugc 4440

agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac cuggcaagcc 4500

acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug uucgucgaac 4560

agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc aagagaguca 4620

uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac agagacaagc 4680

cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac cugggagcac 4740

cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca agcacaaagg 4800

aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa acaagaaucg 4860

aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac aucaucgaaa 4920

aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg gaagaagucg 4980

aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca uacgacgaaa 5040

gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag ccgugggcac 5100

uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga ggaagcccga 5160

agaagaagag aaaggucuaa uagucuagac aucacauuua aaagcaucuc agccuaccau 5220

gagaauaaga gaaagaaaau gaagaucaau agcuuauuca ucucuuuuuc uuuuucguug 5280

guguaaagcc aacacccugu cuaaaaaaca uaaauuucuu uaaucauuuu gccucuuuuc 5340

ucugugcuuc aauuaauaaa aaauggaaag aaccucgaga aaaaaaaaaa aaaaaaaaaa 5400

aaaaaaaaag cgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaccgaaaaa aaaaaaaaaa 5460

aaaaaaaaaa aaaaaaaaaa aaaau 5485

<210> 358

<400> 358
000

<210> 359

<400> 359
000

<210> 360
<211> 3363
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 360
atggacggct ccggcggcgg ctcccccaag aagaagcgga aggtggagga caagcggccc 60

gccgccacca agaaggccgg ccaggccaag aagaagaagg gcggctccgg cggcggcgcc 120

gccttcaagc ccaaccccat caactacatc ctgggcctgg acatcggcat cgcctccgtg 180

ggctgggcca tggtggagat cgacgaggag gagaacccca tccggctgat cgacctgggc 240

gtgcgggtgt tcgagcgggc cgaggtgccc aagaccggcg actccctggc catggcccgg 300

cggctggccc ggtccgtgcg gcggctgacc cggcggcggg cccaccggct gctgcgggcc 360

cggcggctgc tgaagcggga gggcgtgctg caggccgccg acttcgacga gaacggcctg 420

atcaagtccc tgcccaacac cccctggcag ctgcgggccg ccgccctgga ccggaagctg 480

acccccctgg agtggtccgc cgtgctgctg cacctgatca agcaccgggg ctacctgtcc 540

cagcggaaga acgagggcga gaccgccgac aaggagctgg gcgccctgct gaagggcgtg 600

gccaacaacg cccacgccct gcagaccggc gacttccgga cccccgccga gctggccctg 660

aacaagttcg agaaggagtc cggccacatc cggaaccagc ggggcgacta ctcccacacc 720

ttctcccgga aggacctgca ggccgagctg atcctgctgt tcgagaagca gaaggagttc 780

ggcaaccccc acgtgtccgg cggcctgaag gagggcatcg agaccctgct gatgacccag 840

cggcccgccc tgtccggcga cgccgtgcag aagatgctgg gccactgcac cttcgagccc 900

gccgagccca aggccgccaa gaacacctac accgccgagc ggttcatctg gctgaccaag 960

ctgaacaacc tgcggatcct ggagcagggc tccgagcggc ccctgaccga caccgagcgg 1020

gccaccctga tggacgagcc ctaccggaag tccaagctga cctacgccca ggcccggaag 1080

ctgctgggcc tggaggacac cgccttcttc aagggcctgc ggtacggcaa ggacaacgcc 1140

gaggcctcca ccctgatgga gatgaaggcc taccacgcca tctcccgggc cctggagaag 1200

gagggcctga aggacaagaa gtccccccctg aacctgtcct ccgagctgca ggacgagatc 1260

ggcaccgcct tctccctgtt caagaccgac gaggacatca ccggccggct gaaggaccgg 1320

gtgcagcccg agatcctgga ggccctgctg aagcacatct ccttcgacaa gttcgtgcag 1380

atctccctga aggccctgcg gcggatcgtg cccctgatgg agcagggcaa gcggtacgac 1440

gaggcctgcg ccgagatcta cggcgaccac tacggcaaga agaacaccga ggagaagatc 1500

tacctgcccc ccatccccgc cgacgagatc cggaaccccg tggtgctgcg ggccctgtcc 1560

caggcccgga aggtgatcaa cggcgtggtg cggcggtacg gctcccccgc ccggatccac 1620

atcgagaccg cccgggaggt gggcaagtcc ttcaaggacc ggaaggagat cgagaagcgg 1680

caggaggaga accggaagga ccgggagaag gccgccgcca agttccggga gtacttcccc 1740

aacttcgtgg gcgagcccaa gtccaaggac atcctgaagc tgcggctgta cgagcagcag 1800

cacggcaagt gcctgtactc cggcaaggag atcaacctgg tgcggctgaa cgagaagggc 1860

tacgtggaga tcgaccacgc cctgcccttc tcccggacct gggacgactc cttcaacaac 1920

aaggtgctgg tgctgggctc cgagaaccag aacaagggca accagacccc ctacgagtac 1980

ttcaacggca aggacaactc ccgggagtgg caggagttca aggcccgggt ggagacctcc 2040

cggttccccc ggtccaagaa gcagcggatc ctgctgcaga agttcgacga ggacggcttc 2100

aaggagtgca acctgaacga cacccggtac gtgaaccggt tcctgtgcca gttcgtggcc 2160

gaccacatcc tgctgaccgg caagggcaag cggcgggtgt tcgcctccaa cggccagatc 2220

accaacctgc tgcggggctt ctggggcctg cggaaggtgc gggccgagaa cgaccggcac 2280

cacgccctgg acgccgtggt ggtggcctgc tccaccgtgg ccatgcagca gaagatcacc 2340

cggttcgtgc ggtacaagga gatgaacgcc ttcgacggca agaccatcga caaggagacc 2400

ggcaaggtgc tgcaccagaa gacccacttc ccccagccct gggagttctt cgcccaggag 2460

gtgatgatcc gggtgttcgg caagcccgac ggcaagcccg agttcgagga ggccgacacc 2520

cccgagaagc tgcggaccct gctggccgag aagctgtcct cccggcccga ggccgtgcac 2580

gagtacgtga ccccctgtt cgtgtcccgg gcccccaacc ggaagatgtc cggcgcccac 2640

aaggacaccc tgcggtccgc caagcggttc gtgaagcaca acgagaagat ctccgtgaag 2700

cgggtgtggc tgaccgagat caagctggcc gacctggaga acatggtgaa ctacaagaac 2760

ggccgggaga tcgagctgta cgaggccctg aaggcccggc tggaggccta cggcggcaac 2820

gccaagcagg ccttcgaccc caaggacaac cccttctaca agaagggcgg ccagctggtg 2880

aaggccgtgc gggtggagaa gacccaggag tccggcgtgc tgctgaacaa gaagaacgcc 2940

tacaccatcg ccgacaacgg cgacatggtg cgggtggacg tgttctgcaa ggtggacaag 3000

aagggcaaga accagtactt catcgtgccc atctacgcct ggcaggtggc cgagaacatc 3060

ctgcccgaca tcgactgcaa gggctaccgg atcgacgact cctacacctt ctgcttctcc 3120

ctgcacaagt acgacctgat cgccttccag aaggacgaga agtccaaggt ggagttcgcc 3180

tactacatca actgcgactc ctccaacggc cggttctacc tggcctggca cgacaagggc 3240

tccaaggagc agcagttccg gatctccacc cagaacctgg tgctgatcca gaagtaccag 3300

gtgaacgagc tgggcaagga gatccggccc tgccggctga agaagcggcc ccccgtgcgg 3360

tag 3363

<210> 361
<211> 4782
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 361
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacggc 600

accaaggacu ccaccaagga caucccccgag acccccucca aggacgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggauccccu acuacguggg cccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau cacccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccuggggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacggcg gcggcucccc caagaagaag cggaaggugu ga 4782

<210> 362
<211> 4782
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 362
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacggc 600

cgggacgugc ggcagcccga ggugaaggag gagaagcccg aguccgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggauccccu acuacguggg ccccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau cacccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccuggggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacggcg gcggcucccc caagaagaag cggaaggugu ga 4782

<210> 363
<211> 4782
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 363

auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacgag 600

ggcaaguccu ccggcuccgg cuccgagucc aaguccaccg ccggcgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggauccccu acuacguggg ccccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau caccccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccuggggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacggcg gcggcucccc caagaagaag cggaaggugu ga 4782

<210> 364
<211> 4782
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 364
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacacc 600

cccggcuccc ccgccggcuc ccccaccucc accgaggagg gcaccgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggauccccu acuacguggg cccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau cacccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccuggggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacggcg gcggcucccc caagaagaag cggaaggugu ga 4782

<210> 365
<211> 4782
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 365
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacggc 600

uccgagcccg ccaccuccgg cuccgagacc cccggcaccu ccaccgacaa gaaguacucc 660

aucggccugg ccaucggcac caacuccgug ggcugggccg ugaucaccga cgaguacaag 720

gugcccucca agaaguucaa ggugcugggc aacaccgacc ggcacuccau caagaagaac 780

cugaucggcg cccugcuguu cgacuccggc gagaccgccg aggccacccg gcugaagcgg 840

accgcccggc ggcgguacac ccggcggaag aaccggaucu gcuaccugca ggagaucuuc 900

uccaacgaga uggccaaggu ggacgacucc uucuuccacc ggcuggagga guccuuccug 960

guggaggagg acaagaagca cgagcggcac cccaucuucg gcaacaucgu ggacgaggug 1020

gccuaccacg agaaguaccc caccaucuac caccugcgga agaagcuggu ggacuccacc 1080

gacaaggccg accugcggcu gaucuaccug gcccuggccc acaugaucaa guuccggggc 1140

cacuuccuga ucgagggcga ccugaacccc gacaacuccg acguggacaa gcuguucauc 1200

cagcuggugc agaccuacaa ccagcuguuc gaggagaacc ccaucaacgc cuccggcgug 1260

gacgccaagg ccauccuguc cgcccggcug uccaaguccc ggcggcugga gaaccugauc 1320

gcccagcugc ccggcgagaa gaagaacggc cuguucggca accugaucgc ccugucccug 1380

ggccugaccc ccaacuucaa guccaacuuc gaccuggccg aggacgccaa gcugcagcug 1440

uccaaggaca ccuacgacga cgaccuggac aaccugcugg cccagaucgg cgaccaguac 1500

gccgaccugu uccuggccgc caagaaccug uccgacgcca uccugcuguc cgacauccug 1560

cgggugaaca ccgagaucac caaggccccc cuguccgccu ccaugaucaa gcgguacgac 1620

gagcaccacc aggaccugac ccugcugaag gcccuggugc ggcagcagcu gcccgagaag 1680

uacaaggaga ucuucuucga ccaguccaag aacggcuacg ccggcuacau cgacggcggc 1740

gccucccagg aggaguucua caaguucauc aagcccaucc uggagaagau ggacggcacc 1800

gaggagcugc uggugaagcu gaaccgggag gaccugcugc ggaagcagcg gaccuucgac 1860

aacggcucca ucccccacca gauccaccug ggcgagcugc acgccauccu gcggcggcag 1920

gaggacuucu accccuuccu gaaggacaac cgggagaaga ucgagaagau ccugaccuuc 1980

cggaucccu acuacguggg ccccccuggcc cggggcaacu cccgguucgc cuggaugacc 2040

cggaaguccg aggagaccau cacccccugg aacuucgagg agguggugga caagggcgcc 2100

uccgcccagu ccuucaucga gcggaugacc aacuucgaca agaaccugcc caacgagaag 2160

gugcugccca agcacucccu gcuguacgag uacuucaccg uguacaacga gcugaccaag 2220

gugaaguacg ugaccgaggg caugcggaag cccgccuucc uguccggcga gcagaagaag 2280

gccaucgugg accugcuguu caagaccaac cggaagguga ccgugaagca gcugaaggag 2340

gacuacuuca agaagaucga gugcuucgac uccguggaga ucuccggcgu ggaggaccgg 2400

uucaacgccu cccugggcac cuaccacgac cugcugaaga ucaucaagga caaggacuuc 2460

cuggacaacg aggagaacga ggacauccug gaggacaucg ugcugacccu gacccuguuc 2520

gaggaccggg agaugaucga ggagcggcug aagaccuacg cccaccuguu cgacgacaag 2580

gugaugaagc agcugaagcg gcggcgguac accggcuggg gccggcuguc ccggaagcug 2640

aucaacggca uccgggacaa gcaguccggc aagaccaucc uggacuuccu gaaguccgac 2700

ggcuucgcca accggaacuu caugcagcug auccacgacg acucccugac cuucaaggag 2760

gacauccaga aggcccaggu guccggccag ggcgacuccc ugcacgagca caucgccaac 2820

cuggccggcu cccccgccau caagaagggc auccugcaga ccgugaaggu gguggacgag 2880

cuggugaagg ugaugggccg gcacaagccc gagaacaucg ugaucgagau ggcccgggag 2940

aaccagacca cccagaaggg ccagaagaac ucccgggagc ggaugaagcg gaucgaggag 3000

ggcaucaagg agcugggcuc ccagauccug aaggagcacc ccguggagaa cacccagcug 3060

cagaacgaga agcuguaccu guacuaccug cagaacggcc gggacaugua cguggaccag 3120

gagcuggaca ucaaccggcu guccgacuac gacguggacc acaucgugcc ccaguccuuc 3180

cugaaggacg acuccaucga caacaaggug cugacccggu ccgacaagaa ccggggcaag 3240

uccgacaacg ugcccuccga ggagguggug aagaagauga agaacuacug gcggcagcug 3300

cugaacgcca agcugaucac ccagcggaag uucgacaacc ugaccaaggc cgagcggggc 3360

ggccuguccg agcuggacaa ggccggcuuc aucaagcggc agcuggugga gacccggcag 3420

aucaccaagc acguggccca gauccuggac ucccggauga acaccaagua cgacgagaac 3480

gacaagcuga uccgggaggu gaaggugauc acccugaagu ccaagcuggu guccgacuuc 3540

cggaaggacu uccaguucua caaggugcgg gagaucaaca acuaccacca cgcccacgac 3600

gccuaccuga acgccguggu gggcaccgcc cugaucaaga aguaccccaa gcuggagucc 3660

gaguucgugu acggcgacua caagguguac gacgugcgga agaugaucgc caaguccgag 3720

caggagaucg gcaaggccac cgccaaguac uucuucuacu ccaacaucau gaacuucuuc 3780

aagaccgaga ucacccuggc caacggcgag auccggaagc ggccccugau cgagaccaac 3840

ggcgagaccg gcgagaucgu gugggacaag ggccgggacu ucgccaccgu gcggaaggug 3900

cuguccaugc cccaggugaa caucgugaag aagaccgagg ugcagaccgg cggcuucucc 3960

aaggagucca uccugcccaa gcggaacucc gacaagcuga ucgcccggaa gaaggacugg 4020

gaccccaaga aguacggcgg cuucgacucc cccaccgugg ccuacuccgu gcugguggug 4080

gccaaggugg agaagggcaa guccaagaag cugaaguccg ugaaggagcu gcugggcauc 4140

accaucaugg agcgguccuc cuucgagaag aaccccaucg acuuccugga ggccaagggc 4200

uacaaggagg ugaagaagga ccugaucauc aagcugccca aguacucccu guucgagcug 4260

gagaacggcc ggaagcggau gcuggccucc gccggcgagc ugcagaaggg caacgagcug 4320

gcccugcccu ccaaguacgu gaacuuccug uaccuggccu cccacuacga gaagcugaag 4380

ggcucccccg aggacaacga gcagaagcag cuguucgugg agcagcacaa gcacuaccug 4440

gacgagauca ucgagcagau cuccgaguuc uccaagcggg ugauccuggc cgacgccaac 4500

cuggacaagg ugcuguccgc cuacaacaag caccgggaca agcccauccg ggagcaggcc 4560

gagaacauca uccaccuguu cacccugacc aaccugggcg cccccgccgc cuucaaguac 4620

uucgacacca ccaucgaccg gaagcgguac accuccacca aggaggugcu ggacgccacc 4680

cugauccacc aguccaucac cggccuguac gagacccgga ucgaccuguc ccagcugggc 4740

ggcgacggcg gcggcucccc caagaagaag cggaaggugu ga 4782

<210> 366
<211> 4779
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 366
auggaggccu cccccgccuc cggccccccgg caccugaugg acccccacau cuucaccucc 60

aacuucaaca acggcaucgg ccggcacaag accuaccugu gcuacgaggu ggagcggcug 120

gacaacggca ccuccgugaa gauggaccag caccggggcu uccugcacaa ccaggccaag 180

aaccugcugu gcggcuucua cggccggcac gccgagcugc gguuccugga ccuggugccc 240

ucccugcagc uggaccccgc ccagaucuac cgggugaccu gguucaucuc cuggucccccc 300

ugcuucuccu ggggcugcgc cggcgaggug cgggccuucc ugcaggagaa cacccacgug 360

cggcugcgga ucuucgccgc ccggaucuac gacuacgacc cccuguacaa ggaggcccug 420

cagaugcugc gggacgccgg cgcccaggug uccaucauga ccuacgacga guucaagcac 480

ugcugggaca ccuucgugga ccaccagggc ugccccuucc agcccuggga cggccuggac 540

gagcacuccc aggcccuguc cggccggcug cgggccaucc ugcagaacca gggcaacgag 600

gccgccgcca aggaggccgc cgccaaggag gccgccgcca aggacaagaa guacuccauc 660

ggccuggcca ucggcaccaa cuccgugggc ugggccguga ucaccgacga guacaaggug 720

cccuccaaga aguucaaggu gcugggcaac accgaccggc acuccaucaa gaagaaccug 780

aucggcgccc ugcuguucga cuccggcgag accgccgagg ccacccggcu gaagcggacc 840

gcccggcggc gguacacccg gcggaagaac cggaucugcu accugcagga gaucuucucc 900

aacgagaugg ccaaggugga cgacuccuuc uuccaccggc uggaggaguc cuuccuggug 960

gaggaggaca agaagcacga gcggcacccc aucuucggca acaucgugga cgagguggcc 1020

uaccacgaga aguaccccac caucuaccac cugcggaaga agcuggugga cuccaccgac 1080

aaggccgacc ugcggcugau cuaccuggcc cuggcccaca ugaucaaguu ccggggccac 1140

uuccugaucg agggcgaccu gaaccccgac aacuccgacg uggacaagcu guucauccag 1200

cuggugcaga ccuacaacca gcuguucgag gagaacccca ucaacgccuc cggcguggac 1260

gccaaggcca uccuguccgc ccggcugucc aagucccggc ggcuggagaa ccugaucgcc 1320

cagcugcccg gcgagaagaa gaacggccug uucggcaacc ugaucgcccu gucccugggc 1380

cugaccccca acuucaaguc caacuucgac cuggccgagg acgccaagcu gcagcugucc 1440

aaggacaccu acgacgacga ccuggacaac cugcuggccc agaucggcga ccaguacgcc 1500

gaccuguucc uggccgccaa gaaccugucc gacgccaucc ugcuguccga cauccugcgg 1560

gugaacaccg agaucaccaa ggccccccug uccgccucca ugaucaagcg guacgacgag 1620

caccaccagg accugacccu gcugaaggcc cuggugcggc agcagcugcc cgagaaguac 1680

aaggagaucu ucuucgacca guccaagaac ggcuacgccg gcuacaucga cggcggcgcc 1740

ucccaggagg aguucuacaa guucaucaag cccauccugg agaagaugga cggcaccgag 1800

gagcugcugg ugaagcugaa ccgggaggac cugcugcgga agcagcggac cuucgacaac 1860

ggcuccaucc cccaccagau ccaccugggc gagcugcacg ccauccugcg gcggcaggag 1920

gacuucuacc ccuuccugaa ggacaaccgg gagaagaucg agaagauccu gaccuuccgg 1980

aucccccuacu acgugggccc ccuggcccgg ggcaacuccc gguucgccug gaugacccgg 2040

aaguccgagg agaccaucac ccccuggaac uucgaggagg ugguggacaa gggcgccucc 2100

gcccaguccu ucaucgagcg gaugaccaac uucgacaaga accugcccaa cgagaaggug 2160

cugcccaagc acucccugcu guacgaguac uucaccgugu acaacgagcu gaccaaggug 2220

aaguacguga ccgagggcau gcggaagccc gccuuccugu ccggcgagca gaagaaggcc 2280

aucguggacc ugcuguucaa gaccaaccgg aaggugaccg ugaagcagcu gaaggaggac 2340

uacuucaaga agaucgagug cuucgacucc guggagaucu ccggcgugga ggaccgguuc 2400

aacgccuccc ugggcaccua ccacgaccug cugaagauca ucaaggacaa ggacuuccug 2460

gacaacgagg agaacgagga cauccuggag gacaucgugc ugacccugac ccuguucgag 2520

gaccgggaga ugaucgagga gcggcugaag accuacgccc accuguucga cgacaaggug 2580

augaagcagc ugaagcggcg gcgguacacc ggcuggggcc ggcugucccg gaagcugauc 2640

aacggcaucc gggacaagca guccggcaag accauccugg acuuccugaa guccgacggc 2700

uucgccaacc ggaacuucau gcagcugauc cacgacgacu cccugaccuu caaggaggac 2760

auccagaagg cccagguguc cggccagggc gacucccugc acgagcacau cgccaaccug 2820

gccggcuccc ccgccaucaa gaagggcauc cugcagaccg ugaagguggu ggacgagcug 2880

gugaagguga ugggccggca caagcccgag aacaucguga ucgagauggc ccgggagaac 2940

cagaccaccc agaagggcca gaagaacucc cgggagcgga ugaagcggau cgaggagggc 3000

aucaaggagc ugggcuccca gauccugaag gagcaccccg uggagaacac ccagcugcag 3060

aacgagaagc uguaccugua cuaccugcag aacggccggg acauguacgu ggaccaggag 3120

cuggacauca accggcuguc cgacuacgac guggaccaca ucgugcccca guccuuccug 3180

aaggacgacu ccaucgacaa caaggugcug acccgguccg acaagaaccg gggcaagucc 3240

gacaacgugc ccuccgagga gguggugaag aagaugaaga acuacuggcg gcagcugcug 3300

aacgccaagc ugaucaccca gcggaaguuc gacaaccuga ccaaggccga gcggggcggc 3360

cuguccgagc uggacaaggc cggcuucauc aagcggcagc ugguggagac ccggcagauc 3420

accaagcacg uggcccagau ccuggacucc cggaugaaca ccaaguacga cgagaacgac 3480

aagcugaucc gggaggugaa ggugaucacc cugaagucca agcugguguc cgacuuccgg 3540

aaggacuucc aguucuacaa ggugcgggag aucaacaacu accaccacgc ccacgacgcc 3600

uaccugaacg ccgugguggg caccgcccug aucaagaagu accccaagcu ggaguccgag 3660

uucguguacg gcgacuacaa gguguacgac gugcggaaga ugaucgccaa guccgagcag 3720

gagaucggca aggccaccgc caaguacuuc uucuacucca acaucaugaa cuucuucaag 3780

accgagauca cccuggccaa cggcgagauc cggaagcggc cccugaucga gaccaacggc 3840

gagaccggcg agaucgugug ggacaagggc cgggacuucg ccaccgugcg gaaggugcug 3900

uccaugcccc aggugaacau cgugaagaag accgaggugc agaccggcgg cuucuccaag 3960

gaguccaucc ugcccaagcg gaacuccgac aagcugaucg cccggaagaa ggacugggac 4020

cccaagaagu acggcggcuu cgacuccccc accguggccu acuccgugcu ggugguggcc 4080

aagguggaga agggcaaguc caagaagcug aaguccguga aggagcugcu gggcaucacc 4140

aucauggagc gguccuccuu cgagaagaac cccaucgacu uccuggaggc caagggcuac 4200

aaggagguga agaaggaccu gaucaucaag cugcccaagu acucccuguu cgagcuggag 4260

aacggccgga agcggaugcu ggccuccgcc ggcgagcugc agaagggcaa cgagcuggcc 4320

cugcccucca aguacgugaa cuuccuguac cuggccuccc acuacgagaa gcugaagggc 4380

ucccccgagg acaacgagca gaagcagcug uucguggagc agcacaagca cuaccuggac 4440

gagaucaucg agcagaucuc cgaguucucc aagcggguga uccuggccga cgccaaccug 4500

gacaaggugc uguccgccua caacaagcac cgggacaagc ccauccggga gcaggccgag 4560

aacaucaucc accuguucac ccugaccaac cugggcgccc ccgccgccuu caaguacuuc 4620

gacaccacca ucgaccggaa gcgguacacc uccaccaagg aggugcugga cgccacccug 4680

auccaccagu ccaucaccgg ccuguacgag acccggaucg accuguccca gcugggcggc 4740

gacggcggcg gcucccccaa gaagaagcgg aagguguga 4779

<210> 367
<211> 5042
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 367
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgga 600

ggaggaggaa gcgaagcagc agcaaaggaa gcagcagcaa aggacaagaa guacagcauc 660

ggacuggcca ucggaacaaa cagcgucgga ugggcaguca ucacagacga auacaagguc 720

ccgagcaaga aguucaaggu ccugggaaac acagacagac acagcaucaa gaagaaccug 780

aucggagcac ugcuguucga cagcggagaa acagcagaag caacaagacu gaagagaaca 840

gcaagaagaa gauacacaag aagaaagaac agaaucugcu accugcagga aaucuucagc 900

aacgaaaugg caaaggucga cgacagcuuc uuccacagac uggaagaaag cuuccugguc 960

gaagaagaca agaagcacga aagacacccg aucuucggaa acaucgucga cgaagucgca 1020

uaccacgaaa aguacccgac aaucuaccac cugagaaaga agcuggucga cagcacagac 1080

aaggcagacc ugagacugau cuaccuggca cuggcacaca ugaucaaguu cagaggacac 1140

uuccugaucg aaggagaccu gaacccggac aacagcgacg ucgacaagcu guucauccag 1200

cugguccaga cauacaacca gcuguucgaa gaaaacccga ucaacgcaag cggagucgac 1260

gcaaaggcaa uccugagcgc aagacugagc aagagcagaa gacuggaaaa ccugaucgca 1320

cagcugccgg gagaaaagaa gaacggacug uucggaaacc ugaucgcacu gagccuggga 1380

cugacaccga acuucaagag caacuucgac cuggcagaag acgcaaagcu gcagcugagc 1440

aaggacacau acgacgacga ccuggacaac cugcuggcac agaucggaga ccaguacgca 1500

gaccuguucc uggcagcaaa gaaccugagc gacgcaaucc ugcugagcga cauccugaga 1560

gucaacacag aaaucacaaa ggcaccgcug agcgcaagca ugaucaagag auacgacgaa 1620

caccaccagg accugacacu gcugaaggca cuggucagac agcagcugcc ggaaaaguac 1680

aaggaaaucu ucuucgacca gagcaagaac ggauacgcag gauacaucga cggaggagca 1740

agccaggaag aauucuacaa guucaucaag ccgauccugg aaaagaugga cggaacagaa 1800

gaacugcugg ucaagcugaa cagagaagac cugcugagaa agcagagaac auucgacaac 1860

ggaagcaucc cgcaccagau ccaccuggga gaacugcacg caauccugag aagacaggaa 1920

gacuucuacc cguuccugaa ggacaacaga gaaaagaucg aaaagauccu gacauucaga 1980

aucccguacu acgucggacc gcuggcaaga ggaaacagca gauucgcaug gaugacaaga 2040

aagagcgaag aaacaaucac accguggaac uucgaagaag ucgucgacaa gggagcaagc 2100

gcacagagcu ucaucgaaag aaugacaaac uucgacaaga accugccgaa cgaaaagguc 2160

cugccgaagc acagccugcu guacgaauac uucacagucu acaacgaacu gacaaagguc 2220

aaguacguca cagaaggaau gagaaagccg gcauuccuga gcggagaaca gaagaaggca 2280

aucgucgacc ugcuguucaa gacaaacaga aaggucacag ucaagcagcu gaaggaagac 2340

uacuucaaga agaucgaaug cuucgacagc gucgaaauca gcggagucga agacagauuc 2400

aacgcaagcc ugggaacaua ccacgaccug cugaagauca ucaaggacaa ggacuuccug 2460

gacaacgaag aaaacgaaga cauccuggaa gacaucgucc ugacacugac acuguucgaa 2520

gacagagaaa ugaucgaaga aagacugaag acauacgcac accuguucga cgacaagguc 2580

augaagcagc ugaagagaag aagauacaca ggaugggga gacugagcag aaagcugauc 2640

aacggaauca gagacaagca gagcggaaag acaauccugg acuuccugaa gagcgacgga 2700

uucgcaaaca gaaacuucau gcagcugauc cacgacgaca gccugacauu caaggaagac 2760

auccagaagg cacaggucag cggacaggga gacagccugc acgaacacau cgcaaaccug 2820

gcaggaagcc cggcaaucaa gaagggaauc cugcagacag ucaaggucgu cgacgaacug 2880

gucaagguca ugggaagaca caagccggaa aacaucguca ucgaaauggc aagagaaaac 2940

cagacaacac agaagggaca gaagaacagc agagaaagaa ugaagagaau cgaagaagga 3000

aucaaggaac ugggaagcca gauccugaag gaacacccgg ucgaaaacac acagcugcag 3060

aacgaaaagc uguaccugua cuaccugcag aacggaagag acauguacgu cgaccaggaa 3120

cuggacauca acagacugag cgacuacgac gucgaccaca ucgucccgca gagcuuccug 3180

aaggacgaca gcaucgacaa caagguccug acaagaagcg acaagaacag aggaaagagc 3240

gacaacgucc cgagcgaaga agucgucaag aagaugaaga acuacuggag acagcugcug 3300

aacgcaaagc ugaucacaca gagaaaguuc gacaaccuga caaaggcaga gagaggagga 3360

cugagcgaac uggacaaggc aggauucauc aagagacagc uggucgaaac aagacagauc 3420

acaaagcacg ucgcacagau ccuggacagc agaaugaaca caaaguacga cgaaaacgac 3480

aagcugauca gagaagucaa ggucaucaca cugaagagca agcuggucag cgacuucaga 3540

aaggacuucc aguucuacaa ggucagagaa aucaacaacu accaccacgc acacgacgca 3600

uaccugaacg cagucgucgg aacagcacug aucaagaagu acccgaagcu ggaaagcgaa 3660

uucgucuacg gagacuacaa ggucuacgac gucagaaaga ugaucgcaaa gagcgaacag 3720

gaaaucggaa aggcaacagc aaaguacuuc uucuacagca acaucaugaa cuucuucaag 3780

acagaaauca cacuggcaaa cggagaaauc agaaagagac cgcugaucga aacaaacgga 3840

gaaacaggag aaaucgucug ggacaaggga agagacuucg caacagucag aaagguccug 3900

agcaugccgc aggucaacau cgucaagaag acagaagucc agacaggagg auucagcaag 3960

gaaagcaucc ugccgaagag aaacagcgac aagcugaucg caagaaagaa ggacugggac 4020

ccgaagaagu acggaggauu cgacagcccg acagucgcau acagcguccu ggucgucgca 4080

aaggucgaaa agggaaagag caagaagcug aagagcguca aggaacugcu gggaaucaca 4140

aucauggaaa gaagcagcuu cgaaaagaac ccgaucgacu uccuggaagc aaagggauac 4200

aaggaaguca agaaggaccu gaucaucaag cugccgaagu acagccuguu cgaacuggaa 4260

aacggaagaa agagaaugcu ggcaagcgca ggagaacugc agaagggaaa cgaacuggca 4320

cugccgagca aguacgucaa cuuccuguac cuggcaagcc acuacgaaaa gcugaaggga 4380

agcccggaag acaacgaaca gaagcagcug uucgucgaac agcacaagca cuaccuggac 4440

gaaaucaucg aacagaucag cgaauucagc aagagaguca uccuggcaga cgcaaaccug 4500

gacaaggucc ugagcgcaua caacaagcac agagacaagc cgaucagaga acaggcagaa 4560

aacaucaucc accuguucac acugacaaac cugggagcac cggcagcauu caaguacuuc 4620

gacacaacaa ucgacagaaa gagauacaca agcacaaagg aaguccugga cgcaacacug 4680

auccaccaga gcaucacagg acuguacgaa acaagaaucg aucugagcca gcugggagga 4740

gacagcggag gaagcacaaa ccugagcgac aucaucgaaa aggaaacagg aaagcagcug 4800

gucauccagg aaagcauccu gaugcugccg gaagaagucg aagaagucau cggaaacaag 4860

ccggaaagcg acauccuggu ccacacagca uacgacgaaa gcacagacga aaacgucaug 4920

cugcugacaa gcgacgcacc ggaauacaag ccgugggcac uggucaucca ggacagcaac 4980

ggagaaaaca agaucaagau gcugagcgga ggaagcccga agaagaagag aaaggucuaa 5040

ua 5042

<210> 368
<211> 5043
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 368
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag 600

gccgccgcca agggcggcgg cggcagcggc ggcggcggca gcgacaagaa guacagcauc 660

ggacuggcca ucggaacaaa cagcgucgga ugggcaguca ucacagacga auacaagguc 720

ccgagcaaga aguucaaggu ccugggaaac acagacagac acagcaucaa gaagaaccug 780

aucggagcac ugcuguucga cagcggagaa acagcagaag caacaagacu gaagagaaca 840

gcaagaagaa gauacacaag aagaaagaac agaaucugcu accugcagga aaucuucagc 900

aacgaaaugg caaaggucga cgacagcuuc uuccacagac uggaagaaag cuuccugguc 960

gaagaagaca agaagcacga aagacacccg aucuucggaa acaucgucga cgaagucgca 1020

uaccacgaaa aguacccgac aaucuaccac cugagaaaga agcuggucga cagcacagac 1080

aaggcagacc ugagacugau cuaccuggca cuggcacaca ugaucaaguu cagaggacac 1140

uuccugaucg aaggagaccu gaacccggac aacagcgacg ucgacaagcu guucauccag 1200

cugguccaga cauacaacca gcuguucgaa gaaaacccga ucaacgcaag cggagucgac 1260

gcaaaggcaa uccugagcgc aagacugagc aagagcagaa gacuggaaaa ccugaucgca 1320

cagcugccgg gagaaaagaa gaacggacug uucggaaacc ugaucgcacu gagccuggga 1380

cugacaccga acuucaagag caacuucgac cuggcagaag acgcaaagcu gcagcugagc 1440

aaggacacau acgacgacga ccuggacaac cugcuggcac agaucggaga ccaguacgca 1500

gaccuguucc uggcagcaaa gaaccugagc gacgcaaucc ugcugagcga cauccugaga 1560

gucaacacag aaaucacaaa ggcaccgcug agcgcaagca ugaucaagag auacgacgaa 1620

caccaccagg accugacacu gcugaaggca cuggucagac agcagcugcc ggaaaaguac 1680

aaggaaaucu ucuucgacca gagcaagaac ggauacgcag gauacaucga cggaggagca 1740

agccaggaag aauucuacaa guucaucaag ccgauccugg aaaagaugga cggaacagaa 1800

gaacugcugg ucaagcugaa cagagaagac cugcugagaa agcagagaac auucgacaac 1860

ggaagcaucc cgcaccagau ccaccuggga gaacugcacg caauccugag aagacaggaa 1920

gacuucuacc cguuccugaa ggacaacaga gaaaagaucg aaaagauccu gacauucaga 1980

aucccguacu acgucggacc gcuggcaaga ggaaacagca gauucgcaug gaugacaaga 2040

aagagcgaag aaacaaucac accguggaac uucgaagaag ucgucgacaa gggagcaagc 2100

gcacagagcu ucaucgaaag aaugacaaac uucgacaaga accugccgaa cgaaaagguc 2160

cugccgaagc acagccugcu guacgaauac uucacagucu acaacgaacu gacaaagguc 2220

aaguacguca cagaaggaau gagaaagccg gcauuccuga gcggagaaca gaagaaggca 2280

aucgucgacc ugcuguucaa gacaaacaga aaggucacag ucaagcagcu gaaggaagac 2340

uacuucaaga agaucgaaug cuucgacagc gucgaaauca gcggagucga agacagauuc 2400

aacgcaagcc ugggaacaua ccacgaccug cugaagauca ucaaggacaa ggacuuccug 2460

gacaacgaag aaaacgaaga cauccuggaa gacaucgucc ugacacugac acuguucgaa 2520

gacagagaaa ugaucgaaga aagacugaag acauacgcac accuguucga cgacaagguc 2580

augaagcagc ugaagagaag aagauacaca ggaugggga gacugagcag aaagcugauc 2640

aacggaauca gagacaagca gagcggaaag acaauccugg acuuccugaa gagcgacgga 2700

uucgcaaaca gaaacuucau gcagcugauc cacgacgaca gccugacauu caaggaagac 2760

auccagaagg cacaggucag cggacaggga gacagccugc acgaacacau cgcaaaccug 2820

gcaggaagcc cggcaaucaa gaagggaauc cugcagacag ucaaggucgu cgacgaacug 2880

gucaagguca ugggaagaca caagccggaa aacaucguca ucgaaauggc aagagaaaac 2940

cagacaacac agaagggaca gaagaacagc agagaaagaa ugaagagaau cgaagaagga 3000

aucaaggaac ugggaagcca gauccugaag gaacacccgg ucgaaaacac acagcugcag 3060

aacgaaaagc uguaccugua cuaccugcag aacggaagag acauguacgu cgaccaggaa 3120

cuggacauca acagacugag cgacuacgac gucgaccaca ucgucccgca gagcuuccug 3180

aaggacgaca gcaucgacaa caagguccug acaagaagcg acaagaacag aggaaagagc 3240

gacaacgucc cgagcgaaga agucgucaag aagaugaaga acuacuggag acagcugcug 3300

aacgcaaagc ugaucacaca gagaaaguuc gacaaccuga caaaggcaga gagaggagga 3360

cugagcgaac uggacaaggc aggauucauc aagagacagc uggucgaaac aagacagauc 3420

acaaagcacg ucgcacagau ccuggacagc agaaugaaca caaaguacga cgaaaacgac 3480

aagcugauca gagaagucaa ggucaucaca cugaagagca agcuggucag cgacuucaga 3540

aaggacuucc aguucuacaa ggucagagaa aucaacaacu accaccacgc acacgacgca 3600

uaccugaacg cagucgucgg aacagcacug aucaagaagu acccgaagcu ggaaagcgaa 3660

uucgucuacg gagacuacaa ggucuacgac gucagaaaga ugaucgcaaa gagcgaacag 3720

gaaaucggaa aggcaacagc aaaguacuuc uucuacagca acaucaugaa cuucuucaag 3780

acagaaauca cacuggcaaa cggagaaauc agaaagagac cgcugaucga aacaaacgga 3840

gaaacaggag aaaucgucug ggacaaggga agagacuucg caacagucag aaagguccug 3900

agcaugccgc aggucaacau cgucaagaag acagaagucc agacaggagg auucagcaag 3960

gaaagcaucc ugccgaagag aaacagcgac aagcugaucg caagaaagaa ggacugggac 4020

ccgaagaagu acggaggauu cgacagcccg acagucgcau acagcguccu ggucgucgca 4080

aaggucgaaa agggaaagag caagaagcug aagagcguca aggaacugcu gggaaucaca 4140

aucauggaaa gaagcagcuu cgaaaagaac ccgaucgacu uccuggaagc aaagggauac 4200

aaggaaguca agaaggaccu gaucaucaag cugccgaagu acagccuguu cgaacuggaa 4260

aacggaagaa agagaaugcu ggcaagcgca ggagaacugc agaagggaaa cgaacuggca 4320

cugccgagca aguacgucaa cuuccuguac cuggcaagcc acuacgaaaa gcugaaggga 4380

agcccggaag acaacgaaca gaagcagcug uucgucgaac agcacaagca cuaccuggac 4440

gaaaucaucg aacagaucag cgaauucagc aagagaguca uccuggcaga cgcaaaccug 4500

gacaaggucc ugagcgcaua caacaagcac agagacaagc cgaucagaga acaggcagaa 4560

aacaucaucc accuguucac acugacaaac cugggagcac cggcagcauu caaguacuuc 4620

gacacaacaa ucgacagaaa gagauacaca agcacaaagg aaguccugga cgcaacacug 4680

auccaccaga gcaucacagg acuguacgaa acaagaaucg aucugagcca gcugggagga 4740

gacagcggag gaagcacaaa ccugagcgac aucaucgaaa aggaaacagg aaagcagcug 4800

gucauccagg aaagcauccu gaugcugccg gaagaagucg aagaagucau cggaaacaag 4860

ccggaaagcg acauccuggu ccacacagca uacgacgaaa gcacagacga aaacgucaug 4920

cugcugacaa gcgacgcacc ggaauacaag ccgugggcac uggucaucca ggacagcaac 4980

ggagaaaaca agaucaagau gcugagcgga ggaagcccga agaagaagag aaaggucuaa 5040

uag 5043

<210> 369
<211> 5043
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 369
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag 600

gccgccgcca aggaggccgc cgccaaggag gccgccgcca aggacaagaa guacagcauc 660

ggacuggcca ucggaacaaa cagcgucgga ugggcaguca ucacagacga auacaagguc 720

ccgagcaaga aguucaaggu ccugggaaac acagacagac acagcaucaa gaagaaccug 780

aucggagcac ugcuguucga cagcggagaa acagcagaag caacaagacu gaagagaaca 840

gcaagaagaa gauacacaag aagaaagaac agaaucugcu accugcagga aaucuucagc 900

aacgaaaugg caaaggucga cgacagcuuc uuccacagac uggaagaaag cuuccugguc 960

gaagaagaca agaagcacga aagacacccg aucuucggaa acaucgucga cgaagucgca 1020

uaccacgaaa aguacccgac aaucuaccac cugagaaaga agcuggucga cagcacagac 1080

aaggcagacc ugagacugau cuaccuggca cuggcacaca ugaucaaguu cagaggacac 1140

uuccugaucg aaggagaccu gaacccggac aacagcgacg ucgacaagcu guucauccag 1200

cugguccaga cauacaacca gcuguucgaa gaaaacccga ucaacgcaag cggagucgac 1260

gcaaaggcaa uccugagcgc aagacugagc aagagcagaa gacuggaaaa ccugaucgca 1320

cagcugccgg gagaaaagaa gaacggacug uucggaaacc ugaucgcacu gagccuggga 1380

cugacaccga acuucaagag caacuucgac cuggcagaag acgcaaagcu gcagcugagc 1440

aaggacacau acgacgacga ccuggacaac cugcuggcac agaucggaga ccaguacgca 1500

gaccuguucc uggcagcaaa gaaccugagc gacgcaaucc ugcugagcga cauccugaga 1560

gucaacacag aaaucacaaa ggcaccgcug agcgcaagca ugaucaagag auacgacgaa 1620

caccaccagg accugacacu gcugaaggca cuggucagac agcagcugcc ggaaaaguac 1680

aaggaaaucu ucuucgacca gagcaagaac ggauacgcag gauacaucga cggaggagca 1740

agccaggaag aauucuacaa guucaucaag ccgauccugg aaaagaugga cggaacagaa 1800

gaacugcugg ucaagcugaa cagagaagac cugcugagaa agcagagaac auucgacaac 1860

ggaagcaucc cgcaccagau ccaccuggga gaacugcacg caauccugag aagacaggaa 1920

gacuucuacc cguuccugaa ggacaacaga gaaaagaucg aaaagauccu gacauucaga 1980

aucccguacu acgucggacc gcuggcaaga ggaaacagca gauucgcaug gaugacaaga 2040

aagagcgaag aaacaaucac accguggaac uucgaagaag ucgucgacaa gggagcaagc 2100

gcacagagcu ucaucgaaag aaugacaaac uucgacaaga accugccgaa cgaaaagguc 2160

cugccgaagc acagccugcu guacgaauac uucacagucu acaacgaacu gacaaagguc 2220

aaguacguca cagaaggaau gagaaagccg gcauuccuga gcggagaaca gaagaaggca 2280

aucgucgacc ugcuguucaa gacaaacaga aaggucacag ucaagcagcu gaaggaagac 2340

uacuucaaga agaucgaaug cuucgacagc gucgaaauca gcggagucga agacagauuc 2400

aacgcaagcc ugggaacaua ccacgaccug cugaagauca ucaaggacaa ggacuuccug 2460

gacaacgaag aaaacgaaga cauccuggaa gacaucgucc ugacacugac acuguucgaa 2520

gacagagaaa ugaucgaaga aagacugaag acauacgcac accuguucga cgacaagguc 2580

augaagcagc ugaagagaag aagauacaca ggauggggaa gacugagcag aaagcugauc 2640

aacggaauca gagacaagca gagcggaaag acaauccugg acuuccugaa gagcgacgga 2700

uucgcaaaca gaaacuucau gcagcugauc cacgacgaca gccugacauu caaggaagac 2760

auccagaagg cacaggucag cggacaggga gacagccugc acgaacacau cgcaaaccug 2820

gcaggaagcc cggcaaucaa gaagggaauc cugcagacag ucaaggucgu cgacgaacug 2880

gucaagguca ugggaagaca caagccggaa aacaucguca ucgaaauggc aagagaaaac 2940

cagacaacac agaagggaca gaagaacagc agagaaagaa ugaagagaau cgaagaagga 3000

aucaaggaac ugggaagcca gauccugaag gaacacccgg ucgaaaacac acagcugcag 3060

aacgaaaagc uguaccugua cuaccugcag aacggaagag acauguacgu cgaccaggaa 3120

cuggacauca acagacugag cgacuacgac gucgaccaca ucgucccgca gagcuuccug 3180

aaggacgaca gcaucgacaa caagguccug acaagaagcg acaagaacag aggaaagagc 3240

gacaacgucc cgagcgaaga agucgucaag aagaugaaga acuacuggag acagcugcug 3300

aacgcaaagc ugaucacaca gagaaaguuc gacaaccuga caaaggcaga gagaggagga 3360

cugagcgaac uggacaaggc aggauucauc aagagacagc uggucgaaac aagacagauc 3420

acaaagcacg ucgcacagau ccuggacagc agaaugaaca caaaguacga cgaaaacgac 3480

aagcugauca gagaagucaa ggucaucaca cugaagagca agcuggucag cgacuucaga 3540

aaggacuucc aguucuacaa ggucagagaa aucaacaacu accaccacgc acacgacgca 3600

uaccugaacg cagucgucgg aacagcacug aucaagaagu acccgaagcu ggaaagcgaa 3660

uucgucuacg gagacuacaa ggucuacgac gucagaaaga ugaucgcaaa gagcgaacag 3720

gaaaucggaa aggcaacagc aaaguacuuc uucuacagca acaucaugaa cuucuucaag 3780

acagaaauca cacuggcaaa cggagaaauc agaaagagac cgcugaucga aacaaacgga 3840

gaaacaggag aaaucgucug ggacaaggga agagacuucg caacagucag aaagguccug 3900

agcaugccgc aggucaacau cgucaagaag acagaagucc agacaggagg auucagcaag 3960

gaaagcaucc ugccgaagag aaacagcgac aagcugaucg caagaaagaa ggacugggac 4020

ccgaagaagu acggaggauu cgacagcccg acagucgcau acagcguccu ggucgucgca 4080

aaggucgaaa agggaaagag caagaagcug aagagcguca aggaacugcu gggaaucaca 4140

aucauggaaa gaagcagcuu cgaaaagaac ccgaucgacu uccuggaagc aaagggauac 4200

aaggaaguca agaaggaccu gaucaucaag cugccgaagu acagccuguu cgaacuggaa 4260

aacggaagaa agagaaugcu ggcaagcgca ggagaacugc agaagggaaa cgaacuggca 4320

cugccgagca aguacgucaa cuuccuguac cuggcaagcc acuacgaaaa gcugaaggga 4380

agcccggaag acaacgaaca gaagcagcug uucgucgaac agcacaagca cuaccuggac 4440

gaaaucaucg aacagaucag cgaauucagc aagagaguca uccuggcaga cgcaaaccug 4500

gacaaggucc ugagcgcaua caacaagcac agagacaagc cgaucagaga acaggcagaa 4560

aacaucaucc accuguucac acugacaaac cugggagcac cggcagcauu caaguacuuc 4620

gacacaacaa ucgacagaaa gagauacaca agcacaaagg aaguccugga cgcaacacug 4680

auccaccaga gcaucacagg acuguacgaa acaagaaucg aucugagcca gcugggagga 4740

gacagcggag gaagcacaaa ccugagcgac aucaucgaaa aggaaacagg aaagcagcug 4800

gucauccagg aaagcauccu gaugcugccg gaagaagucg aagaagucau cggaaacaag 4860

ccggaaagcg acauccuggu ccacacagca uacgacgaaa gcacagacga aaacgucaug 4920

cugcugacaa gcgacgcacc ggaauacaag ccgugggcac uggucaucca ggacagcaac 4980

ggagaaaaca agaucaagau gcugagcgga ggaagcccga agaagaagag aaaggucuaa 5040

uag 5043

<210> 370
<211> 5058
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 370
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacggc 600

ggcggcggca gcggcggcgg cggcagcggc ggcggcggca gcggcggcgg cggcagcgac 660

aagaaguaca gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca 720

gacgaauaca aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc 780

aucaagaaga accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca 840

agacugaaga gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug 900

caggaaaucu ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa 960

gaaagcuucc uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc 1020

gucgacgaag ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug 1080

gucgacagca cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc 1140

aaguucagag gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac 1200

aagcuguuca uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac 1260

gcaagcggag ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug 1320

gaaaaccuga ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc 1380

gcacugagcc ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca 1440

aagcugcagc ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc 1500

ggagaccagu acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug 1560

agcgacaucc ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc 1620

aagagauacg acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag 1680

cugccggaaa aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac 1740

aucgacggag gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag 1800

auggacggaa cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag 1860

agaacauucg acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc 1920

cugagaagac aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag 1980

auccugacau ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc 2040

gcauggauga caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc 2100

gacaagggag caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug 2160

ccgaacgaaa agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac 2220

gaacugacaa aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga 2280

gaacagaaga aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag 2340

cagcugaagg aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga 2400

gucgaagaca gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag 2460

gacaaggacu uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca 2520

cugacacugu ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug 2580

uucgacgaca aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug 2640

agcagaaagc ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc 2700

cugaagagcg acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug 2760

acauucaagg aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa 2820

cacaucgcaa accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag 2880

gucgucgacg aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa 2940

auggcaagag aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag 3000

agaaucgaag aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa 3060

aacacacagc ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug 3120

uacgucgacc aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc 3180

ccgcagagcu uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag 3240

aacagaggaa agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac 3300

uggagacagc ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag 3360

gcagagagag gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc 3420

gaaacaagac agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag 3480

uacgacgaaa acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug 3540

gucagcgacu ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac 3600

cacgcacacg acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg 3660

aagcuggaaa gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc 3720

gcaaagagcg aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc 3780

augaacuucu ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug 3840

aucgaaacaa acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca 3900

gucagaaagg uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca 3960

ggaggauuca gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga 4020

aagaaggacu gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc 4080

guccuggucg ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa 4140

cugcugggaa ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug 4200

gaagcaaagg gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc 4260

cuguucgaac uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag 4320

ggaaacgaac uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac 4380

gaaaagcuga agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac 4440

aagcacuacc uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug 4500

gcagacgcaa accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc 4560

agagaacagg cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca 4620

gcauucaagu acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc 4680

cuggacgcaa cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug 4740

agccagcugg gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa 4800

acaggaaagc agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa 4860

gucaucggaa acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca 4920

gacgaaaacg ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc 4980

auccaggaca gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag 5040

aagagaaagg ucuaauag 5058

<210> 371
<211> 5055
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 371

auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacggc 600

ggcggcggca gcggaggagg aggaagcgaa gccgccgcca aagaggccgc cgccaaggac 660

aagaaguaca gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca 720

gacgaauaca aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc 780

aucaagaaga accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca 840

agacugaaga gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug 900

caggaaaucu ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa 960

gaaagcuucc uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc 1020

gucgacgaag ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug 1080

gucgacagca cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc 1140

aaguucagag gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac 1200

aagcuguuca uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac 1260

gcaagcggag ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug 1320

gaaaaccuga ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc 1380

gcacugagcc ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca 1440

aagcugcagc ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc 1500

ggagaccagu acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug 1560

agcgacaucc ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc 1620

aagagauacg acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag 1680

cugccggaaa aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac 1740

aucgacggag gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag 1800

auggacggaa cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag 1860

agaacauucg acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc 1920

cugagaagac aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag 1980

auccugacau ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc 2040

gcauggauga caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc 2100

gacaagggag caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug 2160

ccgaacgaaa agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac 2220

gaacugacaa aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga 2280

gaacagaaga aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag 2340

cagcugaagg aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga 2400

gucgaagaca gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag 2460

gacaaggacu uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca 2520

cugacacugu ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug 2580

uucgacgaca aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug 2640

agcagaaagc ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc 2700

cugaagagcg acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug 2760

acauucaagg aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa 2820

cacaucgcaa accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag 2880

gucgucgacg aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa 2940

auggcaagag aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag 3000

agaaucgaag aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa 3060

aacacacagc ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug 3120

uacgucgacc aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc 3180

ccgcagagcu uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag 3240

aacagaggaa agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac 3300

uggagacagc ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag 3360

gcagagagag gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc 3420

gaaacaagac agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag 3480

uacgacgaaa acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug 3540

gucagcgacu ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac 3600

cacgcacacg acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg 3660

aagcuggaaa gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc 3720

gcaaagagcg aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc 3780

augaacuucu ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug 3840

aucgaaacaa acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca 3900

gucagaaagg uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca 3960

ggaggauuca gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga 4020

aagaaggacu gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc 4080

guccuggucg ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa 4140

cugcugggaa ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug 4200

gaagcaaagg gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc 4260

cuguucgaac uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag 4320

ggaaacgaac uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac 4380

gaaaagcuga agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac 4440

aagcacuacc uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug 4500

gcagacgcaa accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc 4560

agagaacagg cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca 4620

gcauucaagu acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc 4680

cuggacgcaa cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug 4740

agccagcugg gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa 4800

acaggaaagc agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa 4860

gucaucggaa acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca 4920

gacgaaaacg ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc 4980

auccaggaca gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag 5040

aagagaaagg ucuaa 5055

<210> 372
<211> 5055
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 372
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacggc 600

ggcggcggaa gcgaggccgc ugccaaaggc ggaggaggaa gcggaggagg cggcagcgac 660

aagaaguaca gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca 720

gacgaauaca aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc 780

aucaagaaga accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca 840

agacugaaga gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug 900

caggaaaucu ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa 960

gaaagcuucc uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc 1020

gucgacgaag ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug 1080

gucgacagca cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc 1140

aaguucagag gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac 1200

aagcuguuca uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac 1260

gcaagcggag ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug 1320

gaaaaccuga ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc 1380

gcacugagcc ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca 1440

aagcugcagc ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc 1500

ggagaccagu acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug 1560

agcgacaucc ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc 1620

aagagauacg acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag 1680

cugccggaaa aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac 1740

aucgacggag gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag 1800

auggacggaa cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag 1860

agaacauucg acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc 1920

cugagaagac aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag 1980

auccugacau ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc 2040

gcauggauga caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc 2100

gacaagggag caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug 2160

ccgaacgaaa agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac 2220

gaacugacaa aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga 2280

gaacagaaga aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag 2340

cagcugaagg aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga 2400

gucgaagaca gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag 2460

gacaaggacu uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca 2520

cugacacugu ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug 2580

uucgacgaca aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug 2640

agcagaaagc ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc 2700

cugaagagcg acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug 2760

acauucaagg aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa 2820

cacaucgcaa accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag 2880

gucgucgacg aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa 2940

auggcaagag aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag 3000

agaaucgaag aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa 3060

aacacacagc ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug 3120

uacgucgacc aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc 3180

ccgcagagcu uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag 3240

aacagaggaa agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac 3300

uggagacagc ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag 3360

gcagagagag gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc 3420

gaaacaagac agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag 3480

uacgacgaaa acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug 3540

gucagcgacu ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac 3600

cacgcacacg acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg 3660

aagcuggaaa gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc 3720

gcaaagagcg aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc 3780

augaacuucu ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug 3840

aucgaaacaa acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca 3900

gucagaaagg uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca 3960

ggaggauuca gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga 4020

aagaaggacu gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc 4080

guccuggucg ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa 4140

cugcugggaa ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug 4200

gaagcaaagg gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc 4260

cuguucgaac uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag 4320

ggaaacgaac uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac 4380

gaaaagcuga agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac 4440

aagcacuacc uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug 4500

gcagacgcaa accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc 4560

agagaacagg cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca 4620

gcauucaagu acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc 4680

cuggacgcaa cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug 4740

agccagcugg gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa 4800

acaggaaagc agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa 4860

gucaucggaa acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca 4920

gacgaaaacg ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc 4980

auccaggaca gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag 5040

aagagaaagg ucuaa 5055

<210> 373
<211> 5055
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 373
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag 600

gccgccgcca aggaggccgc cgccaaggag gccgccgcca aggaggccgc cgccaaggac 660

aagaaguaca gcaucggacu ggccaucgga acaaacagcg ucggaugggc agucaucaca 720

gacgaauaca aggucccgag caagaaguuc aagguccugg gaaacacaga cagacacagc 780

aucaagaaga accugaucgg agcacugcug uucgacagcg gagaaacagc agaagcaaca 840

agacugaaga gaacagcaag aagaagauac acaagaagaa agaacagaau cugcuaccug 900

caggaaaucu ucagcaacga aauggcaaag gucgacgaca gcuucuucca cagacuggaa 960

gaaagcuucc uggucgaaga agacaagaag cacgaaagac acccgaucuu cggaaacauc 1020

gucgacgaag ucgcauacca cgaaaaguac ccgacaaucu accaccugag aaagaagcug 1080

gucgacagca cagacaaggc agaccugaga cugaucuacc uggcacuggc acacaugauc 1140

aaguucagag gacacuuccu gaucgaagga gaccugaacc cggacaacag cgacgucgac 1200

aagcuguuca uccagcuggu ccagacauac aaccagcugu ucgaagaaaa cccgaucaac 1260

gcaagcggag ucgacgcaaa ggcaauccug agcgcaagac ugagcaagag cagaagacug 1320

gaaaaccuga ucgcacagcu gccgggagaa aagaagaacg gacuguucgg aaaccugauc 1380

gcacugagcc ugggacugac accgaacuuc aagagcaacu ucgaccuggc agaagacgca 1440

aagcugcagc ugagcaagga cacauacgac gacgaccugg acaaccugcu ggcacagauc 1500

ggagaccagu acgcagaccu guuccuggca gcaaagaacc ugagcgacgc aauccugcug 1560

agcgacaucc ugagagucaa cacagaaauc acaaaggcac cgcugagcgc aagcaugauc 1620

aagagauacg acgaacacca ccaggaccug acacugcuga aggcacuggu cagacagcag 1680

cugccggaaa aguacaagga aaucuucuuc gaccagagca agaacggaua cgcaggauac 1740

aucgacggag gagcaagcca ggaagaauuc uacaaguuca ucaagccgau ccuggaaaag 1800

auggacggaa cagaagaacu gcuggucaag cugaacagag aagaccugcu gagaaagcag 1860

agaacauucg acaacggaag caucccgcac cagauccacc ugggagaacu gcacgcaauc 1920

cugagaagac aggaagacuu cuacccguuc cugaaggaca acagagaaaa gaucgaaaag 1980

auccugacau ucagaauccc guacuacguc ggaccgcugg caagaggaaa cagcagauuc 2040

gcauggauga caagaaagag cgaagaaaca aucacaccgu ggaacuucga agaagucguc 2100

gacaagggag caagcgcaca gagcuucauc gaaagaauga caaacuucga caagaaccug 2160

ccgaacgaaa agguccugcc gaagcacagc cugcuguacg aauacuucac agucuacaac 2220

gaacugacaa aggucaagua cgucacagaa ggaaugagaa agccggcauu ccugagcgga 2280

gaacagaaga aggcaaucgu cgaccugcug uucaagacaa acagaaaggu cacagucaag 2340

cagcugaagg aagacuacuu caagaagauc gaaugcuucg acagcgucga aaucagcgga 2400

gucgaagaca gauucaacgc aagccuggga acauaccacg accugcugaa gaucaucaag 2460

gacaaggacu uccuggacaa cgaagaaaac gaagacaucc uggaagacau cguccugaca 2520

cugacacugu ucgaagacag agaaaugauc gaagaaagac ugaagacaua cgcacaccug 2580

uucgacgaca aggucaugaa gcagcugaag agaagaagau acacaggaug gggaagacug 2640

agcagaaagc ugaucaacgg aaucagagac aagcagagcg gaaagacaau ccuggacuuc 2700

cugaagagcg acggauucgc aaacagaaac uucaugcagc ugauccacga cgacagccug 2760

acauucaagg aagacaucca gaaggcacag gucagcggac agggagacag ccugcacgaa 2820

cacaucgcaa accuggcagg aagcccggca aucaagaagg gaauccugca gacagucaag 2880

gucgucgacg aacuggucaa ggucauggga agacacaagc cggaaaacau cgucaucgaa 2940

auggcaagag aaaaccagac aacacagaag ggacagaaga acagcagaga aagaaugaag 3000

agaaucgaag aaggaaucaa ggaacuggga agccagaucc ugaaggaaca cccggucgaa 3060

aacacacagc ugcagaacga aaagcuguac cuguacuacc ugcagaacgg aagagacaug 3120

uacgucgacc aggaacugga caucaacaga cugagcgacu acgacgucga ccacaucguc 3180

ccgcagagcu uccugaagga cgacagcauc gacaacaagg uccugacaag aagcgacaag 3240

aacagaggaa agagcgacaa cgucccgagc gaagaagucg ucaagaagau gaagaacuac 3300

uggagacagc ugcugaacgc aaagcugauc acacagagaa aguucgacaa ccugacaaag 3360

gcagagagag gaggacugag cgaacuggac aaggcaggau ucaucaagag acagcugguc 3420

gaaacaagac agaucacaaa gcacgucgca cagauccugg acagcagaau gaacacaaag 3480

uacgacgaaa acgacaagcu gaucagagaa gucaagguca ucacacugaa gagcaagcug 3540

gucagcgacu ucagaaagga cuuccaguuc uacaagguca gagaaaucaa caacuaccac 3600

cacgcacacg acgcauaccu gaacgcaguc gucggaacag cacugaucaa gaaguacccg 3660

aagcuggaaa gcgaauucgu cuacggagac uacaaggucu acgacgucag aaagaugauc 3720

gcaaagagcg aacaggaaau cggaaaggca acagcaaagu acuucuucua cagcaacauc 3780

augaacuucu ucaagacaga aaucacacug gcaaacggag aaaucagaaa gagaccgcug 3840

aucgaaacaa acggagaaac aggagaaauc gucugggaca agggaagaga cuucgcaaca 3900

gucagaaagg uccugagcau gccgcagguc aacaucguca agaagacaga aguccagaca 3960

ggaggauuca gcaaggaaag cauccugccg aagagaaaca gcgacaagcu gaucgcaaga 4020

aagaaggacu gggacccgaa gaaguacgga ggauucgaca gcccgacagu cgcauacagc 4080

guccuggucg ucgcaaaggu cgaaaaggga aagagcaaga agcugaagag cgucaaggaa 4140

cugcugggaa ucacaaucau ggaaagaagc agcuucgaaa agaacccgau cgacuuccug 4200

gaagcaaagg gauacaagga agucaagaag gaccugauca ucaagcugcc gaaguacagc 4260

cuguucgaac uggaaaacgg aagaaagaga augcuggcaa gcgcaggaga acugcagaag 4320

ggaaacgaac uggcacugcc gagcaaguac gucaacuucc uguaccuggc aagccacuac 4380

gaaaagcuga agggaagccc ggaagacaac gaacagaagc agcuguucgu cgaacagcac 4440

aagcacuacc uggacgaaau caucgaacag aucagcgaau ucagcaagag agucauccug 4500

gcagacgcaa accuggacaa gguccugagc gcauacaaca agcacagaga caagccgauc 4560

agagaacagg cagaaaacau cauccaccug uucacacuga caaaccuggg agcaccggca 4620

gcauucaagu acuucgacac aacaaucgac agaaagagau acacaagcac aaaggaaguc 4680

cuggacgcaa cacugaucca ccagagcauc acaggacugu acgaaacaag aaucgaucug 4740

agccagcugg gaggagacag cggaggaagc acaaaccuga gcgacaucau cgaaaaggaa 4800

acaggaaagc agcuggucau ccaggaaagc auccugaugc ugccggaaga agucgaagaa 4860

gucaucggaa acaagccgga aagcgacauc cugguccaca cagcauacga cgaaagcaca 4920

gacgaaaacg ucaugcugcu gacaagcgac gcaccggaau acaagccgug ggcacugguc 4980

auccaggaca gcaacggaga aaacaagauc aagaugcuga gcggaggaag cccgaagaag 5040

aagagaaagg ucuaa 5055

<210> 374
<211> 5070
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 374
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgga 600

ggaggaggaa gcgaagccgc cgccaaagag gccgccgcca agggaggcgg cggcagcgag 660

gccgccgcca aggacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga 720

ugggcaguca ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac 780

acagacagac acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa 840

acagcagaag caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac 900

agaaucugcu accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc 960

uuccacagac uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg 1020

aucuucggaa acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac 1080

cugagaaaga agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca 1140

cuggcacaca ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac 1200

aacagcgacg ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa 1260

gaaaacccga ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc 1320

aagagcagaa gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug 1380

uucggaaacc ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac 1440

cuggcagaag acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac 1500

cugcuggcac agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc 1560

gacgcaaucc ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug 1620

agcgcaagca ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca 1680

cuggucagac agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac 1740

ggauacgcag gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag 1800

ccgauccugg aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac 1860

cugcugagaa agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga 1920

gaacugcacg caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga 1980

gaaaagaucg aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga 2040

ggaaacagca gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac 2100

uucgaagaag ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac 2160

uucgacaaga accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac 2220

uucacagucu acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg 2280

gcauuccuga gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga 2340

aaggucacag ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc 2400

gucgaaauca gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug 2460

cugaagauca ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa 2520

gacaucgucc ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag 2580

acauacgcac accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca 2640

ggaugggaa gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag 2700

acaauccugg acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc 2760

cacgacgaca gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga 2820

gacagccugc acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc 2880

cugcagacag ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa 2940

aacaucguca ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc 3000

agagaaagaa ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag 3060

gaacacccgg ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag 3120

aacggaagag acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac 3180

gucgaccaca ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug 3240

acaagaagcg acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag 3300

aagaugaaga acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc 3360

gacaaccuga caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc 3420

aagagacagc uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc 3480

agaaugaaca caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca 3540

cugaagagca agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa 3600

aucaacaacu accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug 3660

aucaagaagu acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac 3720

gucagaaaga ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc 3780

uucuacagca acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc 3840

agaaagagac cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga 3900

agagacuucg caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag 3960

acagaagucc agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac 4020

aagcugaucg caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg 4080

acagucgcau acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug 4140

aagagcguca aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac 4200

ccgaucgacu uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag 4260

cugccgaagu acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca 4320

ggagaacugc agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac 4380

cuggcaagcc acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug 4440

uucgucgaac agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc 4500

aagagaguca uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac 4560

agagacaagc cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac 4620

cugggagcac cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca 4680

agcacaaagg aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa 4740

acaagaaucg aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac 4800

aucaucgaaa aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg 4860

gaagaagucg aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca 4920

uacgacgaaa gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag 4980

ccgugggcac uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga 5040

ggaagcccga agaagaagag aaaggucuaa 5070

<210> 375
<211> 5070
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 375
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgag 600

gccgccgcca aggaggccgc cgccaagggc ggcggcggca gcggcggcgg cggcagcggc 660

ggcggcggca gcgacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga 720

ugggcaguca ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac 780

acagacagac acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa 840

acagcagaag caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac 900

agaaucugcu accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc 960

uuccacagac uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg 1020

aucuucggaa acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac 1080

cugagaaaga agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca 1140

cuggcacaca ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac 1200

aacagcgacg ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa 1260

gaaaacccga ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc 1320

aagagcagaa gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug 1380

uucggaaacc ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac 1440

cuggcagaag acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac 1500

cugcuggcac agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc 1560

gacgcaaucc ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug 1620

agcgcaagca ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca 1680

cuggucagac agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac 1740

ggauacgcag gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag 1800

ccgauccugg aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac 1860

cugcugagaa agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga 1920

gaacugcacg caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga 1980

gaaaagaucg aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga 2040

ggaaacagca gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac 2100

uucgaagaag ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac 2160

uucgacaaga accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac 2220

uucacagucu acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg 2280

gcauuccuga gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga 2340

aaggucacag ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc 2400

gucgaaauca gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug 2460

cugaagauca ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa 2520

gacaucgucc ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag 2580

acauacgcac accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca 2640

ggaugggaa gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag 2700

acaauccugg acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc 2760

cacgacgaca gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga 2820

gacagccugc acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc 2880

cugcagacag ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa 2940

aacaucguca ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc 3000

agagaaagaa ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag 3060

gaacacccgg ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag 3120

aacggaagag acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac 3180

gucgaccaca ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug 3240

acaagaagcg acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag 3300

aagaugaaga acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc 3360

gacaaccuga caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc 3420

aagagacagc uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc 3480

agaaugaaca caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca 3540

cugaagagca agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa 3600

aucaacaacu accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug 3660

aucaagaagu acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac 3720

gucagaaaga ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc 3780

uucuacagca acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc 3840

agaaagagac cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga 3900

agagacuucg caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag 3960

acagaagucc agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac 4020

aagcugaucg caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg 4080

acagucgcau acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug 4140

aagagcguca aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac 4200

ccgaucgacu uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag 4260

cugccgaagu acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca 4320

ggagaacugc agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac 4380

cuggcaagcc acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug 4440

uucgucgaac agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc 4500

aagagaguca uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac 4560

agagacaagc cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac 4620

cugggagcac cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca 4680

agcacaaagg aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa 4740

acaagaaucg aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac 4800

aucaucgaaa aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg 4860

gaagaagucg aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca 4920

uacgacgaaa gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag 4980

ccgugggcac uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga 5040

ggaagcccga agaagaagag aaaggucuaa 5070

<210> 376
<211> 5070
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 376

auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccuggucccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccgуggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgaa 600

gcagcagcaa aggaagcagc agcaaaggga ggaggaggaa gcggaggagg aggaagcgaa 660

gcagcagcaa aggacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga 720

ugggcaguca ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac 780

acagacagac acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa 840

acagcagaag caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac 900

agaaucugcu accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc 960

uuccacagac uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg 1020

aucuucggaa acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac 1080

cugagaaaga agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca 1140

cuggcacaca ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac 1200

aacagcgacg ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa 1260

gaaaacccga ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc 1320

aagagcagaa gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug 1380

uucggaaacc ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac 1440

cuggcagaag acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac 1500

cugcuggcac agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc 1560

gacgcaaucc ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug 1620

agcgcaagca ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca 1680

cuggucagac agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac 1740

ggauacgcag gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag 1800

ccgauccugg aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac 1860

cugcugagaa agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga 1920

gaacugcacg caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga 1980

gaaaagaucg aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga 2040

ggaaacagca gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac 2100

uucgaagaag ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac 2160

uucgacaaga accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac 2220

uucacagucu acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg 2280

gcauuccuga gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga 2340

aaggucacag ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc 2400

gucgaaauca gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug 2460

cugaagauca ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa 2520

gacaucgucc ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag 2580

acauacgcac accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca 2640

ggaugggga gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag 2700

acaauccugg acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc 2760

cacgacgaca gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga 2820

gacagccugc acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc 2880

cugcagacag ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa 2940

aacaucguca ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc 3000

agagaaagaa ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag 3060

gaacacccgg ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag 3120

aacggaagag acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac 3180

gucgaccaca ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug 3240

acaagaagcg acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag 3300

aagaugaaga acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc 3360

gacaaccuga caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc 3420

aagagacagc uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc 3480

agaaugaaca caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca 3540

cugaagagca agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa 3600

aucaacaacu accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug 3660

aucaagaagu acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac 3720

gucagaaaga ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc 3780

uucuacagca acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc 3840

agaaagagac cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga 3900

agagacuucg caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag 3960

acagaagucc agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac 4020

aagcugaucg caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg 4080

acagucgcau acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug 4140

aagagcguca aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac 4200

ccgaucgacu uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag 4260

cugccgaagu acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca 4320

ggagaacugc agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac 4380

cuggcaagcc acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug 4440

uucgucgaac agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc 4500

aagagaguca uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac 4560

agagacaagc cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac 4620

cugggagcac cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca 4680

agcacaaagg aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa 4740

acaagaaucg aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac 4800

aucaucgaaa aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg 4860

gaagaagucg aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca 4920

uacgacgaaa gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag 4980

ccgugggcac uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga 5040

ggaagcccga agaagaagag aaaggucuaa 5070

<210> 377
<211> 5070
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 377
auggaagcaa gcccggcaag cggaccgaga caccugaugg acccgcacau cuucacaagc 60

aacuucaaca acggaaucgg aagacacaag acauaccugu gcuacgaagu cgaaagacug 120

gacaacggaa caagcgucaa gauggaccag cacagaggau uccugcacaa ccaggcaaag 180

aaccugcugu gcggauucua cggaagacac gcagaacuga gauuccugga ccugguccg 240

agccugcagc uggacccggc acagaucuac agagucacau gguucaucag cuggagcccg 300

ugcuucagcu ggggaugcgc aggagaaguc agagcauuuc ugcaggaaaa cacacacguc 360

agacugagaa ucuucgcagc aagaaucuac gacuacgacc cgcuguacaa ggaagcacug 420

cagaugcuga gagacgcagg agcacagguc agcaucauga cauacgacga auucaagcac 480

ugcugggaca cauucgucga ccaccaggga ugcccguucc agccguggga cggacuggac 540

gaacacagcc aggcacugag cggaagacug agagcaaucc ugcagaacca gggaaacgaa 600

gcagcagcaa aggaagcagc agcaaaggaa gcagcagcaa agggaggagg aggaagcgga 660

ggaggaggaa gcgacaagaa guacagcauc ggacuggcca ucggaacaaa cagcgucgga 720

ugggcaguca ucacagacga auacaagguc ccgagcaaga aguucaaggu ccugggaaac 780

acagacagac acagcaucaa gaagaaccug aucggagcac ugcuguucga cagcggagaa 840

acagcagaag caacaagacu gaagagaaca gcaagaagaa gauacacaag aagaaagaac 900

agaaucugcu accugcagga aaucuucagc aacgaaaugg caaaggucga cgacagcuuc 960

uuccacagac uggaagaaag cuuccugguc gaagaagaca agaagcacga aagacacccg 1020

aucuucggaa acaucgucga cgaagucgca uaccacgaaa aguacccgac aaucuaccac 1080

cugagaaaga agcuggucga cagcacagac aaggcagacc ugagacugau cuaccuggca 1140

cuggcacaca ugaucaaguu cagaggacac uuccugaucg aaggagaccu gaacccggac 1200

aacagcgacg ucgacaagcu guucauccag cugguccaga cauacaacca gcuguucgaa 1260

gaaaacccga ucaacgcaag cggagucgac gcaaaggcaa uccugagcgc aagacugagc 1320

aagagcagaa gacuggaaaa ccugaucgca cagcugccgg gagaaaagaa gaacggacug 1380

uucggaaacc ugaucgcacu gagccuggga cugacaccga acuucaagag caacuucgac 1440

cuggcagaag acgcaaagcu gcagcugagc aaggacacau acgacgacga ccuggacaac 1500

cugcuggcac agaucggaga ccaguacgca gaccuguucc uggcagcaaa gaaccugagc 1560

gacgcaaucc ugcugagcga cauccugaga gucaacacag aaaucacaaa ggcaccgcug 1620

agcgcaagca ugaucaagag auacgacgaa caccaccagg accugacacu gcugaaggca 1680

cuggucagac agcagcugcc ggaaaaguac aaggaaaucu ucuucgacca gagcaagaac 1740

ggauacgcag gauacaucga cggaggagca agccaggaag aauucuacaa guucaucaag 1800

ccgauccugg aaaagaugga cggaacagaa gaacugcugg ucaagcugaa cagagaagac 1860

cugcugagaa agcagagaac auucgacaac ggaagcaucc cgcaccagau ccaccuggga 1920

gaacugcacg caauccugag aagacaggaa gacuucuacc cguuccugaa ggacaacaga 1980

gaaaagaucg aaaagauccu gacauucaga aucccguacu acgucggacc gcuggcaaga 2040

ggaaacagca gauucgcaug gaugacaaga aagagcgaag aaacaaucac accguggaac 2100

uucgaagaag ucgucgacaa gggagcaagc gcacagagcu ucaucgaaag aaugacaaac 2160

uucgacaaga accugccgaa cgaaaagguc cugccgaagc acagccugcu guacgaauac 2220

uucacagucu acaacgaacu gacaaagguc aaguacguca cagaaggaau gagaaagccg 2280

gcauuccuga gcggagaaca gaagaaggca aucgucgacc ugcuguucaa gacaaacaga 2340

aaggucacag ucaagcagcu gaaggaagac uacuucaaga agaucgaaug cuucgacagc 2400

gucgaaauca gcggagucga agacagauuc aacgcaagcc ugggaacaua ccacgaccug 2460

cugaagauca ucaaggacaa ggacuuccug gacaacgaag aaaacgaaga cauccuggaa 2520

gacaucgucc ugacacugac acuguucgaa gacagagaaa ugaucgaaga aagacugaag 2580

acauacgcac accuguucga cgacaagguc augaagcagc ugaagagaag aagauacaca 2640

ggaugggggaa gacugagcag aaagcugauc aacggaauca gagacaagca gagcggaaag 2700

acaauccugg acuuccugaa gagcgacgga uucgcaaaca gaaacuucau gcagcugauc 2760

cacgacgaca gccugacauu caaggaagac auccagaagg cacaggucag cggacaggga 2820

gacagccugc acgaacacau cgcaaaccug gcaggaagcc cggcaaucaa gaagggaauc 2880

cugcagacag ucaaggucgu cgacgaacug gucaagguca ugggaagaca caagccggaa 2940

aacaucguca ucgaaauggc aagagaaaac cagacaacac agaagggaca gaagaacagc 3000

agagaaagaa ugaagagaau cgaagaagga aucaaggaac ugggaagcca gauccugaag 3060

gaacacccgg ucgaaaacac acagcugcag aacgaaaagc uguaccugua cuaccugcag 3120

aacggaagag acauguacgu cgaccaggaa cuggacauca acagacugag cgacuacgac 3180

gucgaccaca ucgucccgca gagcuuccug aaggacgaca gcaucgacaa caagguccug 3240

acaagaagcg acaagaacag aggaaagagc gacaacgucc cgagcgaaga agucgucaag 3300

aagaugaaga acuacuggag acagcugcug aacgcaaagc ugaucacaca gagaaaguuc 3360

gacaaccuga caaaggcaga gagaggagga cugagcgaac uggacaaggc aggauucauc 3420

aagagacagc uggucgaaac aagacagauc acaaagcacg ucgcacagau ccuggacagc 3480

agaaugaaca caaaguacga cgaaaacgac aagcugauca gagaagucaa ggucaucaca 3540

cugaagagca agcuggucag cgacuucaga aaggacuucc aguucuacaa ggucagagaa 3600

aucaacaacu accaccacgc acacgacgca uaccugaacg cagucgucgg aacagcacug 3660

aucaagaagu acccgaagcu ggaaagcgaa uucgucuacg gagacuacaa ggucuacgac 3720

gucagaaaga ugaucgcaaa gagcgaacag gaaaucggaa aggcaacagc aaaguacuuc 3780

uucuacagca acaucaugaa cuucuucaag acagaaauca cacuggcaaa cggagaaauc 3840

agaaagagac cgcugaucga aacaaacgga gaaacaggag aaaucgucug ggacaaggga 3900

agagacuucg caacagucag aaagguccug agcaugccgc aggucaacau cgucaagaag 3960

acagaagucc agacaggagg auucagcaag gaaagcaucc ugccgaagag aaacagcgac 4020

aagcugaucg caagaaagaa ggacugggac ccgaagaagu acggaggauu cgacagcccg 4080

acagucgcau acagcguccu ggucgucgca aaggucgaaa agggaaagag caagaagcug 4140

aagagcguca aggaacugcu gggaaucaca aucauggaaa gaagcagcuu cgaaaagaac 4200

ccgaucgacu uccuggaagc aaagggauac aaggaaguca agaaggaccu gaucaucaag 4260

cugccgaagu acagccuguu cgaacuggaa aacggaagaa agagaaugcu ggcaagcgca 4320

ggagaacugc agaagggaaa cgaacuggca cugccgagca aguacgucaa cuuccuguac 4380

cuggcaagcc acuacgaaaa gcugaaggga agcccggaag acaacgaaca gaagcagcug 4440

uucgucgaac agcacaagca cuaccuggac gaaaucaucg aacagaucag cgaauucagc 4500

aagagaguca uccuggcaga cgcaaaccug gacaaggucc ugagcgcaua caacaagcac 4560

agagacaagc cgaucagaga acaggcagaa aacaucaucc accuguucac acugacaaac 4620

cugggagcac cggcagcauu caaguacuuc gacacaacaa ucgacagaaa gagauacaca 4680

agcacaaagg aaguccugga cgcaacacug auccaccaga gcaucacagg acuguacgaa 4740

acaagaaucg aucugagcca gcugggagga gacagcggag gaagcacaaa ccugagcgac 4800

aucaucgaaa aggaaacagg aaagcagcug gucauccagg aaagcauccu gaugcugccg 4860

gaagaagucg aagaagucau cggaaacaag ccggaaagcg acauccuggu ccacacagca 4920

uacgacgaaa gcacagacga aaacgucaug cugcugacaa gcgacgcacc ggaauacaag 4980

ccgugggcac uggucaucca ggacagcaac ggagaaaaca agaucaagau gcugagcgga 5040

ggaagcccga agaagaagag aaaggucuaa 5070

<210> 378

<400> 378
000

<210> 379

<400> 379
000

<210> 380

<400> 380
000

<210> 381

<400> 381
000

<210> 382

<400> 382
000

<210> 383

<400> 383
000

<210> 384

<400> 384
000

<210> 385

<400> 385
000

<210> 386

<400> 386
000

<210> 387
<211> 1082
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 387

```
Met Ala Ala Phe Lys Pro Asn Pro Ile Asn Tyr Ile Leu Gly Leu Ala
1               5                   10                  15


Ile Gly Ile Ala Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Glu
            20                  25                  30


Glu Asn Pro Ile Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg
        35                  40                  45


Ala Glu Val Pro Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu
    50                  55                  60


Ala Arg Ser Val Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu
65                  70                  75                  80


Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp
                85                  90                  95


Phe Asp Glu Asn Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln
            100                 105                 110


Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser
        115                 120                 125


Ala Val Leu Leu His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg
    130                 135                 140


Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys
145                 150                 155                 160


Gly Val Ala Asn Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr
                165                 170                 175
```

```
Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile
            180                 185                 190

Arg Asn Gln Arg Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu
        195                 200                 205

Gln Ala Glu Leu Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn
    210                 215                 220

Pro His Val Ser Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met
225                 230                 235                 240

Thr Gln Arg Pro Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly
                245                 250                 255

His Cys Thr Phe Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr
            260                 265                 270

Thr Ala Glu Arg Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile
        275                 280                 285

Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr
    290                 295                 300

Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala
305                 310                 315                 320

Arg Lys Leu Leu Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg
                325                 330                 335

Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala
            340                 345                 350

Tyr His Ala Ile Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys
        355                 360                 365

Lys Ser Pro Leu Asn Leu Ser Ser Glu Leu Gln Asp Glu Ile Gly Thr
    370                 375                 380
```

```
Ala Phe Ser Leu Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys
385                 390                 395                 400

Asp Arg Val Gln Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser
                405                 410                 415

Phe Asp Lys Phe Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val
            420                 425                 430

Pro Leu Met Glu Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile
        435                 440                 445

Tyr Gly Asp His Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu
    450                 455                 460

Pro Pro Ile Pro Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala
465                 470                 475                 480

Leu Ser Gln Ala Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly
                485                 490                 495

Ser Pro Ala Arg Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser
            500                 505                 510

Phe Lys Asp Arg Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys
        515                 520                 525

Asp Arg Glu Lys Ala Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe
    530                 535                 540

Val Gly Glu Pro Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu
545                 550                 555                 560

Gln Gln His Gly Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Val
                565                 570                 575
```

```
Arg Leu Asn Glu Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe
            580                 585                 590

Ser Arg Thr Trp Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly
        595                 600                 605

Ser Glu Asn Gln Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn
    610                 615                 620

Gly Lys Asp Asn Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu
625                 630                 635                 640

Thr Ser Arg Phe Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys
                645                 650                 655

Phe Asp Glu Asp Gly Phe Lys Glu Cys Asn Leu Asn Asp Thr Arg Tyr
            660                 665                 670

Val Asn Arg Phe Leu Cys Gln Phe Val Ala Asp His Ile Leu Leu Thr
        675                 680                 685

Gly Lys Gly Lys Arg Arg Val Phe Ala Ser Asn Gly Gln Ile Thr Asn
    690                 695                 700

Leu Leu Arg Gly Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp
705                 710                 715                 720

Arg His His Ala Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala
                725                 730                 735

Met Gln Gln Lys Ile Thr Arg Phe Val Arg Tyr Lys Glu Met Asn Ala
            740                 745                 750

Phe Asp Gly Lys Thr Ile Asp Lys Glu Thr Gly Lys Val Leu His Gln
        755                 760                 765

Lys Thr His Phe Pro Gln Pro Trp Glu Phe Phe Ala Gln Glu Val Met
    770                 775                 780
```

```
Ile Arg Val Phe Gly Lys Pro Asp Gly Lys Pro Glu Phe Glu Glu Ala
785                 790                 795                 800


Asp Thr Pro Glu Lys Leu Arg Thr Leu Leu Ala Glu Lys Leu Ser Ser
                805                 810                 815


Arg Pro Glu Ala Val His Glu Tyr Val Thr Pro Leu Phe Val Ser Arg
            820                 825                 830


Ala Pro Asn Arg Lys Met Ser Gly Ala His Lys Asp Thr Leu Arg Ser
        835                 840                 845


Ala Lys Arg Phe Val Lys His Asn Glu Lys Ile Ser Val Lys Arg Val
    850                 855                 860


Trp Leu Thr Glu Ile Lys Leu Ala Asp Leu Glu Asn Met Val Asn Tyr
865                 870                 875                 880


Lys Asn Gly Arg Glu Ile Glu Leu Tyr Glu Ala Leu Lys Ala Arg Leu
                885                 890                 895


Glu Ala Tyr Gly Gly Asn Ala Lys Gln Ala Phe Asp Pro Lys Asp Asn
            900                 905                 910


Pro Phe Tyr Lys Lys Gly Gly Gln Leu Val Lys Ala Val Arg Val Glu
        915                 920                 925


Lys Thr Gln Glu Ser Gly Val Leu Leu Asn Lys Lys Asn Ala Tyr Thr
    930                 935                 940


Ile Ala Asp Asn Gly Asp Met Val Arg Val Asp Val Phe Cys Lys Val
945                 950                 955                 960


Asp Lys Lys Gly Lys Asn Gln Tyr Phe Ile Val Pro Ile Tyr Ala Trp
                965                 970                 975
```

```
Gln Val Ala Glu Asn Ile Leu Pro Asp Ile Asp Cys Lys Gly Tyr Arg
            980                 985                 990


Ile Asp Asp Ser Tyr Thr Phe Cys  Phe Ser Leu His Lys  Tyr Asp Leu
        995                1000                1005


Ile Ala  Phe Gln Lys Asp Glu  Lys Ser Lys Val Glu  Phe Ala Tyr
    1010                1015                1020


Tyr Ile  Asn Cys Asp Ser Ser  Asn Gly Arg Phe Tyr  Leu Ala Trp
    1025                1030                1035


His Asp  Lys Gly Ser Lys Glu  Gln Gln Phe Arg Ile  Ser Thr Gln
    1040                1045                1050


Asn Leu  Val Leu Ile Gln Lys  Tyr Gln Val Asn Glu  Leu Gly Lys
    1055                1060                1065


Glu Ile  Arg Pro Cys Arg Leu  Lys Lys Arg Pro Pro  Val Arg
    1070                1075                1080


<210>  388
<211>  3243
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  388
gctgctttta agcctaatcc tattaattat attcttggtc ttgctattgg tattgcttct       60

gttggttggg ctatggttga gattgatgag gaggagaatc ctattcgtct tattgatctt      120

ggtgttcgtg tttttgagcg tgctgaggtt cctaagactg gtgattctct tgctatggct      180

cgtcgtcttg ctcgttctgt tcgtcgtctt actcgtcgtc gtgctcatcg tcttcttcgt      240

gctcgtcgtc ttcttaagcg tgagggtgtt cttcaggctg ctgattttga tgagaatggt      300

cttattaagt ctcttcctaa tactccttgg cagcttcgtg ctgctgctct tgatcgtaag      360

cttactcctc ttgagtggtc tgctgttctt cttcatctta ttaagcatcg tggttatctt      420
```

tctcagcgta agaatgaggg tgagactgct gataaggagc ttggtgctct tcttaagggt 480

gttgctaata atgctcatgc tcttcagact ggtgattttc gtactcctgc tgagcttgct 540

cttaataagt ttgagaagga gtctggtcat attcgtaatc agcgtggtga ttattctcat 600

acttttctc gtaaggatct tcaggctgag cttattcttc tttttgagaa gcagaaggag 660

tttggtaatc ctcatgtttc tggtggtctt aaggagggta ttgagactct tcttatgact 720

cagcgtcctg ctctttctgg tgatgctgtt cagaagatgc ttggtcattg tacttttgag 780

cctgctgagc ctaaggctgc taagaatact tatactgctg agcgttttat ttggcttact 840

aagcttaata atcttcgtat tcttgagcag ggttctgagc gtcctcttac tgatactgag 900

cgtgctactc ttatggatga gccttatcgt aagtctaagc ttacttatgc tcaggctcgt 960

aagcttcttg gtcttgagga tactgctttt tttaagggtc ttcgttatgg taaggataat 1020

gctgaggctt ctactcttat ggagatgaag gcttatcatg ctatttctcg tgctcttgag 1080

aaggagggtc ttaaggataa gaagtctcct cttaatcttt cttctgagct tcaggatgag 1140

attggtactg ctttttctct ttttaagact gatgaggata ttactggtcg tcttaaggat 1200

cgtgttcagc ctgagattct tgaggctctt cttaagcata tttcttttga taagtttgtt 1260

cagatttctc ttaaggctct tcgtcgtatt gttcctctta tggagcaggg taagcgttat 1320

gatgaggctt gtgctgagat ttatggtgat cattatggta agaagaatac tgaggagaag 1380

atttatcttc ctcctattcc tgctgatgag attcgtaatc ctgttgttct tcgtgctctt 1440

tctcaggctc gtaaggttat taatggtgtt gttcgtcgtt atggttctcc tgctcgtatt 1500

catattgaga ctgctcgtga ggttggtaag tcttttaagg atcgtaagga gattgagaag 1560

cgtcaggagg agaatcgtaa ggatcgtgag aaggctgctg ctaagtttcg tgagtatttt 1620

cctaattttg ttggtgagcc taagtctaag gatattctta agcttcgtct ttatgagcag 1680

cagcatggta agtgtcttta ttctggtaag gagattaatc ttgttcgtct taatgagaag 1740

ggttatgttg agattgatca tgctcttcct ttttctcgta cttgggatga ttcttttaat 1800

aataaggttc ttgttcttgg ttctgagaat cagaataagg gtaatcagac tccttatgag 1860

tattttaatg gtaaggataa ttctcgtgag tggcaggagt ttaaggctcg tgttgagact 1920

tctcgttttc ctcgttctaa gaagcagcgt attcttcttc agaagtttga tgaggatggt 1980

tttaaggagt gtaatcttaa tgatactcgt tatgttaatc gttttctttg tcagtttgtt 2040

gctgatcata ttcttcttac tggtaagggt aagcgtcgtg tttttgcttc taatggtcag 2100

attactaatc ttcttcgtgg tttttggggt cttcgtaagg ttcgtgctga gaatgatcgt 2160

catcatgctc ttgatgctgt tgttgttgct tgttctactg ttgctatgca gcagaagatt 2220

actcgttttg ttcgttataa ggagatgaat gcttttgatg gtaagactat tgataaggag 2280

actggtaagg ttcttcatca gaagactcat tttcctcagc cttgggagtt ttttgctcag 2340

gaggttatga ttcgtgtttt tggtaagcct gatggtaagc ctgagtttga ggaggctgat 2400

actcctgaga agcttcgtac tcttcttgct gagaagcttt cttctcgtcc tgaggctgtt 2460

catgagtatg ttactcctct ttttgtttct cgtgctccta atcgtaagat gtctggtgct 2520

cataaggata ctcttcgttc tgctaagcgt tttgttaagc ataatgagaa gatttctgtt 2580

aagcgtgttt ggcttactga gattaagctt gctgatcttg agaatatggt taattataag 2640

aatggtcgtg agattgagct ttatgaggct cttaaggctc gtcttgaggc ttatggtggt 2700

aatgctaagc aggcttttga tcctaaggat aatccttttt ataagaaggg tggtcagctt 2760

gttaaggctg ttcgtgttga gaagactcag gagtctggtg ttcttcttaa taagaagaat 2820

gcttatacta ttgctgataa tggtgatatg gttcgtgttg atgtttttttg taaggttgat 2880

aagaagggta agaatcagta ttttattgtt cctatttatg cttggcaggt tgctgagaat 2940

attcttcctg atattgattg taagggttat cgtattgatg attcttatac tttttgtttt 3000

tctcttcata agtatgatct tattgctttt cagaaggatg agaagtctaa ggttgagttt 3060

gcttattata ttaattgtga ttcttctaat ggtcgttttt atcttgcttg gcatgataag 3120

ggttctaagg agcagcagtt tcgtatttct actcagaatc ttgttcttat tcagaagtat 3180

caggttaatg agcttggtaa ggagattcgt ccttgtcgtc ttaagaagcg tcctcctgtt 3240

cgt 3243

<210> 389
<211> 3243

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 389
gccgccttca agcccaaccc catcaactac atcctgggcc tggccatcgg catcgcctcc 60

gtgggctggg ccatggtgga gatcgacgag gaggagaacc ccatccggct gatcgacctg 120

ggcgtgcggg tgttcgagcg ggccgaggtg cccaagaccg gcgactccct ggccatggcc 180

cggcggctgg cccggtccgt gcggcggctg acccggcggc gggcccaccg gctgctgcgg 240

gcccggcggc tgctgaagcg ggagggcgtg ctgcaggccg ccgacttcga cgagaacggc 300

ctgatcaagt ccctgcccaa cacccctgg cagctgcggg ccgccgccct ggaccggaag 360

ctgacccccc tggagtggtc cgccgtgctg ctgcacctga tcaagcaccg gggctacctg 420

tcccagcgga agaacgaggg cgagaccgcc gacaaggagc tgggcgccct gctgaagggc 480

gtggccaaca acgcccacgc cctgcagacc ggcgacttcc ggacccccgc cgagctggcc 540

ctgaacaagt tcgagaagga gtccggccac atccggaacc agcggggcga ctactcccac 600

accttctccc ggaaggacct gcaggccgag ctgatcctgc tgttcgagaa gcagaaggag 660

ttcggcaacc cccacgtgtc cggcggcctg aaggagggca tcgagaccct gctgatgacc 720

cagcggcccg ccctgtccgg cgacgccgtg cagaagatgc tgggccactg caccttcgag 780

cccgccgagc ccaaggccgc caagaacacc tacaccgccg agcggttcat ctggctgacc 840

aagctgaaca acctgcggat cctggagcag ggctccgagc ggcccctgac cgacaccgag 900

cgggccaccc tgatggacga gccctaccgg aagtccaagc tgacctacgc ccaggcccgg 960

aagctgctgg gcctggagga caccgccttc ttcaagggcc tgcggtacgg caaggacaac 1020

gccgaggcct ccaccctgat ggagatgaag gcctaccacg ccatctcccg ggccctggag 1080

aaggagggcc tgaaggacaa gaagtccccc ctgaacctgt cctccgagct gcaggacgag 1140

atcggcaccg ccttctccct gttcaagacc gacgaggaca tcaccggccg gctgaaggac 1200

cgggtgcagc ccgagatcct ggaggccctg ctgaagcaca tctccttcga caagttcgtg 1260

cagatctccc tgaaggccct gcggcggatc gtgcccctga tggagcaggg caagcggtac 1320

gacgaggcct gcgccgagat ctacggcgac cactacggca agaagaacac cgaggagaag 1380

atctacctgc cccccatccc cgccgacgag atccggaacc ccgtggtgct gcgggccctg 1440

tcccaggccc ggaaggtgat caacggcgtg gtgcggcggt acggctcccc cgcccggatc 1500

cacatcgaga ccgcccggga ggtgggcaag tccttcaagg accggaagga gatcgagaag 1560

cggcaggagg agaaccggaa ggaccgggag aaggccgccg ccaagttccg ggagtacttc 1620

cccaacttcg tgggcgagcc caagtccaag gacatcctga agctgcggct gtacgagcag 1680

cagcacggca agtgcctgta ctccggcaag gagatcaacc tggtgcggct gaacgagaag 1740

ggctacgtgg agatcgacca cgccctgccc ttctcccgga cctgggacga ctccttcaac 1800

aacaaggtgc tggtgctggg ctccgagaac cagaacaagg gcaaccagac cccctacgag 1860

tacttcaacg gcaaggacaa ctcccgggag tggcaggagt tcaaggcccg ggtggagacc 1920

tcccggttcc cccggtccaa gaagcagcgg atcctgctgc agaagttcga cgaggacggc 1980

ttcaaggagt gcaacctgaa cgacacccgg tacgtgaacc ggttcctgtg ccagttcgtg 2040

gccgaccaca tcctgctgac cggcaagggc aagcggcggg tgttcgcctc caacggccag 2100

atcaccaacc tgctgcgggg cttctggggc ctgcggaagg tgcgggccga gaacgaccgg 2160

caccacgccc tggacgccgt ggtggtggcc tgctccaccg tggccatgca gcagaagatc 2220

acccggttcg tgcggtacaa ggagatgaac gccttcgacg gcaagaccat cgacaaggag 2280

accggcaagg tgctgcacca gaagacccac ttcccccagc cctgggagtt cttcgcccag 2340

gaggtgatga tccgggtgtt cggcaagccc gacggcaagc ccgagttcga ggaggccgac 2400

acccccgaga agctgcggac cctgctggcc gagaagctgt cctcccggcc cgaggccgtg 2460

cacgagtacg tgacccccct gttcgtgtcc cgggccccca accggaagat gtccggcgcc 2520

cacaaggaca ccctgcggtc cgccaagcgg ttcgtgaagc acaacgagaa gatctccgtg 2580

aagcgggtgt ggctgaccga gatcaagctg gccgacctgg agaacatggt gaactacaag 2640

aacggccggg agatcgagct gtacgaggcc ctgaaggccc ggctggaggc ctacggcggc 2700

aacgccaagc aggccttcga ccccaaggac aacccccttct acaagaaggg cggccagctg 2760

gtgaaggccg tgcgggtgga gaagacccag gagtccggcg tgctgctgaa caagaagaac 2820

gcctacacca tcgccgacaa cggcgacatg gtgcgggtgg acgtgttctg caaggtggac 2880

aagaagggca agaaccagta cttcatcgtg cccatctacg cctggcaggt ggccgagaac 2940

atcctgcccg acatcgactg caagggctac cggatcgacg actcctacac cttctgcttc 3000

tccctgcaca agtacgacct gatcgccttc cagaaggacg agaagtccaa ggtggagttc 3060

gcctactaca tcaactgcga ctcctccaac ggccggttct acctggcctg gcacgacaag 3120

ggctccaagg agcagcagtt ccggatctcc acccagaacc tggtgctgat ccagaagtac 3180

caggtgaacg agctgggcaa ggagatccgg ccctgccggc tgaagaagcg gccccccgtg 3240

cgg 3243

<210> 390
<211> 3243
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 390
gccgccttca agcccaaccc catcaactac atcctgggcc tggccatcgg catcgcctcc 60

gtgggctggg ccatggtgga gatcgacgag gaggagaacc ccatccggct gatcgacctg 120

ggcgtgcggg tgttcgagcg ggccgaggtg cccaagaccg gcgactccct ggccatggcc 180

cggcggctgg cccggtccgt gcggcggctg acccggcggc gggcccaccg gctgctgcgg 240

gcccggcggc tgctgaagcg ggagggcgtg ctgcaggccg ccgacttcga cgagaacggc 300

ctgatcaagt ccctgcccaa cacccctgg cagctgcggg ccgccgccct ggaccggaag 360

ctgacccccc tggagtggtc cgccgtgctg ctgcacctga tcaagcaccg gggctacctg 420

tcccagcgga agaacgaggg cgagaccgcc gacaaggagc tgggcgccct gctgaagggc 480

gtggccaaca acgcccacgc cctgcagacc ggcgacttcc ggacccccgc cgagctggcc 540

ctgaacaagt tcgagaagga gtccggccac atccggaacc agcggggcga ctactcccac 600

accttctccc ggaaggacct gcaggccgag ctgatcctgc tgttcgagaa gcagaaggag 660

ttcggcaacc cccacgtgtc cggcggcctg aaggagggca tcgagaccct gctgatgacc 720

cagcggcccg ccctgtccgg cgacgccgtg cagaagatgc tgggccactg caccttcgag 780

cccgccgagc ccaaggccgc caagaacacc tacaccgccg agcggttcat ctggctgacc 840

aagctgaaca acctgcggat cctggagcag ggctccgagc ggcccctgac cgacaccgag 900

cgggccaccc tgatggacga gccctaccgg aagtccaagc tgacctacgc ccaggcccgg 960

aagctgctgg gcctggagga caccgccttc ttcaagggcc tgcggtacgg caaggacaac 1020

gccgaggcct ccaccctgat ggagatgaag gcctaccacg ccatctcccg ggccctggag 1080

aaggagggcc tgaaggacaa gaagtccccc ctgaacctgt cctccgagct gcaggacgag 1140

atcggcaccg ccttctccct gttcaagacc gacgaggaca tcaccggccg gctgaaggac 1200

cgggtgcagc ccgagatcct ggaggccctg ctgaagcaca tctccttcga caagttcgtg 1260

cagatctccc tgaaggccct gcggcggatc gtgcccctga tggagcaggg caagcggtac 1320

gacgaggcct gcgccgagat ctacggcgac cactacggca agaagaacac cgaggagaag 1380

atctacctgc cccccatccc cgccgacgag atccggaacc ccgtggtgct gcgggccctg 1440

tcccaggccc ggaaggtgat caacggcgtg gtgcggcggt acggctcccc cgcccggatc 1500

cacatcgaga ccgcccggga ggtgggcaag tccttcaagg accggaagga gatcgagaag 1560

cggcaggagg agaaccggaa ggaccgggag aaggccgccg ccaagttccg ggagtacttc 1620

cccaacttcg tgggcgagcc caagtccaag gacatcctga agctgcggct gtacgagcag 1680

cagcacggca agtgcctgta ctccggcaag gagatcaacc tggtgcggct gaacgagaag 1740

ggctacgtgg agatcgacca cgccctgccc ttctcccgga cctgggacga ctccttcaac 1800

aacaaggtgc tggtgctggg ctccgagaac cagaacaagg gcaaccagac cccctacgag 1860

tacttcaacg gcaaggacaa ctcccgggag tggcaggagt tcaaggcccg ggtggagacc 1920

tcccggttcc cccggtccaa gaagcagcgg atcctgctgc agaagttcga cgaggacggc 1980

ttcaaggagt gcaacctgaa cgacacccgg tacgtgaacc ggttcctgtg ccagttcgtg 2040

gccgaccaca tcctgctgac cggcaagggc aagcggcggg tgttcgcctc caacggccag 2100

atcaccaacc tgctgcgggg cttctggggc ctgcggaagg tgcgggccga gaacgaccgg 2160

caccacgccc tggacgccgt ggtggtggcc tgctccaccg tggccatgca gcagaagatc 2220

acccggttcg tgcggtacaa ggagatgaac gccttcgacg gcaagaccat cgacaaggag 2280

accggcaagg tgctgcacca gaagacccac ttcccccagc cctgggagtt cttcgcccag 2340

gaggtgatga tccgggtgtt cggcaagccc gacggcaagc ccgagttcga ggaggccgac 2400

acccccgaga agctgcggac cctgctggcc gagaagctgt cctcccggcc cgaggccgtg 2460

cacgagtacg tgacccccct gttcgtgtcc cgggccccca accggaagat gtccggcgcc 2520

cacaaggaca ccctgcggtc cgccaagcgg ttcgtgaagc acaacgagaa gatctccgtg 2580

aagcgggtgt ggctgaccga gatcaagctg gccgacctgg agaacatggt gaactacaag 2640

aacggccggg agatcgagct gtacgaggcc ctgaaggccc ggctggaggc ctacggcggc 2700

aacgccaagc aggccttcga ccccaaggac aacccccttct acaagaaggg cggccagctg 2760

gtgaaggccg tgcgggtgga gaagacccag gagtccggcg tgctgctgaa caagaagaac 2820

gcctacacca tcgccgacaa cggcgacatg gtgcgggtgg acgtgttctg caaggtggac 2880

aagaagggca agaaccagta cttcatcgtg cccatctacg cctggcaggt ggccgagaac 2940

atcctgcccg acatcgactg caagggctac cggatcgacg actcctacac cttctgcttc 3000

tccctgcaca agtacgacct gatcgccttc cagaaggacg agaagtccaa ggtggagttc 3060

gcctactaca tcaactgcga ctcctccaac ggccggttct acctggcctg gcacgacaag 3120

ggctccaagg agcagcagtt ccggatctcc acccagaacc tggtgctgat ccagaagtac 3180

caggtgaacg agctgggcaa ggagatccgg ccctgccggc tgaagaagcg gccccccgtg 3240

cgg 3243

<210> 391
<211> 3249
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 391
atggctgctt ttaagcctaa tcctattaat tatattcttg gtcttgctat tggtattgct 60

tctgttggtt gggctatggt tgagattgat gaggaggaga atcctattcg tcttattgat 120

cttggtgttc gtgtttttga gcgtgctgag gttcctaaga ctggtgattc tcttgctatg 180

gctcgtcgtc ttgctcgttc tgttcgtcgt cttactcgtc gtcgtgctca tcgtcttctt 240

cgtgctcgtc gtcttcttaa gcgtgagggt gttcttcagg ctgctgattt tgatgagaat 300

ggtcttatta agtctcttcc taatactcct tggcagcttc gtgctgctgc tcttgatcgt 360

aagcttactc ctcttgagtg gtctgctgtt cttcttcatc ttattaagca tcgtggttat 420

ctttctcagc gtaagaatga gggtgagact gctgataagg agcttggtgc tcttcttaag 480

ggtgttgcta ataatgctca tgctcttcag actggtgatt ttcgtactcc tgctgagctt 540

gctcttaata agtttgagaa ggagtctggt catattcgta atcagcgtgg tgattattct 600

catacttttt ctcgtaagga tcttcaggct gagcttattc ttctttttga gaagcagaag 660

gagtttggta atcctcatgt ttctggtggt cttaaggagg gtattgagac tcttcttatg 720

actcagcgtc ctgctctttc tggtgatgct gttcagaaga tgcttggtca ttgtactttt 780

gagcctgctg agcctaaggc tgctaagaat acttatactg ctgagcgttt tatttggctt 840

actaagctta ataatcttcg tattcttgag cagggttctg agcgtcctct tactgatact 900

gagcgtgcta ctcttatgga tgagccttat cgtaagtcta agcttactta tgctcaggct 960

cgtaagcttc ttggtcttga ggatactgct ttttttaagg gtcttcgtta tggtaaggat 1020

aatgctgagg cttctactct tatggagatg aaggcttatc atgctatttc tcgtgctctt 1080

gagaaggagg gtcttaagga taagaagtct cctcttaatc tttcttctga gcttcaggat 1140

gagattggta ctgctttttc tcttttaag actgatgagg atattactgg tcgtcttaag 1200

gatcgtgttc agcctgagat tcttgaggct cttcttaagc atatttcttt tgataagttt 1260

gttcagattt ctcttaaggc tcttcgtcgt attgttcctc ttatggagca gggtaagcgt 1320

tatgatgagg cttgtgctga gatttatggt gatcattatg gtaagaagaa tactgaggag 1380

aagatttatc ttcctcctat tcctgctgat gagattcgta atcctgttgt tcttcgtgct 1440

ctttctcagg ctcgtaaggt tattaatggt gttgttcgtc gttatggttc tcctgctcgt 1500

attcatattg agactgctcg tgaggttggt aagtctttta aggatcgtaa ggagattgag 1560

aagcgtcagg aggagaatcg taaggatcgt gagaaggctg ctgctaagtt tcgtgagtat 1620

tttcctaatt ttgttggtga gcctaagtct aaggatattc ttaagcttcg tctttatgag 1680

cagcagcatg gtaagtgtct ttattctggt aaggagatta atcttgttcg tcttaatgag 1740

aagggttatg ttgagattga tcatgctctt ccttttttctc gtacttggga tgattctttt 1800

aataataagg ttcttgttct tggttctgag aatcagaata agggtaatca gactccttat 1860

gagtatttta atggtaagga taattctcgt gagtggcagg agtttaaggc tcgtgttgag 1920

acttctcgtt ttcctcgttc taagaagcag cgtattcttc ttcagaagtt tgatgaggat 1980

ggttttaagg agtgtaatct taatgatact cgttatgtta atcgttttct ttgtcagttt 2040

gttgctgatc atattcttct tactggtaag ggtaagcgtc gtgtttttgc ttctaatggt 2100

cagattacta atcttcttcg tggtttttgg ggtcttcgta aggttcgtgc tgagaatgat 2160

cgtcatcatg ctcttgatgc tgttgttgtt gcttgttcta ctgttgctat gcagcagaag 2220

attactcgtt ttgttcgtta taaggagatg aatgcttttg atggtaagac tattgataag 2280

gagactggta aggttcttca tcagaagact cattttcctc agccttggga gtttttgct 2340

caggaggtta tgattcgtgt ttttggtaag cctgatggta agcctgagtt tgaggaggct 2400

gatactcctg agaagcttcg tactcttctt gctgagaagc tttcttctcg tcctgaggct 2460

gttcatgagt atgttactcc tctttttgtt tctcgtgctc ctaatcgtaa gatgtctggt 2520

gctcataagg atactcttcg ttctgctaag cgttttgtta agcataatga gaagatttct 2580

gttaagcgtg tttggcttac tgagattaag cttgctgatc ttgagaatat ggttaattat 2640

aagaatggtc gtgagattga gctttatgag gctcttaagg ctcgtcttga ggcttatggt 2700

ggtaatgcta agcaggcttt tgatcctaag gataatcctt tttataagaa gggtggtcag 2760

cttgttaagg ctgttcgtgt tgagaagact caggagtctg gtgttcttct taataagaag 2820

aatgcttata ctattgctga taatggtgat atggttcgtg ttgatgtttt ttgtaaggtt 2880

gataagaagg gtaagaatca gtattttatt gttcctattt atgcttggca ggttgctgag 2940

aatattcttc ctgatattga ttgtaagggt tatcgtattg atgattctta tacttttgt 3000

ttttctcttc ataagtatga tcttattgct tttcagaagg atgagaagtc taaggttgag 3060

tttgcttatt atattaattg tgattcttct aatggtcgtt tttatcttgc ttggcatgat 3120

aagggttcta aggagcagca gtttcgtatt tctactcaga atcttgttct tattcagaag 3180

tatcaggtta atgagcttgg taaggagatt cgtccttgtc gtcttaagaa gcgtcctcct 3240

gttcgtuga 3249

<210> 392
<211> 3249
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 392
atggccgcct tcaagcccaa ccccatcaac tacatcctgg gcctggccat cggcatcgcc 60

tccgtgggct gggccatggt ggagatcgac gaggaggaga accccatccg gctgatcgac 120

ctgggcgtgc gggtgttcga gcgggccgag gtgcccaaga ccggcgactc cctggccatg 180

gcccggcggc tggcccggtc cgtgcggcgg ctgacccggc ggcgggccca ccggctgctg 240

cgggcccggc ggctgctgaa gcgggagggc gtgctgcagg ccgccgactt cgacgagaac 300

ggcctgatca agtccctgcc caacaccccc tggcagctgc gggccgccgc cctggaccgg 360

aagctgaccc ccctggagtg gtccgccgtg ctgctgcacc tgatcaagca ccggggctac 420

ctgtcccagc ggaagaacga gggcgagacc gccgacaagg agctgggcgc cctgctgaag 480

ggcgtggcca acaacgccca cgccctgcag accggcgact tccggacccc cgccgagctg 540

gccctgaaca agttcgagaa ggagtccggc cacatccgga accagcgggg cgactactcc 600

cacaccttct cccggaagga cctgcaggcc gagctgatcc tgctgttcga gaagcagaag 660

gagttcggca acccccacgt gtccggcggc ctgaaggagg gcatcgagac cctgctgatg 720

acccagcggc ccgccctgtc cggcgacgcc gtgcagaaga tgctgggcca ctgcaccttc 780

gagcccgccg agcccaaggc cgccaagaac acctacaccg ccgagcggtt catctggctg 840

accaagctga acaacctgcg gatcctggag cagggctccg agcggccccc gaccgacacc 900

gagcgggcca ccctgatgga cgagccctac cggaagtcca agctgaccta cgcccaggcc 960

cggaagctgc tgggcctgga ggacaccgcc ttcttcaagg gcctgcggta cggcaaggac 1020

aacgccgagg cctccaccct gatggagatg aaggcctacc acgccatctc ccgggccctg 1080

gagaaggagg gcctgaagga caagaagtcc cccctgaacc tgtcctccga gctgcaggac 1140

gagatcggca ccgccttctc cctgttcaag accgacgagg acatcaccgg ccggctgaag 1200

gaccgggtgc agcccgagat cctggaggcc ctgctgaagc acatctcctt cgacaagttc 1260

gtgcagatct ccctgaaggc cctgcggcgg atcgtgcccc tgatggagca gggcaagcgg 1320

tacgacgagg cctgcgccga gatctacggc gaccactacg gcaagaagaa caccgaggag 1380

aagatctacc tgccccccat ccccgccgac gagatccgga accccgtggt gctgcgggcc 1440

ctgtcccagg cccggaaggt gatcaacggc gtggtgcggc ggtacggctc ccccgcccgg 1500

atccacatcg agaccgcccg ggaggtgggc aagtccttca aggaccggaa ggagatcgag 1560

aagcggcagg aggagaaccg gaaggaccgg gagaaggccg ccgccaagtt ccgggagtac 1620

ttccccaact tcgtgggcga gcccaagtcc aaggacatcc tgaagctgcg gctgtacgag 1680

cagcagcacg gcaagtgcct gtactccggc aaggagatca acctggtgcg gctgaacgag 1740

aagggctacg tggagatcga ccacgccctg cccttctccc ggacctggga cgactccttc 1800

aacaacaagg tgctggtgct gggctccgag aaccagaaca agggcaacca gacccctac 1860

gagtacttca acggcaagga caactcccgg gagtggcagg agttcaaggc ccgggtggag 1920

acctcccggt tcccccggtc caagaagcag cggatcctgc tgcagaagtt cgacgaggac 1980

ggcttcaagg agtgcaacct gaacgacacc cggtacgtga accggttcct gtgccagttc 2040

gtggccgacc acatcctgct gaccggcaag ggcaagcggc gggtgttcgc ctccaacggc 2100

cagatcacca acctgctgcg gggcttctgg ggcctgcgga aggtgcgggc cgagaacgac 2160

cggcaccacg ccctggacgc cgtggtggtg gcctgctcca ccgtggccat gcagcagaag 2220

atcacccggt tcgtgcggta caaggagatg aacgccttcg acggcaagac catcgacaag 2280

gagaccggca aggtgctgca ccagaagacc cacttccccc agccctggga gttcttcgcc 2340

caggaggtga tgatccgggt gttcggcaag cccgacggca agcccgagtt cgaggaggcc 2400

gacacccccg agaagctgcg gaccctgctg gccgagaagc tgtcctcccg gcccgaggcc 2460

gtgcacgagt acgtgacccc cctgttcgtg tcccgggccc ccaaccggaa gatgtccggc 2520

gcccacaagg acaccctgcg gtccgccaag cggttcgtga agcacaacga gaagatctcc 2580

gtgaagcggg tgtggctgac cgagatcaag ctggccgacc tggagaacat ggtgaactac 2640

aagaacggcc gggagatcga gctgtacgag gccctgaagg cccggctgga ggcctacggc 2700

ggcaacgcca agcaggcctt cgaccccaag gacaacccct tctacaagaa gggcggccag 2760

ctggtgaagg ccgtgcgggt ggagaagacc caggagtccg gcgtgctgct gaacaagaag 2820

aacgcctaca ccatcgccga caacggcgac atggtgcggg tggacgtgtt ctgcaaggtg 2880

gacaagaagg gcaagaacca gtacttcatc gtgcccatct acgcctggca ggtggccgag 2940

aacatcctgc ccgacatcga ctgcaagggc taccggatcg acgactccta caccttctgc 3000

ttctccctgc acaagtacga cctgatcgcc ttccagaagg acgagaagtc caaggtggag 3060

ttcgcctact acatcaactg cgactcctcc aacggccggt tctacctggc ctggcacgac 3120

aagggctcca aggagcagca gttccggatc tccacccaga acctggtgct gatccagaag 3180

taccaggtga acgagctggg caaggagatc cggccctgcc ggctgaagaa gcggcccccc 3240

gtgcgguga 3249

<210> 393
<211> 3249
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 393
atggccgcct tcaagcccaa ccccatcaac tacatcctgg gcctggccat cggcatcgcc 60

tccgtgggct gggccatggt ggagatcgac gaggaggaga accccatccg gctgatcgac 120

ctgggcgtgc gggtgttcga gcgggccgag gtgcccaaga ccggcgactc cctggccatg 180

gcccggcggc tggcccggtc cgtgcggcgg ctgacccggc ggcgggccca ccggctgctg 240

cgggcccggc ggctgctgaa gcgggagggc gtgctgcagg ccgccgactt cgacgagaac 300

ggcctgatca agtccctgcc caacaccccc tggcagctgc gggccgccgc cctggaccgg 360

aagctgaccc ccctggagtg gtccgccgtg ctgctgcacc tgatcaagca ccggggctac 420

ctgtcccagc ggaagaacga gggcgagacc gccgacaagg agctgggcgc cctgctgaag 480

ggcgtggcca acaacgccca cgccctgcag accggcgact tccggacccc cgccgagctg 540

gccctgaaca agttcgagaa ggagtccggc cacatccgga accagcgggg cgactactcc 600

cacaccttct cccggaagga cctgcaggcc gagctgatcc tgctgttcga gaagcagaag 660

gagttcggca acccccacgt gtccggcggc ctgaaggagg gcatcgagac cctgctgatg 720

acccagcggc ccgccctgtc cggcgacgcc gtgcagaaga tgctgggcca ctgcaccttc 780

gagcccgccg agcccaaggc cgccaagaac acctacaccg ccgagcggtt catctggctg 840

accaagctga acaacctgcg gatcctggag cagggctccg agcggcccct gaccgacacc 900

gagcgggcca ccctgatgga cgagccctac cggaagtcca agctgaccta cgcccaggcc 960

cggaagctgc tgggcctgga ggacaccgcc ttcttcaagg gcctgcggta cggcaaggac 1020

aacgccgagg cctccaccct gatggagatg aaggcctacc acgccatctc ccgggccctg 1080

gagaaggagg gcctgaagga caagaagtcc cccctgaacc tgtcctccga gctgcaggac 1140

gagatcggca ccgccttctc cctgttcaag accgacgagg acatcaccgg ccggctgaag 1200

gaccgggtgc agcccgagat cctggaggcc ctgctgaagc acatctcctt cgacaagttc 1260

gtgcagatct ccctgaaggc cctgcggcgg atcgtgcccc tgatggagca gggcaagcgg 1320

tacgacgagg cctgcgccga gatctacggc gaccactacg gcaagaagaa caccgaggag 1380

aagatctacc tgcccccccat ccccgccgac gagatccgga accccgtggt gctgcgggcc 1440

ctgtcccagg cccggaaggt gatcaacggc gtggtgcggc ggtacggctc ccccgcccgg 1500

atccacatcg agaccgcccg ggaggtgggc aagtccttca aggaccggaa ggagatcgag 1560

aagcggcagg aggagaaccg gaaggaccgg gagaaggccg ccgccaagtt ccgggagtac 1620

ttccccaact tcgtgggcga gcccaagtcc aaggacatcc tgaagctgcg gctgtacgag 1680

cagcagcacg gcaagtgcct gtactccggc aaggagatca acctggtgcg gctgaacgag 1740

aagggctacg tggagatcga ccacgccctg cccttctccc ggacctggga cgactccttc 1800

aacaacaagg tgctggtgct gggctccgag aaccagaaca agggcaacca gacccctac 1860

gagtacttca acggcaagga caactcccgg gagtggcagg agttcaaggc ccgggtggag 1920

acctcccggt tcccccggtc caagaagcag cggatcctgc tgcagaagtt cgacgaggac 1980

ggcttcaagg agtgcaacct gaacgacacc cggtacgtga accggttcct gtgccagttc 2040

gtggccgacc acatcctgct gaccggcaag ggcaagcggc gggtgttcgc ctccaacggc 2100

cagatcacca acctgctgcg gggcttctgg ggcctgcgga aggtgcgggc cgagaacgac 2160

cggcaccacg ccctggacgc cgtggtggtg gcctgctcca ccgtggccat gcagcagaag 2220

atcacccggt tcgtgcggta caaggagatg aacgccttcg acggcaagac catcgacaag 2280

gagaccggca aggtgctgca ccagaagacc cacttccccc agccctggga gttcttcgcc 2340

caggaggtga tgatccgggt gttcggcaag cccgacggca agcccgagtt cgaggaggcc 2400

gacacccccg agaagctgcg gaccctgctg gccgagaagc tgtcctcccg gcccgaggcc 2460

gtgcacgagt acgtgacccc cctgttcgtg tcccgggccc ccaaccggaa gatgtccggc 2520

gcccacaagg acaccctgcg gtccgccaag cggttcgtga agcacaacga gaagatctcc 2580

gtgaagcggg tgtggctgac cgagatcaag ctggccgacc tggagaacat ggtgaactac 2640

aagaacggcc gggagatcga gctgtacgag gccctgaagg cccggctgga ggcctacggc 2700

ggcaacgcca agcaggcctt cgaccccaag gacaacccct tctacaagaa gggcggccag 2760

ctggtgaagg ccgtgcgggt ggagaagacc caggagtccg gcgtgctgct gaacaagaag 2820

aacgcctaca ccatcgccga caacggcgac atggtgcggg tggacgtgtt ctgcaaggtg 2880

gacaagaagg gcaagaacca gtacttcatc gtgcccatct acgcctggca ggtggccgag 2940

aacatcctgc ccgacatcga ctgcaagggc taccggatcg acgactccta caccttctgc 3000

ttctccctgc acaagtacga cctgatcgcc ttccagaagg acgagaagtc caaggtggag 3060

ttcgcctact acatcaactg cgactcctcc aacggccggt tctacctggc ctggcacgac 3120

aagggctcca aggagcagca gttccggatc tccacccaga acctggtgct gatccagaag 3180

taccaggtga acgagctggg caaggagatc cggccctgcc ggctgaagaa gcggcccccc 3240

gtgcgguaa 3249

<210> 394

<400> 394
000

<210> 395

<400> 395
000

<210> 396

<400> 396
000

<210> 397

<400> 397
000

<210> 398

<400> 398
000

<210> 399

<400> 399
000

<210> 400

<400> 400
000

<210> 401
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 401
cgcacgggua ccaggggcca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 402
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 402
gcacggguac caggggccac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 403
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 403
cacggguacc aggggccacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 404
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 404
gggaggcgcc ccguggcccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 405
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 405
gcgaucaggg aggcgccccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 406
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 406
cgggcuggcc ucccacaagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 407
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 407
cuccuugugg gaggccagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 408
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 408
ggcuggccuc ccacaaggag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 409
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 409
uugucucccc uccuuguggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 410
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 410
ccacaaggag gggagacaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 411
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 411
cacaaggagg ggagacaauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 412
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 412
caauugucuc cccuccuugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 413
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 413
ccaauugucu ccccuccuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 414
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 414
aucccucgaa uacugaugag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 415
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 415
aaccacucau caguauucga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 416
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 416
gaaccacuca ucaguauucg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 417

<400> 417
000

<210> 418
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 418
ggcccgugac uuuuccucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 419
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 419
ugcuucacac ucaaugugug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 420
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 420
gcuucacacu caaugugugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 421
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 421
cuucacacuc aaugugugug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 422
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 422
uucacacuca augugugugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 423

<400> 423
000

<210> 424
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 424

aggggcccug acccugcuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 425

<400> 425
000

<210> 426

<400> 426
000

<210> 427
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 427
ugagauuucu ugucucacug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 428
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 428
gagauuucuu gucucacuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 429
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 429
uaaagcaccu guuaaaauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 430
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 430
aaucuguccu ucauuuuaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 431

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 431
gucacagggg aaggucccug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 432
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 432
aacaugaaga aagcaggugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 433
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 433
aaacaugaag aaagcaggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 434
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 434
ucaggaaaca ugaagaaagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 435
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 435
uguccuguga gauaccagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 436

<400> 436
000

<210> 437
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 437
aggcugaugc cugagguccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 438
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 438
ggcugaugcc ugagguccuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 439
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 439
cacaauaucc caaggaccuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 440
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 440
gguccuuggg auauuguguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 441
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 441
guccuuggga uauuguguuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 442
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 442
cucccaaaca caauauccca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 443
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 443
uccucuagcc acaucuucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 444

<400> 444
000

<210> 445
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 445
ccucuagcca caucuucugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 446
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 446
cccacagaag auguggcuag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 447
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 447
gucagauccc acagaagaug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 448
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 448
aucuucugug ggaucugacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 449
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 449
cccaggcagu gacagugccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 450
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 450
cugggcacug ucacugccug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 451
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 451
ccugggcacu gucacugccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 452
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 452
uggagggccu gaugugugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 453
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 453
gccugaugug uguugggugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 454

<400> 454
000

<210> 455
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
for detailed description of substitutions and preferred
embodiments

<400> 455
cccaacaccc aacacacauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 456
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
for detailed description of substitutions and preferred
embodiments

<400> 456
uuggguguug ggcggaacag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 457
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 457
uggauguauu gagcaugcga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 458
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 458
ggauguauug agcaugcgau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 459
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 459
acggccaucc ucggcgucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 460
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 460
gacggccauc cucggcgucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 461
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 461
gacgccgagg auggccguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 462
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 462
ugacggccau ccucggcguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 463
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 463
ggcgccauga cggccauccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 464
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 464
acagcgacgc cgcgagccag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 465
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 465
cgacgccgcg agccagagga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 466
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 466

cgagccagag gauggagccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 467
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 467
cggcuccauc cucuggcucg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 468
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 468
gagccagagg auggagccgc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 469
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 469
gcgcccgcgg cuccauccuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 470
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 470
gcccguccgu gggggaugag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 471
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 471
ucaucccccca cggacgggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 472
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 472
aucucggacc cggagacugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 473
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 473
gggguccccgc ggcuucgggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 474
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 474
cucggggucc cgcggcuucg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 475
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 475
ucucgggguc ccgcggcuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 476
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 476
gucucggggu cccgcggcuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 477
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 477
gcaagggucu cggggucccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 478
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 478
ggaccccgag acccuugccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 479
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 479
gaccccgaga cccuugcccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 480
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 480
cgagacccuu gccccgggag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 481

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 481
cucccggggc aagggucucg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 482
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 482
ucucccgggg caagggucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 483
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 483
cucucccggg gcaagggucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 484
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 484
ccuugccccg ggagaggccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 485
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 485
cugggccucu cccggggcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 486
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 486
ccugggccuc ucccggggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 487
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 487
aggcgccugg gccucucccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 488
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 488
uuuaggccaa aaaucccccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 489
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 489
aggcauuuug caucugucau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 490
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 490
caggcauuuu gcaucuguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 491
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 491
cccugggcac ugucacugcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 492
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 492
cgcugcagcg cacggguacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 493
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 493
gcugcagcgc acggguacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 494
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 494
cugcagcgca cggguaccag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 495
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 495
uccuuguggg aggccagccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 496
<211> 91
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 496
cuugacgcau cgcgccagga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 497
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 497
aauccuaauu gccuccugca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 498
<211> 91
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 498
acucacgcug gauagccucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 499
<211> 91
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 499
uucaaaaccu gucagugauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 500
<211> 91

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 500
cgcccagguc cucacgucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 501
<211> 91
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 501
gcugcagcgc acggguacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 502
<211> 91
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 502
ccacacccaa aaggccacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 503
<211> 101
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 503
gugugucccu cuccccaccc guccguugua gcucccugaa accguugcua caauaaggcc 60

gucgaaagau gugccgcaac gcucugccuu cuggcaucgu u 101

<210> 504
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 504
aucugccuuu auagggcacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 505
<211> 99
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 505
agucacgaug ccuuuauagg uuuuagagcu agaaauagca aguuaaaaua aggcuagucc 60

guuaucaacu ugaaaaagug gcaccgaguc ggugcuuuu 99

<210> 506
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 506
gguugcaagg aaugagaacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 507
<211> 101
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 507
ggggccagcu ucagacacaa auacguugua gcucccugaa accguugcua caauaaggcc 60

gucgaaagau gugccgcaac gcucugccuu cuggcaucgu u   101

<210> 508
<211> 101
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 508
cagaauugag agacucagcc caggguugua gcucccugaa accguugcua caauaaggcc 60

gucgaaagau gugccgcaac gcucugccuu cuggcaucgu u   101

<210> 509
<211> 101
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 509
gagaagccgu cacacagauc cacaguugua gcucccugaa accguugcua caauaaggcc 60

gucgaaagau gugccgcaac gcucugccuu cuggcaucgu u 101

<210> 510
<211> 93
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 510
ccaagugucu uccaguacga uuugguugua gcucccuucg accguugcua caauaaggcc 60

gucgaugugc cgcaacgcuc ugccggcauc guu 93

<210> 511
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 511
acacaaauac caguccagcg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 512
<211> 145
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (1)..(24)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (21)..(25)
<223> n may or may not be present

<400> 512
nnnnnnnnnn nnnnnnnnnn nnnnguugua gcucccuuuc ucauuucgga aacgaaauga 60

gaaccguugc uacaauaagg ccgucugaaa agaugugccg caacgcucug ccccuuaaag 120

cuucugcuuu aaggggcauc guuua 145

<210> 513
<211> 102
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (1)..(25)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (21)..(25)
<223> n may or may not be present

<400> 513
nnnnnnnnnn nnnnnnnnnn nnnnnguugu agcucccuga aaccguugcu acaauaaggc 60

cgucgaaaga ugugccgcaa cgcucugccu ucuggcaucg uu 102

<210> 514
<211> 106
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (1)..(25)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (21)..(25)
<223> n may or may not be present

<400> 514
nnnnnnnnnn nnnnnnnnnn nnnnnguugu agcucccuga aaccguugcu acaauaaggc 60

cgucgaaaga ugugccgcaa cgcucugccu ucuggcaucg uuuauu 106

<210> 515
<211> 108
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (1)..(25)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (21)..(25)
<223> n may or may not be present

<400> 515
nnnnnnnnnn nnnnnnnnnn nnnnnguugu agcucccugg aaacccguug cuacaauaag 60

gccgucgaaa gaugugccgc aacgcucugc cuucuggcau cguuuauu 108

<210> 516
<211> 77
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 516
guuguagcuc ccugaaaccg uugcuacaau aaggccgucg aaagaugugc cgcaacgcuc 60

ugccuucugg caucguu 77

<210> 517
<211> 77
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 517
guuguagcuc ccugaaaccg uugcuacaau aaggccgucg aaagaugugc cgcaacgcuc 60

ugccuucugg caucguu 77

<210> 518
<211> 94
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains internal linker L1; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(25)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (21)..(25)
<223> n may or may not be present

<400> 518
nnnnnnnnnn nnnnnnnnnn nnnnnguugu agcucccuuc gaccguugcu acaauaaggc 60

cgucgaugug ccgcaacgcu cugccggcau cguu 94

<210> 519
<211> 93
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications and internal linker L1; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(24)
<223> n is a, c, g, or u

<400> 519
nnnnnnnnnn nnnnnnnnnn nnnnguugua gcucccuucg accguugcua caauaaggcc 60

gucgaugugc cgcaacgcuc ugccggcauc guu 93

<210> 520
<211> 91
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is a, c, g, or u

<400> 520
nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 521
<211> 91
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is a, c, g, or u

<400> 521
nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucacg aaagggcacc gagucggugc u 91

<210> 522
<211> 101
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<220>
<221> misc_feature
<222> (1)..(24)
<223> n is a, c, g, or u

<400> 522
nnnnnnnnnn nnnnnnnnnn nnnnguugua gcucccugaa accguugcua caauaaggcc   60

gucgaaagau gugccgcaac gcucugccuu cuggcaucgu u   101

<210> 523
<211> 80
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications and internal linkers L1 and
       L2; see specification as filed for detailed description of
       substitutions and preferred embodiments

<400> 523
acgcaaauau caguccagcg guuuuagagc uauagcaagu uaaaauaagg cguccguuau   60

cacgggcacc gagucggugc   80

<210> 524
<211> 93
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains internal linker L1; see specification as filed for
detailed description of substitutions and preferred embodiments

<220>
<221> misc_feature
<222> (1)..(20)
<223> n is a, c, g, or u

<400> 524
nnnnnnnnnn nnnnnnnnnn guuuuagagc uauagcaagu uaaaauaagg cuaguccguu 60

aucaacuuaa guggcaccga gucggugcuu uuu 93

<210> 525

<400> 525
000

<210> 526

<400> 526
000

<210> 527

<400> 527
000

<210> 528

<400> 528
000

<210> 529
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 529
uucccggccu uuuuaccuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 530
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 530
ucaacugcga ccaguucagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 531
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 531
agcgcaggca guggcaggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 532
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 532
accugcaaca acaggauuca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 533
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 533
ccaggaacac cugcaacaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 534
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 534
cagcagcaag agccuggagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 535
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 535
gcagcagcaa gagccuggag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 536
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 536
caaucucuuc uucuccagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 537
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 537
agcagcucgc ugccagccuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 538
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 538
agcucgcugc cagccuucgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 539
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 539
cucuggacca ggcggccccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 540
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 540
gcagcccucg acagcccccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 541
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 541
cagcccucga cagccccccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 542
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 542
agccaaguac ccccucccag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 543
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 543
cagccaacag caccucagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 544
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 544
ugccgcgccc gcaguguccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 545
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 545
cgacagcccc cccggggccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 546
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 546
ggcagcgagc ugcugggccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 547
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 547
gccagcucug ccagggcccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 548
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 548
agcgagcgaa ggcagggccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 549
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 549
aaggcuggca gcgagcugcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 550
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 550
agcccucgac agccccccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 551
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 551
ggaugcagcg agcgaaggca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 552
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 552
uccaccgagg cagccgccga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 553
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 553
ucuccaacaa gcuuccaaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 554
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 554
agcagccccc ggagggagca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 555
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 555

gccacagccc uacuuugugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 556
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 556
cugcgcccac gaggccgagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 557
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 557
cagccgccga uggcccgagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 558
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 558
cgaggcuccc caauccagag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 559
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 559
cucaacgagg aacuggagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 560
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 560
ccacagcccu acuuugugcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 561
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 561
caagagccug gagcgggaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 562
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 562
gacugccagu caccacagug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 563
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 563
gcccacgagg ccgaggaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 564
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 564
aucagcccag ccagaaagcg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 565
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 565
cagagaagac aaagucguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 566
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 566
ugggaguccc ugcagcagca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 567
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 567
cagcaggcug uugugugaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 568
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 568
cuggucaggg caagagcuau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 569
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 569
cagguucagg caugcugggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 570

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 570
uuuccaagga cuucagcugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 571
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 571
aggccgagga ggcuggaauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 572
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 572
caguggcuga uggagcgaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 573
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 573
gggaugcagc gagcgaaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 574
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 574
aagcugcccu ccacgcucac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 575
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 575
uuccuccugc aaugcuuccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 576
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 576
uccuccugca augcuuccug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 577
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 577
caagcugccc uccacgcuca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 578
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 578
ucccuggacc uccgcagcac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 579
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 579
uccagccucc ucggccucgu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 580
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 580
cuucuccagc cagguccauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 581
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 581
cuuccuccug caaugcuucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 582
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 582
cguccucccc aagcuccagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 583
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 583
ccagcuccuc gaagccgucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 584
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 584
cuuuuccucc cugcagcauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 585
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 585
cggccgccac gaguggcugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 586
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 586
cgcccagguc cucacgucug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 587
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 587
cacgugcgga ccggcaccgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 588
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 588
cagcuuggcc agcucugcca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 589
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 589
ucggacucug cggcccgcgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 590
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 590
cuucccccag cugaaguccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 591
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 591
gcccagcucc caggccagcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 592
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 592
cccuccagcc aguugucaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 593
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 593
gcccuccagc caguugucau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 594
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 594
cccugacgcu ccuccgggac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 595
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 595
cacacugccc ggcacaaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 596
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 596
cagcucacag ugugccacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 597
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 597
agcucggacu cugcggcccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 598
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 598
gacgcccuau uugagcuguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 599
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 599
gccagugcug cggaggucca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 600
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 600
agggcuccca ggcagcgggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 601
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 601
gccgagcccg caggcccgga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 602
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 602
gcucccaggc agcgggcggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 603
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 603
ccucuccagc ugccgggcau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 604
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 604
aggcuuuccc caaacuggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 605
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 605

ccaucuccac ucugccccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 606
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 606
uccuccucac agcccggccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 607
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 607
ccacucugcc ccaugggcuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 608
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 608
cagccucccg cccgcugccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 609
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 609
cccggccgcc ucucuuuucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 610
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 610
ccugggccca cagccacucg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 611
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 611
uccccggcug cagccgcuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 612
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 612
cgcguucugc ucauccuaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 613
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 613
uuccacaucc uucagggacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 614
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 614
cucuccagcu gccgggcauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 615
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 615
gggcccacag ccacucgugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 616
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 616
ccccggccgc cucucuuuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 617
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 617
uccagcugcc gggcauuggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 618
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 618
ccacacccaa aaggccacac 20

<210> 619
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 619
cccaccagcu cagcuccacg 20

<210> 620
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 620
ucccuagcag gaucucauag 20

<210> 621
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 621
ggcucaaaca cagcgaccuc 20

<210> 622
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 622
ggcacaccag uguggccuuu 20

<210> 623
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 623
agguggccga gacccucagg 20

<210> 624
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 624
accugcucua ccccaggccu 20

<210> 625
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 625

cauagaggau gguggcagac 20

<210> 626
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 626
ccacguggag cugagcuggu 20

<210> 627
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 627
ggagaaugac gaguggaccc 20

<210> 628
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 628
ccaguguggc cuuuugggug 20

<210> 629
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 629
caaacacagc gaccucgggu 20

<210> 630
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 630
uaccauggcc aucaacacaa 20

<210> 631
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 631
gcgcugacga ucugggugac 20

<210> 632
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 632
cacggacccg cagccccuca 20

<210> 633
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 633
ucaaacacag cgaccucggg 20

<210> 634

<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 634
cagggaagaa gccuguggcc 20

<210> 635
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 635
caguguggcc uuuugggugu 20

<210> 636
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 636
accacgugga gcugagcugg 20

<210> 637
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 637
gagggcgggc ugcuccuuga 20

<210> 638
<211> 20
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 638
agcucagcuc cacgugguca 20

<210> 639
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 639
gaacaaggug uucccacccg 20

<210> 640
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 640
gcacaccagu guggccuuuu 20

<210> 641
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 641
gagcuggugg gugaauggga 20

<210> 642
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 642
ggcugcuccu ugaggggcug 20

<210> 643
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 643
cgggugggaa caccuuguuc 20

<210> 644
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 644
cagcucagcu ccacgugguc 20

<210> 645
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 645
gacgaucugg gugacggguu 20

<210> 646
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 646
uagcaggauc ucauagagga 20

<210> 647
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 647
gaccagcaca gcauacaggg 20

<210> 648
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 648
cugaccacgu ggagcugagc 20

<210> 649
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 649
ccaacagugu ccuaccagca 20

<210> 650
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 650

cuggugggug aaugggaagg 20

<210> 651
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 651
cuaugagauc cugcuaggga 20

<210> 652
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 652
caggaucuca uagaggaugg 20

<210> 653
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 653
ggccacccug uaugcugugc 20

<210> 654
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 654
caacaguguc cuaccagcaa 20

<210> 655
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 655
agcugagcug gugggugaau 20

<210> 656
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 656
aacagugucc uaccagcaag 20

<210> 657
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 657
aggcuucuuc ccugaccacg 20

<210> 658
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 658
ucuucugcag gucaagagaa 20

<210> 659

<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 659
gagcugagcu ggugggugaa 20

<210> 660
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 660
ggucagcgcc cuuguguuga 20

<210> 661
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 661
agaucgucag cgccgaggcc 20

<210> 662
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 662
gcugcuccuu gaggggcugc 20

<210> 663
<211> 20
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 663
guaucuggag ucauugaggg 20

<210> 664
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 664
auccucuaug agauccugcu 20

<210> 665
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 665
uccucuauga gauccugcua 20

<210> 666
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 666
gcaguaucug gagucauuga 20

<210> 667
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 667
gcugaccagc acagcauaca 20

<210> 668
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 668
acaggguggc cuucccuagc 20

<210> 669
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 669
cgcugaccag cacagcauac 20

<210> 670
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 670
agagaucucc cacacccaaa 20

<210> 671
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 671
ggcucucgga gaaugacgag 20

<210> 672
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 672
uguggccagg cacaccagug 20

<210> 673
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 673
gcggcugcuc aggcaguauc 20

<210> 674
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 674
acacuggugu gccuggccac 20

<210> 675
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 675

cguagaacug gacuugacag 20

<210> 676
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 676
cuugacagcg gaagugguug 20

<210> 677
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 677
acuggacuug acagcggaag 20

<210> 678
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 678
ugagggucuc ggccaccuuc 20

<210> 679
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 679
aguccaguuc uacgggcucu 20

<210> 680
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 680
cacccagauc gucagcgccg 20

<210> 681
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 681
gaucgucagc gccgaggccu 20

<210> 682
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 682
aucgucagcg ccgaggccug 20

<210> 683
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 683
acugccugag cagccgccug 20

<210> 684

<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 684
gagcagccgc cugagggucu 20

<210> 685
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 685
agggcgggcu gcuccuugag 20

<210> 686
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 686
gcugucaagu ccaguucuac 20

<210> 687
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 687
aggccucggc gcugacgauc 20

<210> 688
<211> 20
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 688
cugccugagc agccgccuga 20

<210> 689
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 689
agacccucag gcggcugcuc 20

<210> 690
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 690
cgcugucaag uccaguucua 20

<210> 691
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 691
aaaggccaca cuggugugcc 20

<210> 692
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 692
uugacagcgg aaguggguugc 20

<210> 693
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 693
ccacucaccu gcucuacccc 20

<210> 694
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 694
uggcucaaac acagcgaccu 20

<210> 695
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 695
uggcagacag gaccccuugc 20

<210> 696
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 696
guugcggggg uucugccaga 20

<210> 697
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 697
ugagggcggg cugcuccuug 20

<210> 698
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 698
ucuccgagag cccguagaac 20

<210> 699
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 699
ggccucggcg cugacgaucu 20

<210> 700
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 700

augacgagug gacccaggau 20

<210> 701
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 701
ugacgagugg acccaggaua 20

<210> 702
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 702
gcgggguuc ugccagaagg 20

<210> 703
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 703
ggugcacagu ggggucagca 20

<210> 704
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 704
ggcgcugacg aucuggguga 20

<210> 705
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 705
agacaggacc ccuugcuggu 20

<210> 706
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 706
aacaaaugug ucacaaagua 20

<210> 707
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 707
uuucaaaacc ugucagugau 20

<210> 708
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 708
cuuaccuggg cuggggaaga 20

<210> 709

<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 709
ccgaauccuc cuccugaaag 20

<210> 710
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 710
cugacagguu uugaaaguuu 20

<210> 711
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 711
cuggggaaga aggugucuuc 20

<210> 712
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 712
uccuccuccu gaaaguggcc 20

<210> 713
<211> 20
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 713
ccacuuucag gaggaggauu 20

<210> 714
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 714
auuuguuuga gaaucaaaau 20

<210> 715
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 715
cuucaagagc aacagugcug 20

<210> 716
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 716
agcugcccuu accugggcug 20

<210> 717
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 717
agagcaacag ugcuguggcc 20

<210> 718
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 718
aaagcugccc uuaccugggc 20

<210> 719
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 719
aagcugcccu uaccugggcu 20

<210> 720
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 720
uggaauaaug cuguuguuga 20

<210> 721
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 721
caccaaagcu gcccuuaccu 20

<210> 722
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 722
agagucucuc agcugguaca 20

<210> 723
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 723
acacggcagg gucaggguuc 20

<210> 724
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 724
agcugguaca cggcaggguc 20

<210> 725
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 725

cucucagcug guacacggca 20

<210> 726
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 726
ucucucagcu gguacacggc 20

<210> 727
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 727
uggauuuaga gucucucagc 20

<210> 728
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 728
uaggcagaca gacuugucac 20

<210> 729
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 729
uucggaaccc aaucacugac 20

<210> 730
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 730
gcugguacac ggcaggguca 20

<210> 731
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 731
acaaaacugu gcuagacaug 20

<210> 732
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 732
gagaaucaaa aucggugaau 20

<210> 733
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 733
aaguuccugu gaugucaagc 20

<210> 734

<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 734
ugcucaugac gcugcggcug 20

<210> 735
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 735
uuaaucugcu caugacgcug 20

<210> 736
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 736
gauuaaaccc ggccacuuuc 20

<210> 737
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 737
cucgaccagc uugacaucac 20

<210> 738
<211> 20
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 738
uaaacccggc cacuuucagg 20

<210> 739
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 739
acccggccac uuucaggagg 20

<210> 740
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 740
aaagucagau uuguugcucc 20

<210> 741
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 741
gcaccaaagc ugcccuuacc 20

<210> 742
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 742
auccuccucc ugaaaguggc 20

<210> 743
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 743
uucaaaaccu gucagugauu 20

<210> 744
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 744
ugugcuagac augaggucua 20

<210> 745
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 745
cgucaugagc agauuaaacc 20

<210> 746
<211> 20
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 746
ucaaggcccc ucaccucagc 20

<210> 747
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 747
aacaaaugug ucacaaagua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 748
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 748
uuucaaaacc ugucagugau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 749
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 749
cuuaccuggg cuggggaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 750
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<400> 750
ccgaauccuc cuccugaaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 751
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 751
cugacagguu uugaaaguuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 752
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 752
cuggggaaga aggugucuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 753
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 753
uccuccuccu gaaaguggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 754
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 754
ccacuuucag gaggaggauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 755
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 755
auuuguuuga gaaucaaaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 756
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 756
cuucaagagc aacagugcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 757
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 757
agcugcccuu accugggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 758
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 758
agagcaacag ugcuguggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 759
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 759
aaagcugccc uuaccugggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 760
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 760
aagcugcccu uaccugggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 761
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 761
uggaauaaug cuguuguuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 762
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 762
caccaaagcu gcccuuaccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 763

<400> 763
000

<210> 764

<400> 764
000

<210> 765

<400> 765
000

<210> 766

<400> 766
000

<210> 767

<400> 767
000

<210> 768

<400> 768
000

<210> 769

<400> 769
000

<210> 770

<400> 770
000

<210> 771

<400> 771
000

<210> 772

<400> 772
000

<210> 773

<400> 773
000

<210> 774

<400> 774
000

<210> 775

<400> 775
000

<210> 776

<400> 776
000

<210> 777

<400> 777
000

<210> 778

<400> 778
000

<210> 779

<400> 779
000

<210> 780

<400> 780
000

<210> 781

<400> 781
000

<210> 782

<400> 782
000

<210> 783

<400> 783
000

<210> 784

<400> 784
000

<210> 785

<400> 785
000

<210> 786

<400> 786
000

<210> 787

<400> 787
000

<210> 788

<400> 788
000

<210> 789

<400> 789
000

<210> 790

<400> 790
000

<210> 791

<400> 791
000

<210> 792

<400> 792
000

<210> 793

<400> 793
000

<210> 794

<400> 794
000

<210> 795

<400> 795
000

<210> 796

<400> 796
000

<210> 797

<400> 797
000

<210> 798

<400> 798
000

<210> 799

<400> 799
000

<210> 800

<400> 800
000

<210> 801
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 801
ccacacccaa aaggccacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 802
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 802
cccaccagcu cagcuccacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 803
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 803
ucccuagcag gaucucauag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 804
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 804
ggcucaaaca cagcgaccuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 805
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 805
ggcacaccag uguggccuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 806
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 806
agguggccga gacccucagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 807
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 807
accugcucua ccccaggccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 808
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 808
cauagaggau gguggcagac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 809
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 809
ccacguggag cugagcuggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 810
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 810
ggagaaugac gaguggaccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 811
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 811
ccaguguggc cuuuugggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 812
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 812
caaacacagc gaccucgggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 813
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 813
uaccauggcc aucaacacaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 814
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 814
gcgcugacga ucugggugac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 815
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 815
cacggacccg cagccccuca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 816
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 816

ucaaacacag cgaccucggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 817
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 817
cagggaagaa gccuguggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 818
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 818
caguguggcc uuuugggugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 819
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 819
accacgugga gcugagcugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 820
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 820
gagggcgggc ugcuccuuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 821
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 821
agcucagcuc cacgugguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 822
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 822
gaacaaggug uucccacccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 823
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 823
gcacaccagu guggccuuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 824
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 824
gagcuggugg gugaauggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 825
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 825
ggcugcuccu ugaggggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 826
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 826
cggggugggaa caccuuguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 827
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 827
cagcucagcu ccacgugguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 828
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 828
gacgaucugg gugacggguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 829
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 829
uagcaggauc ucauagagga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 830
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 830
gaccagcaca gcauacaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 831

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 831
cugaccacgu ggagcugagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 832
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 832
ccaacagugu ccuaccagca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 833
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 833
cuggugggug aaugggaagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 834
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 834
cuaugagauc cugcuaggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 835
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 835
caggaucuca uagaggaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 836
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 836
ggccacccug uaugcugugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 837
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 837
caacaguguc cuaccagcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 838
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 838
agcugagcug gugggugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 839
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 839
aacagugucc uaccagcaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 840
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 840
aggcuucuuc ccugaccacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 841
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 841
ucuucugcag gucaagagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 842
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 842
gagcugagcu gguggggugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 843
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 843
ggucagcgcc cuuguguuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 844
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 844
agaucgucag cgccgaggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 845
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 845
gcugcuccuu gaggggcugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 846
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 846
guaucuggag ucauugaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 847
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 847
auccucuaug agauccugcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 848
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 848
uccucuauga gauccugcua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 849
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 849
gcaguaucug gagucauuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 850
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 850
gcugaccagc acagcauaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 851
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 851
acaggguggc cuucccuagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 852
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 852
cgcugaccag cacagcauac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 853
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 853
aacaaaugug ucacaaagua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 854
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 854
uuucaaaacc ugucagugau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 855
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 855
cuuaccuggg cuggggaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 856
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 856
ccgaauccuc cuccugaaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 857
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 857
cugacagguu uugaaaguuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 858
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 858
cuggggaaga aggugucuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 859
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 859
uccuccuccu gaaaguggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 860
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 860
ccacuuucag gaggaggauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 861
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 861
auuuguuuga gaaucaaaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 862
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 862
cuucaagagc aacagugcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 863
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 863
agcugcccuu accugggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 864
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 864
agagcaacag ugcuguggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 865
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 865
aaagcugccc uuaccugggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 866
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 866

aagcugcccu uaccugggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 867
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 867
uggaauaaug cuguuguuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 868
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 868
caccaaagcu gcccuuaccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 869
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 869
gauccucguc guggugcucg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 870
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 870
ccucgucgug gugcucgcgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 871
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 871
gcaucgcgcc aggacggucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 872
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 872
uguacuugac gcaucgcgcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 873
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 873
caccgggcug aacucgcagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 874
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 874
ucgcgguggu cgucccgagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 875
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 875
ccaccgcgag caccacgacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 876
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 876
caucgcgcca ggacggucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 877
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 877
gacggucucg ggaaagcgcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 878
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 878
gcgcuuuccc gagaccgucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 879
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 879
cuugacgcau cgcgccagga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 880
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 880
ugcgaguuca gcccgguguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 881

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 881
ugcucgcggu ggucgucccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 882
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 882
aggguuuguc cccggacacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 883
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 883
gcgaguucag cccggugucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 884
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 884
ggucucggga aagcgcuugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 885
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 885
cgaguucagc ccgguguccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 886
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 886
ccggcagcag gguuuguccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 887
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 887
guugggcucu ccuagagagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 888
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 888
cgagcaccac gacgaggauc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 889
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 889
ggccaacugc gaguucagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 890
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 890
uacugacgcc aagacagagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 891
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 891
ucgccaaccc accucaucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 892
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 892
ccggggacaa acccugcugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 893
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 893
ggaaagcgcu ugguggugcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 894
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 894
ucuccuagag agccggcagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 895
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 895
cgccagcagu ggagcggucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 896
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 896
cuccacugcu ggcgccaccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 897
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 897
caggguuugu ccccggacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 898
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 898
cugucuuggc gucaguaucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 899
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 899
ugcccggacc gcuccacugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 900
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 900
cuccuagaga gccggcagca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 901
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 901
ccugguccug auccucgucg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 902
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 902
ucccgaggug gcgccagcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 903
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 903
gcgccagcag uggagcgguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 904
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 904
uccacugcug gcgccaccuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 905
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 905
aggagagccc aacucugucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 906
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 906
acugacgcca agacagaguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 907
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 907
aacccugcug ccggcucucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 908
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 908
uaucagccac uaaugaaguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 909
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 909
ugaaagcauc ccauacacuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 910
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 910
cccccaauua ccuuguugca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 911
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 911
auguagacug ccaaagugua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 912
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 912
aaguguaugg gaugcuuuca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 913
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 913
aauuaccuug uugcaaggua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 914
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 914
ugaguuccca acuucauuag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 915
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 915
uguagacugc caaaguguau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 916
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 916

aguguauggg augcuuucaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 917
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 917
cuaucagcca cuaaugaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 918
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 918
ucuucuucag uaauguugca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 919
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 919
ucuggcccau caguucacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 920
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 920
gcgggacaug uucacccugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 921
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 921
ccagcgggac auguucaccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 922
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 922
agccuagcag cguguccucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 923
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 923
ugaauucacc acaccaugug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 924
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 924
aucaguucac acagguccag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 925
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 925
gcugaugacc ucacauggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 926
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 926
gccuagcagc guguccucca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 927
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 927
accaugugag gucaucagca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 928
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 928
ucaguucaca cagguccagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 929
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 929
ccagggugaa cauguccgc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 930
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 930
gcuggaccug ugugaacuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 931

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 931
acccuggagg acacgcugcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 932
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 932
cuggaccugu gugaacugau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 933
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 933
ucuggaaaac gguuucccgc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 934

<400> 934
000

<210> 935

<400> 935
000

<210> 936

<400> 936
000

<210> 937
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 937
guuuuccaga auacugaaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 938
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 938
gggauccauu gagcaucaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 939
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 939
caucacaugg accacaucac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 940
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 940
gugguccaug ugaugcucaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 941
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 941
agagaagguu cagacacuag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 942
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 942
ucuagugucu gaaccuucuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 943
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 943
gguucagaca cuagaggccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 944
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 944
ccauaauuug caaccagaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 945
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 945
ccuucucugg uugcaaauua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 946
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 946
agguucagac acuagaggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 947
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 947
cuagaggccu gggugauaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 948
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 948
uuuuagcacu uuugggagug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 949
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 949
ucuuccacca uguaucaccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 950
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 950
cuuccccaga gacuuaugcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 951
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 951
uuaccuguag auauucuugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 952
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 952
gcucuuuuau ggugggaucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 953
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 953
ggaucccacc auaaaagagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 954
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 954

cgauuccagc ucuuuuaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 955
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 955
gauuccagcu cuuuuauggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 956
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 956
aaucgauucc agcucuuuua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 957

<400> 957

000

<210> 958

<400> 958
000

<210> 959

<400> 959
000

<210> 960
<211> 113
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 960

Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1   5   10   15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
            20   25   30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
        35   40   45

Ser Val Trp Arg His Thr Ser Gln Asn Thr Ser Asn His Val Glu Val
    50   55   60

Asn Phe Leu Glu Lys Phe Thr Thr Glu Arg Tyr Phe Arg Pro Asn Thr
65   70   75   80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
                85   90   95

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg His Pro Tyr Val Thr Leu
            100   105   110

Phe

<210> 961
<211> 121
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 961

Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1 5 10 15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
20 25 30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
35 40 45

Ser Val Trp Arg His Thr Ser Gln Asn Thr Ser Asn His Val Glu Val
50 55 60

Asn Phe Leu Glu Lys Phe Thr Thr Glu Arg Tyr Phe Arg Pro Asn Thr
65 70 75 80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
85 90 95

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg His Pro Tyr Val Thr Leu
100 105 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His
115 120

<210> 962
<211> 128
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 962

Met Asn Trp Asn Ala Leu Arg Ser Lys Ala Ile Glu Val Ser Arg His
1   5   10   15

Ala Tyr Ala Pro Tyr Ser Gly Phe Pro Val Gly Ala Ala Ala Leu Val
            20   25   30

Asp Asp Gly Arg Thr Val Thr Gly Cys Asn Val Glu Asn Val Ser Tyr
        35   40   45

Gly Leu Gly Leu Cys Ala Glu Cys Ala Val Val Cys Ala Leu His Ser
    50   55   60

Gly Gly Gly Gly Arg Leu Val Ala Leu Ser Cys Val Gly Pro Asp Gly
65   70   75   80

Gly Val Leu Met Pro Cys Gly Arg Cys Arg Gln Val Leu Leu Glu His
                85   90   95

Gly Gly Pro Glu Leu Leu Ile Asp His Ala His Gly Pro Arg Pro Leu
            100   105   110

Arg Glu Leu Leu Pro Asp Ala Phe Gly Pro Asp Asp Leu Gly Arg Arg
        115   120   125

<210> 963
<211> 130
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 963

Met Thr His His Ala Leu Ile Glu Ala Ala Lys Ala Ala Arg Glu Lys
1   5   10   15

```
Ala Tyr Ala Pro Tyr Ser Asn Phe Lys Val Gly Ala Ala Leu Val Thr
            20                  25                  30

Asn Asp Gly Lys Val Phe His Gly Cys Asn Val Glu Asn Ala Ser Tyr
        35                  40                  45

Gly Leu Cys Asn Cys Ala Glu Arg Thr Ala Leu Phe Ser Ala Leu Ala
    50                  55                  60

Ala Gly Tyr Arg Pro Gly Glu Phe Ala Ala Ile Ala Val Val Gly Glu
65                  70                  75                  80

Thr His Gly Pro Ile Ala Pro Cys Gly Ala Cys Arg Gln Val Met Ile
                85                  90                  95

Glu Leu Gly Lys Pro Thr Leu Glu Val Val Leu Thr Asn Met Gln Gly
            100                 105                 110

Asp Val Arg Val Thr Ser Ala Gly Asp Leu Leu Pro Asp Ala Phe Tyr
        115                 120                 125

Leu Ala
    130
```

<210> 964
<211> 132
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 964

```
Met Pro Asp Ile Asp Trp Lys Gln Leu Arg Asp Lys Ala Thr Gln Val
1               5                   10                  15

Ala Ala Gly Ala Tyr Ala Pro Tyr Ser Arg Phe Pro Val Gly Ala Ala
            20                  25                  30
```

Ala Leu Val Asp Asp Gly Arg Val Val Thr Gly Cys Asn Val Glu Asn
        35   40   45

Val Ser Tyr Gly Leu Ala Leu Cys Ala Glu Cys Gly Val Val Cys Ala
    50   55   60

Leu His Ala Thr Gly Gly Gly Arg Leu Val Ala Leu Ala Cys Val Asp
65   70   75   80

Gly Arg Gly Ala Pro Leu Met Pro Cys Gly Arg Cys Arg Gln Leu Leu
                85   90   95

Phe Glu His Gly Gly Pro Glu Leu Leu Val Asp His Leu Ala Gly Pro
            100   105   110

Arg Arg Leu Gly Asp Leu Leu Pro Glu Pro Phe His Ala Asp Leu Thr
        115   120   125

Gly Glu Pro Thr
    130

<210> 965
<211> 132
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 965

Met Asn Ser Lys Gln Leu Ile Gln Glu Ala Ile Glu Ala Arg Lys Gln
1   5   10   15

Ala Tyr Val Pro Tyr Ser Lys Phe Gln Val Gly Ala Ala Leu Leu Thr
            20   25   30

Gln Asp Gly Lys Val Tyr Arg Gly Cys Asn Val Glu Asn Ala Ser Tyr
        35   40   45

```
Gly Leu Cys Asn Cys Ala Glu Arg Thr Ala Leu Phe Lys Ala Val Ser
    50                  55                  60

Glu Gly Asp Lys Glu Phe Val Ala Ile Ala Ile Val Ala Asp Thr Lys
65                  70                  75                  80

Arg Pro Val Pro Pro Cys Gly Ala Cys Arg Gln Val Met Val Glu Leu
                85                  90                  95

Cys Lys Gln Asp Thr Lys Val Tyr Leu Ser Asn Leu His Gly Asp Val
            100                 105                 110

Gln Glu Thr Thr Val Gly Glu Leu Leu Pro Gly Ala Phe Leu Ala Glu
        115                 120                 125

Asp Leu His Glu
    130
```

<210> 966
<211> 133
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 966

```
Met Pro Asp Val Asp Trp Asn Met Leu Arg Gly Asn Ala Thr Gln Ala
1               5                   10                  15

Ala Ala Gly Ala Tyr Val Pro Tyr Ser Arg Phe Ala Val Gly Ala Ala
            20                  25                  30

Ala Leu Val Asp Asp Gly Arg Val Val Thr Gly Cys Asn Val Glu Asn
        35                  40                  45

Val Ser Tyr Gly Leu Thr Leu Cys Ala Glu Cys Ala Val Val Cys Ala
    50                  55                  60
```

```
Leu His Ser Thr Gly Gly Gly Arg Leu Leu Ala Leu Ala Cys Val Asp
65                  70                  75                  80


Gly His Gly Ser Val Leu Met Pro Cys Gly Arg Cys Arg Gln Val Leu
                85                  90                  95


Leu Glu His Gly Gly Ser Glu Leu Leu Ile Asp His Pro Val Arg Pro
            100                 105                 110


Arg Arg Leu Gly Asp Leu Leu Pro Asp Ala Phe Gly Leu Asp Asp Leu
        115                 120                 125


Pro Arg Glu Arg Arg
    130


<210>  967
<211>  134
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  967

Met Gly Asp Val Asn Trp Asp Thr Leu Gln Lys Ala Ala Val Ala Ala
1               5                   10                  15


Arg Ala Asn Ser Tyr Ala Pro Tyr Ser Asn Phe Pro Val Gly Val Ala
            20                  25                  30


Gly Phe Val Asn Asp Gly Arg Leu Ile Thr Gly Val Asn Val Glu Asn
        35                  40                  45


Ala Ser Tyr Gly Leu Ala Leu Cys Ala Glu Cys Ser Met Ile Ser Ala
    50                  55                  60


Leu Tyr Ala Thr Gly Gly Gly Arg Leu Val Ala Val Tyr Cys Val Asp
65                  70                  75                  80
```

Gly Asn Gly Asp Ser Leu Met Pro Cys Gly Arg Cys Arg Gln Leu Leu
85 90 95

Tyr Glu His Gly Gly Pro Glu Leu Lys Ile Met Thr Pro Lys Gly Val
100 105 110

Gln Thr Met Ala Gln Leu Leu Pro Gln Ala Phe Asn Pro Gln Glu Arg
115 120 125

Ile Phe Gly Asn Asp Glu
130

<210> 968
<211> 136
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 968

Met Asn Arg Gln Glu Leu Ile Thr Glu Ala Leu Lys Ala Arg Asp Met
1 5 10 15

Ala Tyr Ala Pro Tyr Ser Lys Phe Gln Val Gly Ala Ala Leu Leu Thr
20 25 30

Lys Asp Gly Lys Val Tyr Arg Gly Cys Asn Ile Glu Asn Ala Ala Tyr
35 40 45

Ser Met Cys Asn Cys Ala Glu Arg Thr Ala Leu Phe Lys Ala Val Ser
50 55 60

Glu Gly Asp Thr Glu Phe Gln Met Leu Ala Val Ala Ala Asp Thr Pro
65 70 75 80

Gly Pro Val Ser Pro Cys Gly Ala Cys Arg Gln Val Ile Ser Glu Leu
85 90 95

Cys Thr Lys Asp Val Ile Val Val Leu Thr Asn Leu Gln Gly Gln Ile
100 105 110

Lys Glu Met Thr Val Glu Glu Leu Leu Pro Gly Ala Phe Ser Ser Glu
115 120 125

Asp Leu His Asp Glu Arg Lys Leu
130 135

<210> 969
<211> 142
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 969

Met Lys Val Gly Gly Ile Glu Asp Arg Gln Leu Glu Ala Leu Lys Arg
1 5 10 15

Ala Ala Leu Lys Ala Cys Glu Leu Ser Tyr Ser Pro Tyr Ser His Phe
20 25 30

Arg Val Gly Cys Ser Ile Leu Thr Asn Asn Asp Val Ile Phe Thr Gly
35 40 45

Ala Asn Val Glu Asn Ala Ser Tyr Ser Asn Cys Ile Cys Ala Glu Arg
50 55 60

Ser Ala Met Ile Gln Val Leu Met Ala Gly His Arg Ser Gly Trp Lys
65 70 75 80

Cys Met Val Ile Cys Gly Asp Ser Glu Asp Gln Cys Val Ser Pro Cys
85 90 95

Gly Val Cys Arg Gln Phe Ile Asn Glu Phe Val Val Lys Asp Phe Pro
100 105 110

Ile Val Met Leu Asn Ser Thr Gly Ser Arg Ser Lys Val Met Thr Met
        115   120   125

Gly Glu Leu Leu Pro Met Ala Phe Gly Pro Ser His Leu Asn
    130   135   140

<210> 970
<211> 142
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 970

Met Asn Ile Glu Gln Gln Leu Tyr Asp Val Val Lys Gln Leu Ile Glu
1   5   10   15

Gln Arg Tyr Pro Asn Asp Trp Gly Gly Ala Ala Ala Ile Arg Val Glu
            20   25   30

Asp Gly Thr Ile Tyr Thr Ser Val Ala Pro Asp Val Ile Asn Ala Ser
        35   40   45

Thr Glu Leu Cys Met Glu Thr Gly Ala Ile Leu Glu Ala His Lys Phe
    50   55   60

Gln Lys Lys Val Thr His Ser Ile Cys Leu Ala Arg Glu Asn Glu His
65   70   75   80

Ser Glu Leu Lys Val Leu Ser Pro Cys Gly Val Cys Gln Glu Arg Leu
                85   90   95

Phe Tyr Trp Gly Pro Glu Val Gln Cys Ala Ile Thr Asn Ala Lys Gln
            100   105   110

Asp Ile Ile Phe Lys Pro Leu Lys Glu Leu Gln Pro Tyr His Trp Thr
        115   120   125

Glu Ala Tyr His Asp Glu Met Val Lys Glu Trp Ser Thr Arg
130 135 140

<210> 971
<211> 146
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 971

Met Ala Gln Glu Arg Pro Ser Cys Ala Val Glu Pro Glu His Val Gln
1 5 10 15

Arg Leu Leu Leu Ser Ser Arg Glu Ala Lys Lys Ser Ala Tyr Cys Pro
20 25 30

Tyr Ser Arg Phe Pro Val Gly Ala Ala Leu Leu Thr Gly Asp Gly Arg
35 40 45

Ile Phe Ser Gly Cys Asn Ile Glu Asn Ala Cys Tyr Pro Leu Gly Val
50 55 60

Cys Ala Glu Arg Thr Ala Ile Gln Lys Ala Ile Ser Glu Gly Tyr Lys
65 70 75 80

Asp Phe Arg Ala Ile Ala Ile Ser Ser Asp Leu Gln Glu Glu Phe Ile
85 90 95

Ser Pro Cys Gly Ala Cys Arg Gln Val Met Arg Glu Phe Gly Thr Asp
100 105 110

Trp Ala Val Tyr Met Thr Lys Pro Asp Gly Thr Phe Val Val Arg Thr
115 120 125

Val Gln Glu Leu Leu Pro Ala Ser Phe Gly Pro Glu Asp Leu Gln Lys
130 135 140

Ile Gln
145

<210> 972
<211> 146
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 972

Met Ala Gln Lys Arg Pro Ala Cys Thr Leu Lys Pro Glu Cys Val Gln
1 5 10 15

Gln Leu Leu Val Cys Ser Gln Glu Ala Lys Lys Ser Ala Tyr Cys Pro
20 25 30

Tyr Ser His Phe Pro Val Gly Ala Ala Leu Leu Thr Gln Glu Gly Arg
35 40 45

Ile Phe Lys Gly Cys Asn Ile Glu Asn Ala Cys Tyr Pro Leu Gly Ile
50 55 60

Cys Ala Glu Arg Thr Ala Ile Gln Lys Ala Val Ser Glu Gly Tyr Lys
65 70 75 80

Asp Phe Arg Ala Ile Ala Ile Ala Ser Asp Met Gln Asp Asp Phe Ile
85 90 95

Ser Pro Cys Gly Ala Cys Arg Gln Val Met Arg Glu Phe Gly Thr Asn
100 105 110

Trp Pro Val Tyr Met Thr Lys Pro Asp Gly Thr Tyr Ile Val Met Thr
115 120 125

Val Gln Glu Leu Leu Pro Ser Ser Phe Gly Pro Glu Asp Leu Gln Lys
130 135 140

Thr Gln
145

<210> 973
<211> 158
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 973

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 974
<211> 158
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 974

Met Thr Thr Thr Lys Ala Asn Leu Thr Glu Phe Glu Gln Gln Leu Val
1 5 10 15

Asp Lys Ala Ile Gly Ala Met Glu Asn Ala Tyr Cys Lys Tyr Ser Asn
20 25 30

Phe Lys Val Gly Ala Ala Leu Val Cys Asp Asp Gly Glu Ile Ile Ile
35 40 45

Gly Ala Asn His Glu Asn Ala Ser Tyr Gly Ala Thr Ile Cys Ala Glu
50 55 60

Arg Ser Ala Ile Val Thr Ala Leu Thr Lys Gly His Arg Lys Phe Lys
65 70 75 80

Tyr Ile Val Val Ala Thr Glu Leu Glu Ala Pro Cys Ser Pro Cys Gly
85 90 95

Val Cys Arg Gln Val Leu Ile Glu Phe Gly Asp Tyr Lys Val Ile Leu
100 105 110

Gly Ser Ser Thr Ser Asp Gln Ile Ile Glu Thr Thr Thr Tyr Glu Leu
115 120 125

Leu Pro Tyr Ala Phe Thr Pro Lys Ser Leu Asp Asp His Glu Lys Glu
130 135 140

Thr Glu Glu Arg Lys His His Asn Asp His Asn Asn Lys Glu
145 150 155

<210> 975
<211> 166
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 975

Met Ala Ala Asn Ser Leu Pro Gln Asp Ile Ser Asp Val Glu Leu Val
1 5 10 15

His Leu Ala Arg Ala Ala Met Lys Arg Ala His Cys Pro Tyr Ser Lys
20 25 30

Phe Pro Val Gly Ala Ala Leu Leu Thr Glu Ser Gly Glu Ile Val Gln
35 40 45

Gly Cys Asn Val Glu Asn Ala Ser Tyr Gly Gly Thr Ile Cys Ala Glu
50 55 60

Arg Ser Ala Ile Val Ser Ala Val Ser Gln Gly Tyr Thr Lys Phe Arg
65 70 75 80

Ala Ile Ala Val Val Thr Glu Leu Ser Glu Pro Ala Ser Pro Cys Gly
85 90 95

Leu Cys Arg Gln Phe Leu Val Glu Phe Gly Asp Tyr Lys Val Val Val
100 105 110

Gly Thr Ala Ser Asn Asn Lys Ile Leu Ile Thr Ser Thr Arg Ala Leu
115 120 125

Leu Pro Phe Ala Phe Thr Pro Glu Ser Leu Asp Thr Phe Glu Gln Glu
130 135 140

Lys Ala Ser Glu Ala Lys Gly Leu Lys Gln Asp Asp Ala Thr Glu His
145 150 155 160

Asn Val Thr Val Val Ser
165

<210> 976
<211> 190
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 976

Met Asn Pro Gln Ile Arg Asn Pro Met Lys Ala Met Tyr Pro Gly Thr
1 5 10 15

Phe Tyr Phe Gln Phe Lys Asn Leu Trp Glu Ala Asn Asp Arg Asn Glu
20 25 30

Thr Trp Leu Cys Phe Thr Val Glu Gly Ile Lys Arg Arg Ser Val Val
35 40 45

Ser Trp Lys Thr Gly Val Phe Arg Asn Gln Val Asp Ser Glu Thr His
50 55 60

Cys His Ala Glu Arg Cys Phe Leu Ser Trp Phe Cys Asp Asp Ile Leu
65 70 75 80

Ser Pro Asn Thr Lys Tyr Gln Val Thr Trp Tyr Thr Ser Trp Ser Pro
85 90 95

Cys Pro Asp Cys Ala Gly Glu Val Ala Glu Phe Leu Ala Arg His Ser
100 105 110

Asn Val Asn Leu Thr Ile Phe Thr Ala Arg Leu Tyr Tyr Phe Gln Tyr
115 120 125

```
Pro Cys Tyr Gln Glu Gly Leu Arg Ser Leu Ser Gln Glu Gly Val Ala
    130                 135                 140

Val Glu Ile Met Asp Tyr Glu Asp Phe Lys Tyr Cys Trp Glu Asn Phe
145                 150                 155                 160

Val Tyr Asn Asp Asn Glu Pro Phe Lys Pro Trp Lys Gly Leu Lys Thr
                165                 170                 175

Asn Phe Arg Leu Leu Lys Arg Arg Leu Arg Glu Ser Leu Gln
            180                 185                 190
```

<210> 977
<211> 190
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 977

```
Met Asn Pro Gln Ile Arg Asn Pro Met Glu Ala Met Tyr Pro His Ile
1               5                   10                  15

Phe Tyr Phe His Phe Lys Asn Leu Leu Lys Ala Cys Gly Arg Asn Glu
            20                  25                  30

Ser Trp Leu Cys Phe Thr Met Glu Val Thr Lys His His Ser Ala Val
        35                  40                  45

Phe Arg Lys Arg Gly Val Phe Arg Asn Gln Val Asp Pro Glu Thr His
    50                  55                  60

Cys His Ala Glu Arg Cys Phe Leu Ser Trp Phe Cys Asp Asp Ile Leu
65                  70                  75                  80

Ser Pro Asn Thr Asn Tyr Glu Val Thr Trp Tyr Thr Ser Trp Ser Pro
                85                  90                  95
```

```
Cys Pro Glu Cys Ala Gly Glu Val Ala Glu Phe Leu Ala Arg His Ser
            100                 105                 110

Asn Val Asn Leu Thr Ile Phe Thr Ala Arg Leu Cys Tyr Phe Trp Asp
        115                 120                 125

Thr Asp Tyr Gln Glu Gly Leu Cys Ser Leu Ser Gln Glu Gly Ala Ser
    130                 135                 140

Val Lys Ile Met Gly Tyr Lys Asp Phe Val Ser Cys Trp Lys Asn Phe
145                 150                 155                 160

Val Tyr Ser Asp Asp Glu Pro Phe Lys Pro Trp Lys Gly Leu Gln Thr
                165                 170                 175

Asn Phe Arg Leu Leu Lys Arg Arg Leu Arg Glu Ile Leu Gln
            180                 185                 190
```

<210> 978
<211> 193
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 978

```
Met Asn Asp Ala Leu His Ile Gly Leu Pro Pro Phe Leu Val Gln Ala
1               5                   10                  15

Asn Asn Glu Pro Arg Val Leu Ala Ala Pro Glu Ala Arg Met Gly Tyr
            20                  25                  30

Val Leu Glu Leu Val Arg Ala Asn Ile Ala Ala Asp Gly Gly Pro Phe
        35                  40                  45

Ala Ala Ala Val Phe Glu Arg Asp Ser Gly Leu Leu Ile Ala Ala Gly
    50                  55                  60
```

```
Thr Asn Arg Val Val Pro Gly Arg Cys Ser Ala Ala His Ala Glu Ile
65                  70                  75                  80


Leu Ala Leu Ser Leu Ala Gln Ala Lys Leu Asp Thr His Asp Leu Ser
                85                  90                  95


Ala Asp Gly Leu Pro Ala Cys Glu Leu Val Thr Ser Ala Glu Pro Cys
            100                 105                 110


Val Met Cys Phe Gly Ala Val Ile Trp Ser Gly Val Arg Ser Leu Val
        115                 120                 125


Cys Ala Ala Arg Ser Asp Asp Val Glu Ala Ile Gly Phe Asp Glu Gly
    130                 135                 140


Pro Arg Pro Glu Asn Trp Met Gly Gly Leu Glu Ala Arg Gly Ile Thr
145                 150                 155                 160


Val Thr Thr Gly Leu Leu Arg Asp Ala Ala Cys Ala Leu Leu Arg Glu
                165                 170                 175


Tyr Asn Ala Cys Asn Gly Val Ile Tyr Asn Ala Arg Cys Gly Val His
            180                 185                 190


Lys
```

<210> 979
<211> 198
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 979

```
Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys
1               5                   10                  15
```

```
Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val
            20                  25                  30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr
        35                  40                  45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr
    50                  55                  60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp
65                  70                  75                  80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp
                85                  90                  95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg
            100                 105                 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg
        115                 120                 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr
    130                 135                 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys
145                 150                 155                 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu
                165                 170                 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala
            180                 185                 190

Phe Arg Thr Leu Gly Leu
        195
```

<210> 980
<211> 198
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 980

Met Asp Ser Leu Leu Met Lys Gln Lys Lys Phe Leu Tyr His Phe Lys
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val
20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Cys Ser Leu Asp Phe Gly His
35 40 45

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp
65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Glu
85 90 95

Phe Leu Arg Trp Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg
100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg
115 120 125

Leu His Arg Ala Gly Val Gln Ile Gly Ile Met Thr Phe Lys Asp Tyr
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Arg Thr Phe Lys
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Thr Arg Gln Leu
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala
180 185 190

Phe Arg Met Leu Gly Phe
195

<210> 981
<211> 200
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 981

Met Ala Leu Leu Thr Ala Glu Thr Phe Arg Leu Gln Phe Asn Asn Lys
1 5 10 15

Arg Arg Leu Arg Arg Pro Tyr Tyr Pro Arg Lys Ala Leu Leu Cys Tyr
20 25 30

Gln Leu Thr Pro Gln Asn Gly Ser Thr Pro Thr Arg Gly Tyr Phe Glu
35 40 45

Asn Lys Lys Lys Cys His Ala Glu Ile Cys Phe Ile Asn Glu Ile Lys
50 55 60

Ser Met Gly Leu Asp Glu Thr Gln Cys Tyr Gln Val Thr Cys Tyr Leu
65 70 75 80

Thr Trp Ser Pro Cys Ser Ser Cys Ala Trp Glu Leu Val Asp Phe Ile
85 90 95

Lys Ala His Asp His Leu Asn Leu Gly Ile Phe Ala Ser Arg Leu Tyr
100 105 110

```
Tyr His Trp Cys Lys Pro Gln Gln Lys Gly Leu Arg Leu Leu Cys Gly
        115                 120                 125

Ser Gln Val Pro Val Glu Val Met Gly Phe Pro Lys Phe Ala Asp Cys
    130                 135                 140

Trp Glu Asn Phe Val Asp His Glu Lys Pro Leu Ser Phe Asn Pro Tyr
145                 150                 155                 160

Lys Met Leu Glu Glu Leu Asp Lys Asn Ser Arg Ala Ile Lys Arg Arg
                165                 170                 175

Leu Glu Arg Ile Lys Ile Pro Gly Val Arg Ala Gln Gly Arg Tyr Met
            180                 185                 190

Asp Ile Leu Cys Asp Ala Glu Val
        195                 200
```

<210> 982
<211> 224
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 982

```
Met Ala Gln Lys Glu Glu Ala Ala Val Ala Thr Glu Ala Ala Ser Gln
1               5                   10                  15

Asn Gly Glu Asp Leu Glu Asn Leu Asp Asp Pro Glu Lys Leu Lys Glu
            20                  25                  30

Leu Ile Glu Leu Pro Pro Phe Glu Ile Val Thr Gly Glu Arg Leu Pro
        35                  40                  45

Ala Asn Phe Phe Lys Phe Gln Phe Arg Asn Val Glu Tyr Ser Ser Gly
    50                  55                  60
```

```
Arg Asn Lys Thr Phe Leu Cys Tyr Val Val Glu Ala Gln Gly Lys Gly
65                  70                  75                  80

Gly Gln Val Gln Ala Ser Arg Gly Tyr Leu Glu Asp Glu His Ala Ala
                85                  90                  95

Ala His Ala Glu Glu Ala Phe Phe Asn Thr Ile Leu Pro Ala Phe Asp
            100                 105                 110

Pro Ala Leu Arg Tyr Asn Val Thr Trp Tyr Val Ser Ser Ser Pro Cys
        115                 120                 125

Ala Ala Cys Ala Asp Arg Ile Ile Lys Thr Leu Ser Lys Thr Lys Asn
    130                 135                 140

Leu Arg Leu Leu Ile Leu Val Gly Arg Leu Phe Met Trp Glu Glu Pro
145                 150                 155                 160

Glu Ile Gln Ala Ala Leu Lys Lys Leu Lys Glu Ala Gly Cys Lys Leu
                165                 170                 175

Arg Ile Met Lys Pro Gln Asp Phe Glu Tyr Val Trp Gln Asn Phe Val
            180                 185                 190

Glu Gln Glu Glu Gly Glu Ser Lys Ala Phe Gln Pro Trp Glu Asp Ile
        195                 200                 205

Gln Glu Asn Phe Leu Tyr Tyr Glu Glu Lys Leu Ala Asp Ile Leu Lys
    210                 215                 220
```

<210> 983
<211> 224
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 983

```
Met Ala Gln Lys Glu Glu Ala Ala Glu Ala Ala Ala Pro Ala Ser Gln
1               5                  10                  15

Asn Gly Asp Asp Leu Glu Asn Leu Glu Asp Pro Glu Lys Leu Lys Glu
            20                  25                  30

Leu Ile Asp Leu Pro Pro Phe Glu Ile Val Thr Gly Val Arg Leu Pro
        35                  40                  45

Val Asn Phe Phe Lys Phe Gln Phe Arg Asn Val Glu Tyr Ser Ser Gly
    50                  55                  60

Arg Asn Lys Thr Phe Leu Cys Tyr Val Val Glu Val Gln Ser Lys Gly
65                  70                  75                  80

Gly Gln Ala Gln Ala Thr Gln Gly Tyr Leu Glu Asp Glu His Ala Gly
                85                  90                  95

Ala His Ala Glu Glu Ala Phe Phe Asn Thr Ile Leu Pro Ala Phe Asp
            100                 105                 110

Pro Ala Leu Lys Tyr Asn Val Thr Trp Tyr Val Ser Ser Ser Pro Cys
        115                 120                 125

Ala Ala Cys Ala Asp Arg Ile Leu Lys Thr Leu Ser Lys Thr Lys Asn
    130                 135                 140

Leu Arg Leu Leu Ile Leu Val Ser Arg Leu Phe Met Trp Glu Glu Pro
145                 150                 155                 160

Glu Val Gln Ala Ala Leu Lys Lys Leu Lys Glu Ala Gly Cys Lys Leu
                165                 170                 175

Arg Ile Met Lys Pro Gln Asp Phe Glu Tyr Ile Trp Gln Asn Phe Val
            180                 185                 190

Glu Gln Glu Glu Gly Glu Ser Lys Ala Phe Glu Pro Trp Glu Asp Ile
        195                 200                 205
```

Gln Glu Asn Phe Leu Tyr Tyr Glu Glu Lys Leu Ala Asp Ile Leu Lys
210 215 220

<210> 984
<211> 229
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 984

Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1 5 10 15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
20 25 30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
35 40 45

Ser Val Trp Arg His Thr Ser Gln Asn Thr Ser Asn His Val Glu Val
50 55 60

Asn Phe Leu Glu Lys Phe Thr Thr Glu Arg Tyr Phe Arg Pro Asn Thr
65 70 75 80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
85 90 95

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg His Pro Tyr Val Thr Leu
100 105 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His His Thr Asp Gln Arg Asn Arg
115 120 125

Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met
130 135 140

```
Thr Glu Gln Glu Tyr Cys Tyr Cys Trp Arg Asn Phe Val Asn Tyr Pro
145                 150                 155                 160


Pro Ser Asn Glu Ala Tyr Trp Pro Arg Tyr Pro His Leu Trp Val Lys
                165                 170                 175


Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
            180                 185                 190


Leu Lys Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile
        195                 200                 205


Thr Leu Gln Thr Cys His Tyr Gln Arg Ile Pro Pro His Leu Leu Trp
    210                 215                 220


Ala Thr Gly Leu Lys
225
```

<210> 985
<211> 235
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 985

```
Met Asn Ser Lys Thr Gly Pro Ser Val Gly Asp Ala Thr Leu Arg Arg
1               5                   10                  15


Arg Ile Lys Pro Trp Glu Phe Val Ala Phe Phe Asn Pro Gln Glu Leu
            20                  25                  30


Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Lys Trp Gly Asn Gln Asn
        35                  40                  45


Ile Trp Arg His Ser Asn Gln Asn Thr Ser Gln His Ala Glu Ile Asn
    50                  55                  60
```

```
Phe Met Glu Lys Phe Thr Ala Glu Arg His Phe Asn Ser Ser Val Arg
65                  70                  75                  80

Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Trp Glu Cys Ser
                85                  90                  95

Lys Ala Ile Arg Lys Phe Leu Asp His Tyr Pro Asn Val Thr Leu Ala
            100                 105                 110

Ile Phe Ile Ser Arg Leu Tyr Trp His Met Asp Gln Gln His Arg Gln
        115                 120                 125

Gly Leu Lys Glu Leu Val His Ser Gly Val Thr Ile Gln Ile Met Ser
    130                 135                 140

Tyr Ser Glu Tyr His Tyr Cys Trp Arg Asn Phe Val Asp Tyr Pro Gln
145                 150                 155                 160

Gly Glu Glu Asp Tyr Trp Pro Lys Tyr Pro Tyr Leu Trp Ile Met Leu
                165                 170                 175

Tyr Val Leu Glu Leu His Cys Ile Ile Leu Gly Leu Pro Pro Cys Leu
            180                 185                 190

Lys Ile Ser Gly Ser His Ser Asn Gln Leu Ala Leu Phe Ser Leu Asp
        195                 200                 205

Leu Gln Asp Cys His Tyr Gln Lys Ile Pro Tyr Asn Val Leu Val Ala
    210                 215                 220

Thr Gly Leu Val Gln Pro Phe Val Thr Trp Arg
225                 230                 235
```

<210> 986
<211> 236
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 986

```
Met Thr Ser Glu Lys Gly Pro Ser Thr Gly Asp Pro Thr Leu Arg Arg
1               5                   10                  15


Arg Ile Glu Pro Trp Glu Phe Asp Val Phe Tyr Asp Pro Arg Glu Leu
            20                  25                  30


Arg Lys Glu Ala Cys Leu Leu Tyr Glu Ile Lys Trp Gly Met Ser Arg
        35                  40                  45


Lys Ile Trp Arg Ser Ser Gly Lys Asn Thr Thr Asn His Val Glu Val
    50                  55                  60


Asn Phe Ile Lys Lys Phe Thr Ser Glu Arg Asp Phe His Pro Ser Met
65                  70                  75                  80


Ser Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Trp Glu Cys
                85                  90                  95


Ser Gln Ala Ile Arg Glu Phe Leu Ser Arg His Pro Gly Val Thr Leu
            100                 105                 110


Val Ile Tyr Val Ala Arg Leu Phe Trp His Met Asp Gln Gln Asn Arg
        115                 120                 125


Gln Gly Leu Arg Asp Leu Val Asn Ser Gly Val Thr Ile Gln Ile Met
    130                 135                 140


Arg Ala Ser Glu Tyr Tyr His Cys Trp Arg Asn Phe Val Asn Tyr Pro
145                 150                 155                 160


Pro Gly Asp Glu Ala His Trp Pro Gln Tyr Pro Pro Leu Trp Met Met
                165                 170                 175
```

```
Leu Tyr Ala Leu Glu Leu His Cys Ile Ile Leu Ser Leu Pro Pro Cys
            180                 185                 190


Leu Lys Ile Ser Arg Arg Trp Gln Asn His Leu Thr Phe Phe Arg Leu
        195                 200                 205


His Leu Gln Asn Cys His Tyr Gln Thr Ile Pro Pro His Ile Leu Leu
    210                 215                 220


Ala Thr Gly Leu Ile His Pro Ser Val Ala Trp Arg
225                 230                 235


<210>  987
<211>  236
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  987

Met Ala Ser Glu Lys Gly Pro Ser Asn Lys Asp Tyr Thr Leu Arg Arg
1               5                   10                  15


Arg Ile Glu Pro Trp Glu Phe Glu Val Phe Phe Asp Pro Gln Glu Leu
            20                  25                  30


Arg Lys Glu Ala Cys Leu Leu Tyr Glu Ile Lys Trp Gly Ala Ser Ser
        35                  40                  45


Lys Thr Trp Arg Ser Ser Gly Lys Asn Thr Thr Asn His Val Glu Val
    50                  55                  60


Asn Phe Leu Glu Lys Leu Thr Ser Glu Gly Arg Leu Gly Pro Ser Thr
65                  70                  75                  80


Cys Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Trp Glu Cys
                85                  90                  95
```

```
Ser Met Ala Ile Arg Glu Phe Leu Ser Gln His Pro Gly Val Thr Leu
            100                 105                 110

Ile Ile Phe Val Ala Arg Leu Phe Gln His Met Asp Arg Arg Asn Arg
        115                 120                 125

Gln Gly Leu Lys Asp Leu Val Thr Ser Gly Val Thr Val Arg Val Met
    130                 135                 140

Ser Val Ser Glu Tyr Cys Tyr Cys Trp Glu Asn Phe Val Asn Tyr Pro
145                 150                 155                 160

Pro Gly Lys Ala Ala Gln Trp Pro Arg Tyr Pro Pro Arg Trp Met Leu
                165                 170                 175

Met Tyr Ala Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
            180                 185                 190

Leu Lys Ile Ser Arg Arg His Gln Lys Gln Leu Thr Phe Phe Ser Leu
        195                 200                 205

Thr Pro Gln Tyr Cys His Tyr Lys Met Ile Pro Pro Tyr Ile Leu Leu
    210                 215                 220

Ala Thr Gly Leu Leu Gln Pro Ser Val Pro Trp Arg
225                 230                 235


<210>  988
<211>  278
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  988

Met Lys Val Ser Leu Ala Gly Gln Thr Val Asp Val Lys Lys Ile Leu
1               5                   10                  15
```

```
Asn Glu Ile Pro Lys Arg Thr Val Thr Ala Ala Leu Leu Glu Gly Gly
            20                  25                  30

Glu Ile Val Ala Val Glu Glu Ala Asp Asp Glu His Ala Glu Arg Lys
        35                  40                  45

Leu Val Arg Arg His Asp Val Glu Gly Lys Val Val Phe Val Thr Ala
    50                  55                  60

Arg Pro Cys Leu Tyr Cys Ala Arg Glu Leu Ala Glu Ala Gly Val Ala
65                  70                  75                  80

Gly Val Val Tyr Leu Gly Arg Gly Arg Gly Leu Gly Pro Tyr Tyr Leu
                85                  90                  95

Ala Arg Ser Gly Val Glu Val Val Glu Val His Pro Asp Glu Pro Leu
            100                 105                 110

Gly Tyr Asp Pro Val Asp Arg Leu Asp Val Leu Leu Thr Phe Gly Gly
        115                 120                 125

Asn Pro Tyr Leu Thr Glu Glu Asp Val Ala Ala Arg Val Tyr Cys Leu
    130                 135                 140

Leu Thr Gly Arg Gly Phe Asp Ala Asp Ile Ala Pro Ala Pro Glu Asn
145                 150                 155                 160

Leu Ser Gly Arg Val Glu Ile Met Val Thr Arg Gly Asp Pro Asp Glu
                165                 170                 175

Ala Val Glu Leu Leu Lys Glu Glu Leu Pro Val Phe Arg Ile Arg Arg
            180                 185                 190

Phe Leu Ile Ser Gly Glu Phe Asp Arg Asp Glu Leu Arg Glu Arg Ile
        195                 200                 205

Leu Glu Asp Ile Glu Pro Arg Ile Leu Asp Pro Phe Ala Val Arg Ala
    210                 215                 220
```

```
Arg Ile Ala Arg Ala Gly Ala Phe Ser Ser Ser Arg Glu Ala Glu Val
225                 230                 235                 240


Phe Ile Gly Asp Val Leu Thr Ser Val Gly Arg Glu Val Asn Leu Asn
                245                 250                 255


Asp Pro Arg Thr Val Val Thr Val Asp Val Leu Gly Pro Arg Val Ser
            260                 265                 270


Val Gly Val Glu Lys Arg
        275


<210>  989
<211>  294
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  989

Met His Pro Arg Phe Gln Thr Ala Phe Ala Gln Leu Ala Asp Asn Leu
1               5                   10                  15


Gln Ser Ala Leu Glu Pro Ile Leu Ala Asp Lys Tyr Phe Pro Ala Leu
            20                  25                  30


Leu Thr Gly Glu Gln Val Ser Ser Leu Lys Ser Ala Thr Gly Leu Asp
        35                  40                  45


Glu Asp Ala Leu Ala Phe Ala Leu Leu Pro Leu Ala Ala Ala Cys Ala
    50                  55                  60


Arg Thr Pro Leu Ser Asn Phe Asn Val Gly Ala Ile Ala Arg Gly Val
65                  70                  75                  80


Ser Gly Thr Trp Tyr Phe Gly Ala Asn Met Glu Phe Ile Gly Ala Thr
                85                  90                  95
```

```
Met Gln Gln Thr Val His Ala Glu Gln Ser Ala Ile Ser His Ala Trp
            100                 105                 110

Leu Ser Gly Glu Lys Ala Leu Ala Ala Ile Thr Val Asn Tyr Thr Pro
        115                 120                 125

Cys Gly His Cys Arg Gln Phe Met Asn Glu Leu Asn Ser Gly Leu Asp
    130                 135                 140

Leu Arg Ile His Leu Pro Gly Arg Glu Ala His Ala Leu Arg Asp Tyr
145                 150                 155                 160

Leu Pro Asp Ala Phe Gly Pro Lys Asp Leu Glu Ile Lys Thr Leu Leu
                165                 170                 175

Met Asp Glu Gln Asp His Gly Tyr Ala Leu Thr Gly Asp Ala Leu Ser
            180                 185                 190

Gln Ala Ala Ile Ala Ala Ala Asn Arg Ser His Met Pro Tyr Ser Lys
        195                 200                 205

Ser Pro Ser Gly Val Ala Leu Glu Cys Lys Asp Gly Arg Ile Phe Ser
    210                 215                 220

Gly Ser Tyr Ala Glu Asn Ala Ala Phe Asn Pro Thr Leu Pro Pro Leu
225                 230                 235                 240

Gln Gly Ala Leu Ile Leu Leu Asn Leu Lys Gly Tyr Asp Tyr Pro Asp
                245                 250                 255

Ile Gln Arg Ala Val Leu Ala Glu Lys Ala Asp Ala Pro Leu Ile Gln
            260                 265                 270

Trp Asp Ala Thr Ser Ala Thr Leu Lys Ala Leu Gly Cys His Ser Ile
        275                 280                 285
```

Asp Arg Val Leu Leu Ala
290

<210> 990
<211> 295
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 990

Met Arg Asn Arg Ile Glu Gln Ala Leu Gln Gln Met Pro Ala Ser Phe
1 5 10 15

Ala Pro Tyr Leu Arg Glu Leu Val Leu Ala Lys Asp Phe Asp Ala Thr
20 25 30

Phe Ser Ala Glu Gln Tyr Gln Gln Leu Leu Thr Leu Ser Gly Leu Glu
35 40 45

Asp Ala Asp Leu Arg Val Ala Leu Leu Pro Ile Ala Ala Ala Tyr Ser
50 55 60

Tyr Ala Pro Ile Ser Glu Phe Tyr Val Gly Ala Ile Val Arg Gly Ile
65 70 75 80

Ser Gly Arg Leu Tyr Leu Gly Ala Asn Met Glu Phe Thr Gly Ala Gln
85 90 95

Leu Gly Gln Thr Val His Ala Glu Gln Cys Ala Ile Ser His Ala Trp
100 105 110

Met Lys Gly Glu Lys Gly Val Ala Asp Ile Thr Ile Asn Phe Ser Pro
115 120 125

Cys Gly His Cys Arg Gln Phe Met Asn Glu Leu Thr Thr Ala Ser Ser
130 135 140

```
Leu Lys Ile Gln Leu Pro Lys Arg Ala Ala Lys Thr Leu Gln Glu Tyr
145                 150                 155                 160

Leu Pro Glu Ser Phe Gly Pro Ala Asp Leu Gly Ile Asp Ser Gly Leu
                165                 170                 175

Met Ser Pro Val Asn His Gly Lys Thr Ser Asp Asp Asp Glu Glu Leu
            180                 185                 190

Ile Gln Gln Ala Leu Arg Ala Met Asn Ile Ser His Ser Pro Tyr Thr
        195                 200                 205

Gln Asn Phe Ser Gly Val Ala Leu Lys Met Arg Ser Gly Ala Ile Tyr
    210                 215                 220

Leu Gly Ala Tyr Ala Glu Asn Ala Ala Phe Asn Pro Ser Leu Pro Pro
225                 230                 235                 240

Leu Gln Val Ala Leu Ala Gln Ala Met Met Met Gly Glu Ser Phe Glu
                245                 250                 255

Asp Ile Glu Ala Ala Ala Leu Val Glu Ser Ala Thr Gly Lys Ile Ser
            260                 265                 270

His Leu Ala Asp Thr Gln Ala Thr Leu Glu Val Ile Asn Pro Asp Ile
        275                 280                 285

Pro Leu Ser Tyr Leu Ser Leu
    290                 295
```

<210> 991
<211> 297
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 991

```
Met Ser Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile Val Arg Phe
1               5                  10                  15

Leu Ala Thr His His Asn Leu Ser Leu Asp Ile Phe Ser Ser Arg Leu
            20                  25                  30

Tyr Asn Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys Arg Leu Val
        35                  40                  45

Gln Glu Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu Phe Lys Lys
    50                  55                  60

Cys Trp Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe Arg Pro Trp
65                  70                  75                  80

Lys Arg Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys Leu Gln Glu
                85                  90                  95

Ile Leu Arg Arg Met Asp Pro Leu Ser Glu Glu Glu Phe Tyr Ser Gln
            100                 105                 110

Phe Tyr Asn Gln Arg Val Lys His Leu Cys Tyr Tyr His Arg Met Lys
        115                 120                 125

Pro Tyr Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly Gln Ala Pro Leu
    130                 135                 140

Lys Gly Cys Leu Leu Ser Glu Lys Gly Lys Gln His Ala Glu Ile Leu
145                 150                 155                 160

Phe Leu Asp Lys Ile Arg Ser Met Glu Leu Ser Gln Val Thr Ile Thr
                165                 170                 175

Cys Tyr Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala Trp Gln Leu Ala
            180                 185                 190

Ala Phe Lys Arg Asp Arg Pro Asp Leu Ile Leu His Ile Tyr Thr Ser
        195                 200                 205
```

```
Arg Leu Tyr Phe His Trp Lys Arg Pro Phe Gln Lys Gly Leu Cys Ser
    210                 215                 220

Leu Trp Gln Ser Gly Ile Leu Val Asp Val Met Asp Leu Pro Gln Phe
225                 230                 235                 240

Thr Asp Cys Trp Thr Asn Phe Val Asn Pro Lys Arg Pro Phe Trp Pro
                245                 250                 255

Trp Lys Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln Arg Arg Leu Arg
            260                 265                 270

Arg Ile Lys Glu Ser Trp Gly Leu Gln Asp Leu Val Asn Asp Phe Gly
        275                 280                 285

Asn Leu Gln Leu Gly Pro Pro Met Ser
    290                 295
```

<210> 992
<211> 301
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 992

```
Met Asp Lys Pro Ser Phe Val Ile Gln Ser Lys Glu Ala Glu Ser Ala
1               5                   10                  15

Ala Lys Gln Leu Gly Val Ser Val Ile Gln Leu Leu Pro Ser Leu Val
            20                  25                  30

Lys Pro Ala Gln Ser Tyr Ala Arg Thr Pro Ile Ser Lys Phe Asn Val
        35                  40                  45

Ala Val Val Gly Leu Gly Ser Ser Gly Arg Ile Phe Leu Gly Val Asn
    50                  55                  60
```

```
Val Glu Phe Pro Asn Leu Pro Leu His His Ser Ile His Ala Glu Gln
65                  70                  75                  80

Phe Leu Val Thr Asn Leu Thr Leu Asn Gly Glu Arg His Leu Asn Phe
                85                  90                  95

Phe Ala Val Ser Ala Ala Pro Cys Gly His Cys Arg Gln Phe Leu Gln
            100                 105                 110

Glu Ile Arg Asp Ala Pro Glu Ile Lys Ile Leu Ile Thr Asp Pro Asn
        115                 120                 125

Asn Ser Ala Asp Ser Asp Ser Ala Ala Asp Ser Asp Gly Phe Leu Arg
    130                 135                 140

Leu Gly Ser Phe Leu Pro His Arg Phe Gly Pro Asp Asp Leu Leu Gly
145                 150                 155                 160

Lys Asp His Pro Leu Leu Leu Glu Ser His Asp Asn His Leu Lys Ile
                165                 170                 175

Ser Asp Leu Asp Ser Ile Cys Asn Gly Asn Thr Asp Ser Ser Ala Asp
            180                 185                 190

Leu Lys Gln Thr Ala Leu Ala Ala Ala Asn Arg Ser Tyr Ala Pro Tyr
        195                 200                 205

Ser Leu Cys Pro Ser Gly Val Ser Leu Val Asp Cys Asp Gly Lys Val
    210                 215                 220

Tyr Arg Gly Trp Tyr Met Glu Ser Ala Ala Tyr Asn Pro Ser Met Gly
225                 230                 235                 240

Pro Val Gln Ala Ala Leu Val Asp Tyr Val Ala Asn Gly Gly Gly Gly
                245                 250                 255
```

Gly Tyr Glu Arg Ile Val Gly Ala Val Leu Val Glu Lys Glu Asp Ala
260 265 270

Val Val Arg Gln Glu His Thr Ala Arg Leu Leu Leu Glu Thr Ile Ser
275 280 285

Pro Lys Cys Glu Phe Lys Val Phe His Cys Tyr Glu Ala
290 295 300

<210> 993
<211> 370
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 993

Met Val Glu Pro Met Asp Pro Arg Thr Phe Val Ser Asn Phe Asn Asn
1 5 10 15

Arg Pro Ile Leu Ser Gly Leu Asn Thr Val Trp Leu Cys Cys Glu Val
20 25 30

Lys Thr Lys Asp Pro Ser Gly Pro Pro Leu Asp Ala Lys Ile Phe Gln
35 40 45

Gly Lys Val Tyr Ser Lys Ala Lys Tyr His Pro Glu Met Arg Phe Leu
50 55 60

Arg Trp Phe His Lys Trp Arg Gln Leu His His Asp Gln Glu Tyr Lys
65 70 75 80

Val Thr Trp Tyr Val Ser Trp Ser Pro Cys Thr Arg Cys Ala Asn Ser
85 90 95

Val Ala Thr Phe Leu Ala Lys Asp Pro Lys Val Thr Leu Thr Ile Phe
100 105 110

```
Val Ala Arg Leu Tyr Tyr Phe Trp Lys Pro Asp Tyr Gln Gln Ala Leu
        115                 120                 125

Arg Ile Leu Cys Gln Lys Arg Gly Gly Pro His Ala Thr Met Lys Ile
    130                 135                 140

Met Asn Tyr Asn Glu Phe Gln Asp Cys Trp Asn Lys Phe Val Asp Gly
145                 150                 155                 160

Arg Gly Lys Pro Phe Lys Pro Arg Asn Asn Leu Pro Lys His Tyr Thr
                165                 170                 175

Leu Leu Gln Ala Thr Leu Gly Glu Leu Leu Arg His Leu Met Asp Pro
            180                 185                 190

Gly Thr Phe Thr Ser Asn Phe Asn Asn Lys Pro Trp Val Ser Gly Gln
        195                 200                 205

His Glu Thr Tyr Leu Cys Tyr Lys Val Glu Arg Leu His Asn Asp Thr
    210                 215                 220

Trp Val Pro Leu Asn Gln His Arg Gly Phe Leu Arg Asn Gln Ala Pro
225                 230                 235                 240

Asn Ile His Gly Phe Pro Lys Gly Arg His Ala Glu Leu Cys Phe Leu
                245                 250                 255

Asp Leu Ile Pro Phe Trp Lys Leu Asp Gly Gln Gln Tyr Arg Val Thr
            260                 265                 270

Cys Phe Thr Ser Trp Ser Pro Cys Phe Ser Cys Ala Gln Glu Met Ala
        275                 280                 285

Lys Phe Ile Ser Asn Asn Glu His Val Ser Leu Cys Ile Phe Ala Ala
    290                 295                 300

Arg Ile Tyr Asp Asp Gln Gly Arg Tyr Gln Glu Gly Leu Arg Ala Leu
305                 310                 315                 320
```

```
His Arg Asp Gly Ala Lys Ile Ala Met Met Asn Tyr Ser Glu Phe Glu
                325                 330                 335


Tyr Cys Trp Asp Thr Phe Val Asp Arg Gln Gly Arg Pro Phe Gln Pro
            340                 345                 350


Trp Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg
        355                 360                 365


Ala Ile
    370


<210>  994
<211>  373
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  994

Met Lys Pro His Phe Arg Asn Thr Val Glu Arg Met Tyr Arg Asp Thr
1               5                   10                  15


Phe Ser Tyr Asn Phe Tyr Asn Arg Pro Ile Leu Ser Arg Arg Asn Thr
            20                  25                  30


Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Gly Pro Ser Arg Pro Arg
        35                  40                  45


Leu Asp Ala Lys Ile Phe Arg Gly Gln Val Tyr Ser Gln Pro Glu His
    50                  55                  60


His Ala Glu Met Cys Phe Leu Ser Trp Phe Cys Gly Asn Gln Leu Pro
65                  70                  75                  80


Ala Tyr Lys Cys Phe Gln Ile Thr Trp Phe Val Ser Trp Thr Pro Cys
                85                  90                  95
```

```
Pro Asp Cys Val Ala Lys Leu Ala Glu Phe Leu Ala Glu His Pro Asn
            100                 105                 110

Val Thr Leu Thr Ile Ser Ala Ala Arg Leu Tyr Tyr Tyr Trp Glu Arg
        115                 120                 125

Asp Tyr Arg Arg Ala Leu Cys Arg Leu Ser Gln Ala Gly Ala Arg Val
    130                 135                 140

Lys Ile Met Asp Asp Glu Glu Phe Ala Tyr Cys Trp Glu Asn Phe Val
145                 150                 155                 160

Tyr Ser Glu Gly Gln Pro Phe Met Pro Trp Tyr Lys Phe Asp Asp Asn
                165                 170                 175

Tyr Ala Phe Leu His Arg Thr Leu Lys Glu Ile Leu Arg Asn Pro Met
            180                 185                 190

Glu Ala Met Tyr Pro His Ile Phe Tyr Phe His Phe Lys Asn Leu Arg
        195                 200                 205

Lys Ala Tyr Gly Arg Asn Glu Ser Trp Leu Cys Phe Thr Met Glu Val
    210                 215                 220

Val Lys His His Ser Pro Val Ser Trp Lys Arg Gly Val Phe Arg Asn
225                 230                 235                 240

Gln Val Asp Pro Glu Thr His Cys His Ala Glu Arg Cys Phe Leu Ser
                245                 250                 255

Trp Phe Cys Asp Asp Ile Leu Ser Pro Asn Thr Asn Tyr Glu Val Thr
            260                 265                 270

Trp Tyr Thr Ser Trp Ser Pro Cys Pro Glu Cys Ala Gly Glu Val Ala
        275                 280                 285
```

```
Glu Phe Leu Ala Arg His Ser Asn Val Asn Leu Thr Ile Phe Thr Ala
    290                 295                 300

Arg Leu Tyr Tyr Phe Trp Asp Thr Asp Tyr Gln Glu Gly Leu Arg Ser
305                 310                 315                 320

Leu Ser Gln Glu Gly Ala Ser Val Glu Ile Met Gly Tyr Lys Asp Phe
                325                 330                 335

Lys Tyr Cys Trp Glu Asn Phe Val Tyr Asn Asp Asp Glu Pro Phe Lys
            340                 345                 350

Pro Trp Lys Gly Leu Lys Tyr Asn Phe Leu Phe Leu Asp Ser Lys Leu
        355                 360                 365

Gln Glu Ile Leu Glu
    370
```

<210> 995
<211> 377
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 995

```
Met Asn Pro Gln Ile Arg Asn Met Val Glu Gln Met Glu Pro Asp Ile
1               5                   10                  15

Phe Val Tyr Tyr Phe Asn Asn Arg Pro Ile Leu Ser Gly Arg Asn Thr
            20                  25                  30

Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Asp Pro Ser Gly Pro Pro
        35                  40                  45

Leu Asp Ala Asn Ile Phe Gln Gly Lys Leu Tyr Pro Glu Ala Lys Asp
    50                  55                  60
```

```
His Pro Glu Met Lys Phe Leu His Trp Phe Arg Lys Trp Arg Gln Leu
65                  70                  75                  80

His Arg Asp Gln Glu Tyr Glu Val Thr Trp Tyr Val Ser Trp Ser Pro
                85                  90                  95

Cys Thr Arg Cys Ala Asn Ser Val Ala Thr Phe Leu Ala Glu Asp Pro
            100                 105                 110

Lys Val Thr Leu Thr Ile Phe Val Ala Arg Leu Tyr Tyr Phe Trp Lys
        115                 120                 125

Pro Asp Tyr Gln Gln Ala Leu Arg Ile Leu Cys Gln Glu Arg Gly Gly
    130                 135                 140

Pro His Ala Thr Met Lys Ile Met Asn Tyr Asn Glu Phe Gln His Cys
145                 150                 155                 160

Trp Asn Glu Phe Val Asp Gly Gln Gly Lys Pro Phe Lys Pro Arg Lys
                165                 170                 175

Asn Leu Pro Lys His Tyr Thr Leu Leu His Ala Thr Leu Gly Glu Leu
            180                 185                 190

Leu Arg His Val Met Asp Pro Gly Thr Phe Thr Ser Asn Phe Asn Asn
        195                 200                 205

Lys Pro Trp Val Ser Gly Gln Arg Glu Thr Tyr Leu Cys Tyr Lys Val
    210                 215                 220

Glu Arg Ser His Asn Asp Thr Trp Val Leu Leu Asn Gln His Arg Gly
225                 230                 235                 240

Phe Leu Arg Asn Gln Ala Pro Asp Arg His Gly Phe Pro Lys Gly Arg
                245                 250                 255

His Ala Glu Leu Cys Phe Leu Asp Leu Ile Pro Phe Trp Lys Leu Asp
            260                 265                 270
```

```
Asp Gln Gln Tyr Arg Val Thr Cys Phe Thr Ser Trp Ser Pro Cys Phe
        275                 280                 285


Ser Cys Ala Gln Lys Met Ala Lys Phe Ile Ser Asn Asn Lys His Val
    290                 295                 300


Ser Leu Cys Ile Phe Ala Ala Arg Ile Tyr Asp Asp Gln Gly Arg Cys
305                 310                 315                 320


Gln Glu Gly Leu Arg Thr Leu His Arg Asp Gly Ala Lys Ile Ala Val
                325                 330                 335


Met Asn Tyr Ser Glu Phe Glu Tyr Cys Trp Asp Thr Phe Val Asp Arg
            340                 345                 350


Gln Gly Arg Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser Gln
        355                 360                 365


Ala Leu Ser Gly Arg Leu Arg Ala Ile
    370                 375


<210>  996
<211>  382
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  996

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Arg Met Tyr Arg Asp Thr
1               5                   10                  15


Phe Tyr Asp Asn Phe Glu Asn Glu Pro Ile Leu Tyr Gly Arg Ser Tyr
            20                  25                  30


Thr Trp Leu Cys Tyr Glu Val Lys Ile Lys Arg Gly Arg Ser Asn Leu
        35                  40                  45
```

```
Leu Trp Asp Thr Gly Val Phe Arg Gly Gln Val Tyr Phe Lys Pro Gln
    50                  55                  60

Tyr His Ala Glu Met Cys Phe Leu Ser Trp Phe Cys Gly Asn Gln Leu
65                  70                  75                  80

Pro Ala Tyr Lys Cys Phe Gln Ile Thr Trp Phe Val Ser Trp Thr Pro
                85                  90                  95

Cys Pro Asp Cys Val Ala Lys Leu Ala Glu Phe Leu Ser Glu His Pro
            100                 105                 110

Asn Val Thr Leu Thr Ile Ser Ala Ala Arg Leu Tyr Tyr Tyr Trp Glu
        115                 120                 125

Arg Asp Tyr Arg Arg Ala Leu Cys Arg Leu Ser Gln Ala Gly Ala Arg
    130                 135                 140

Val Thr Ile Met Asp Tyr Glu Glu Phe Ala Tyr Cys Trp Glu Asn Phe
145                 150                 155                 160

Val Tyr Asn Glu Gly Gln Gln Phe Met Pro Trp Tyr Lys Phe Asp Glu
                165                 170                 175

Asn Tyr Ala Phe Leu His Arg Thr Leu Lys Glu Ile Leu Arg Tyr Leu
            180                 185                 190

Met Asp Pro Asp Thr Phe Thr Phe Asn Phe Asn Asn Asp Pro Leu Val
        195                 200                 205

Leu Arg Arg Arg Gln Thr Tyr Leu Cys Tyr Glu Val Glu Arg Leu Asp
    210                 215                 220

Asn Gly Thr Trp Val Leu Met Asp Gln His Met Gly Phe Leu Cys Asn
225                 230                 235                 240
```

```
Glu Ala Lys Asn Leu Leu Cys Gly Phe Tyr Gly Arg His Ala Glu Leu
                245                 250                 255

Arg Phe Leu Asp Leu Val Pro Ser Leu Gln Leu Asp Pro Ala Gln Ile
            260                 265                 270

Tyr Arg Val Thr Trp Phe Ile Ser Trp Ser Pro Cys Phe Ser Trp Gly
        275                 280                 285

Cys Ala Gly Glu Val Arg Ala Phe Leu Gln Glu Asn Thr His Val Arg
    290                 295                 300

Leu Arg Ile Phe Ala Ala Arg Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys
305                 310                 315                 320

Glu Ala Leu Gln Met Leu Arg Asp Ala Gly Ala Gln Val Ser Ile Met
                325                 330                 335

Thr Tyr Asp Glu Phe Glu Tyr Cys Trp Asp Thr Phe Val Tyr Arg Gln
            340                 345                 350

Gly Cys Pro Phe Gln Pro Trp Asp Gly Leu Glu Glu His Ser Gln Ala
        355                 360                 365

Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn
    370                 375                 380
```

<210> 997
<211> 383
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 997

```
Met Lys Pro Gln Ile Arg Asn Met Val Glu Pro Met Asp Pro Arg Thr
1               5                   10                  15
```

```
Phe Val Ser Asn Phe Asn Asn Arg Pro Ile Leu Ser Gly Leu Asp Thr
            20                  25                  30

Val Trp Leu Cys Cys Glu Val Lys Thr Lys Asp Pro Ser Gly Pro Pro
        35                  40                  45

Leu Asp Ala Lys Ile Phe Gln Gly Lys Val Tyr Pro Lys Ala Lys Tyr
    50                  55                  60

His Pro Glu Met Arg Phe Leu Arg Trp Phe His Lys Trp Arg Gln Leu
65                  70                  75                  80

His His Asp Gln Glu Tyr Lys Val Thr Trp Tyr Val Ser Trp Ser Pro
                85                  90                  95

Cys Thr Arg Cys Ala Asn Ser Val Ala Thr Phe Leu Ala Lys Asp Pro
            100                 105                 110

Lys Val Thr Leu Thr Ile Phe Val Ala Arg Leu Tyr Tyr Phe Trp Lys
        115                 120                 125

Pro Asp Tyr Gln Gln Ala Leu Arg Ile Leu Cys Gln Lys Arg Asp Gly
    130                 135                 140

Pro His Ala Thr Met Lys Ile Met Asn Tyr Asn Glu Phe Gln Asp Cys
145                 150                 155                 160

Trp Asn Lys Phe Val Asp Gly Arg Gly Lys Pro Phe Lys Pro Trp Asn
                165                 170                 175

Asn Leu Pro Lys His Tyr Thr Leu Leu Gln Ala Thr Leu Gly Glu Leu
            180                 185                 190

Leu Arg His Leu Met Asp Pro Gly Thr Phe Thr Ser Asn Phe Asn Asn
        195                 200                 205

Lys Pro Trp Val Ser Gly Gln His Glu Thr Tyr Leu Cys Tyr Lys Val
    210                 215                 220
```

```
Glu Arg Leu His Asn Asp Thr Trp Val Pro Leu Asn Gln His Arg Gly
225                 230                 235                 240

Phe Leu Arg Asn Gln Ala Pro Asn Ile His Gly Phe Pro Lys Gly Arg
                245                 250                 255

His Ala Glu Leu Cys Phe Leu Asp Leu Ile Pro Phe Trp Lys Leu Asp
            260                 265                 270

Gly Gln Gln Tyr Arg Val Thr Cys Phe Thr Ser Trp Ser Pro Cys Phe
        275                 280                 285

Ser Cys Ala Gln Glu Met Ala Lys Phe Ile Ser Asn Asn Glu His Val
    290                 295                 300

Ser Leu Cys Ile Phe Ala Ala Arg Ile Tyr Asp Asp Gln Gly Arg Tyr
305                 310                 315                 320

Gln Glu Gly Leu Arg Thr Leu His Arg Asp Gly Ala Lys Ile Ala Met
                325                 330                 335

Met Asn Tyr Ser Glu Phe Glu Tyr Cys Trp Asp Thr Phe Val Asp Arg
            340                 345                 350

Gln Gly Arg Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser Gln
        355                 360                 365

Ala Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn
    370                 375                 380
```

<210> 998
<211> 384
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 998

```
Met Lys Pro His Phe Arg Asn Thr Val Glu Arg Met Tyr Arg Asp Thr
1               5                  10                  15

Phe Ser Tyr Asn Phe Tyr Asn Arg Pro Ile Leu Ser Arg Arg Asn Thr
            20                  25                  30

Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Gly Pro Ser Arg Pro Pro
        35                  40                  45

Leu Asp Ala Lys Ile Phe Arg Gly Gln Val Tyr Ser Glu Leu Lys Tyr
    50                  55                  60

His Pro Glu Met Arg Phe Phe His Trp Phe Ser Lys Trp Arg Lys Leu
65                  70                  75                  80

His Arg Asp Gln Glu Tyr Glu Val Thr Trp Tyr Ile Ser Trp Ser Pro
                85                  90                  95

Cys Thr Lys Cys Thr Arg Asp Met Ala Thr Phe Leu Ala Glu Asp Pro
            100                 105                 110

Lys Val Thr Leu Thr Ile Phe Val Ala Arg Leu Tyr Tyr Phe Trp Asp
        115                 120                 125

Pro Asp Tyr Gln Glu Ala Leu Arg Ser Leu Cys Gln Lys Arg Asp Gly
    130                 135                 140

Pro Arg Ala Thr Met Lys Ile Met Asn Tyr Asp Glu Phe Gln His Cys
145                 150                 155                 160

Trp Ser Lys Phe Val Tyr Ser Gln Arg Glu Leu Phe Glu Pro Trp Asn
                165                 170                 175

Asn Leu Pro Lys Tyr Tyr Ile Leu Leu His Ile Met Leu Gly Glu Ile
            180                 185                 190
```

```
Leu Arg His Ser Met Asp Pro Pro Thr Phe Thr Phe Asn Phe Asn Asn
        195                 200                 205

Glu Pro Trp Val Arg Gly Arg His Glu Thr Tyr Leu Cys Tyr Glu Val
    210                 215                 220

Glu Arg Met His Asn Asp Thr Trp Val Leu Leu Asn Gln Arg Arg Gly
225                 230                 235                 240

Phe Leu Cys Asn Gln Ala Pro His Lys His Gly Phe Leu Glu Gly Arg
                245                 250                 255

His Ala Glu Leu Cys Phe Leu Asp Val Ile Pro Phe Trp Lys Leu Asp
            260                 265                 270

Leu Asp Gln Asp Tyr Arg Val Thr Cys Phe Thr Ser Trp Ser Pro Cys
        275                 280                 285

Phe Ser Cys Ala Gln Glu Met Ala Lys Phe Ile Ser Lys Asn Lys His
    290                 295                 300

Val Ser Leu Cys Ile Phe Thr Ala Arg Ile Tyr Asp Asp Gln Gly Arg
305                 310                 315                 320

Cys Gln Glu Gly Leu Arg Thr Leu Ala Glu Ala Gly Ala Lys Ile Ser
                325                 330                 335

Ile Met Thr Tyr Ser Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp
            340                 345                 350

His Gln Gly Cys Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser
        355                 360                 365

Gln Asp Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Glu Asn
    370                 375                 380
```

<210> 999
<211> 384

<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 999

Met Lys Pro His Phe Arg Asn Pro Val Glu Arg Met Tyr Gln Asp Thr
1 5 10 15

Phe Ser Asp Asn Phe Tyr Asn Arg Pro Ile Leu Ser His Arg Asn Thr
20 25 30

Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Gly Pro Ser Arg Pro Pro
35 40 45

Leu Asp Ala Lys Ile Phe Arg Gly Gln Val Tyr Ser Lys Leu Lys Tyr
50 55 60

His Pro Glu Met Arg Phe Phe His Trp Phe Ser Lys Trp Arg Lys Leu
65 70 75 80

His Arg Asp Gln Glu Tyr Glu Val Thr Trp Tyr Ile Ser Trp Ser Pro
85 90 95

Cys Thr Lys Cys Thr Arg Asp Val Ala Thr Phe Leu Ala Glu Asp Pro
100 105 110

Lys Val Thr Leu Thr Ile Phe Val Ala Arg Leu Tyr Tyr Phe Trp Asp
115 120 125

Pro Asp Tyr Gln Glu Ala Leu Arg Ser Leu Cys Gln Lys Arg Asp Gly
130 135 140

Pro Arg Ala Thr Met Lys Ile Met Asn Tyr Asp Glu Phe Gln His Cys
145 150 155 160

Trp Ser Lys Phe Val Tyr Ser Gln Arg Glu Leu Phe Glu Pro Trp Asn
165 170 175

```
Asn Leu Pro Lys Tyr Tyr Ile Leu Leu His Ile Met Leu Gly Glu Ile
            180                 185                 190

Leu Arg His Ser Met Asp Pro Pro Thr Phe Thr Ser Asn Phe Asn Asn
        195                 200                 205

Glu Leu Trp Val Arg Gly Arg His Glu Thr Tyr Leu Cys Tyr Glu Val
    210                 215                 220

Glu Arg Leu His Asn Asp Thr Trp Val Leu Leu Asn Gln Arg Arg Gly
225                 230                 235                 240

Phe Leu Cys Asn Gln Ala Pro His Lys His Gly Phe Leu Glu Gly Arg
                245                 250                 255

His Ala Glu Leu Cys Phe Leu Asp Val Ile Pro Phe Trp Lys Leu Asp
            260                 265                 270

Leu His Gln Asp Tyr Arg Val Thr Cys Phe Thr Ser Trp Ser Pro Cys
        275                 280                 285

Phe Ser Cys Ala Gln Glu Met Ala Lys Phe Ile Ser Asn Asn Lys His
    290                 295                 300

Val Ser Leu Cys Ile Phe Ala Ala Arg Ile Tyr Asp Asp Gln Gly Arg
305                 310                 315                 320

Cys Gln Glu Gly Leu Arg Thr Leu Ala Lys Ala Gly Ala Lys Ile Ser
                325                 330                 335

Ile Met Thr Tyr Ser Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp
            340                 345                 350

His Gln Gly Cys Pro Phe Gln Pro Trp Asp Gly Leu Glu Glu His Ser
        355                 360                 365
```

```
Gln Ala Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn
    370                 375                 380


<210>  1000
<211>  386
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  1000

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Arg Met Tyr Arg Asp Thr
1               5                   10                  15


Phe Tyr Asp Asn Phe Glu Asn Glu Pro Ile Leu Tyr Gly Arg Ser Tyr
            20                  25                  30


Thr Trp Leu Cys Tyr Glu Val Lys Ile Lys Arg Gly Arg Ser Asn Leu
        35                  40                  45


Leu Trp Asp Thr Gly Val Phe Arg Gly Pro Val Leu Pro Lys Arg Gln
    50                  55                  60


Ser Asn His Arg Gln Glu Val Tyr Phe Arg Phe Glu Asn His Ala Glu
65                  70                  75                  80


Met Cys Phe Leu Ser Trp Phe Cys Gly Asn Arg Leu Pro Ala Asn Arg
                85                  90                  95


Arg Phe Gln Ile Thr Trp Phe Val Ser Trp Asn Pro Cys Leu Pro Cys
            100                 105                 110


Val Val Lys Val Thr Lys Phe Leu Ala Glu His Pro Asn Val Thr Leu
        115                 120                 125


Thr Ile Ser Ala Ala Arg Leu Tyr Tyr Tyr Arg Asp Arg Asp Trp Arg
    130                 135                 140
```

```
Trp Val Leu Leu Arg Leu His Lys Ala Gly Ala Arg Val Lys Ile Met
145                 150                 155                 160

Asp Tyr Glu Asp Phe Ala Tyr Cys Trp Glu Asn Phe Val Cys Asn Glu
                165                 170                 175

Gly Gln Pro Phe Met Pro Trp Tyr Lys Phe Asp Asp Asn Tyr Ala Ser
            180                 185                 190

Leu His Arg Thr Leu Lys Glu Ile Leu Arg Asn Pro Met Glu Ala Met
        195                 200                 205

Tyr Pro His Ile Phe Tyr Phe His Phe Lys Asn Leu Leu Lys Ala Cys
    210                 215                 220

Gly Arg Asn Glu Ser Trp Leu Cys Phe Thr Met Glu Val Thr Lys His
225                 230                 235                 240

His Ser Ala Val Phe Arg Lys Arg Gly Val Phe Arg Asn Gln Val Asp
                245                 250                 255

Pro Glu Thr His Cys His Ala Glu Arg Cys Phe Leu Ser Trp Phe Cys
            260                 265                 270

Asp Asp Ile Leu Ser Pro Asn Thr Asn Tyr Glu Val Thr Trp Tyr Thr
        275                 280                 285

Ser Trp Ser Pro Cys Pro Glu Cys Ala Gly Glu Val Ala Glu Phe Leu
    290                 295                 300

Ala Arg His Ser Asn Val Asn Leu Thr Ile Phe Thr Ala Arg Leu Cys
305                 310                 315                 320

Tyr Phe Trp Asp Thr Asp Tyr Gln Glu Gly Leu Cys Ser Leu Ser Gln
                325                 330                 335

Glu Gly Ala Ser Val Lys Ile Met Gly Tyr Lys Asp Phe Val Ser Cys
            340                 345                 350
```

```
Trp Lys Asn Phe Val Tyr Ser Asp Asp Glu Pro Phe Lys Pro Trp Lys
        355                 360                 365


Gly Leu Gln Thr Asn Phe Arg Leu Leu Lys Arg Arg Leu Arg Glu Ile
    370                 375                 380


Leu Gln
385
```

<210> 1001
<211> 392
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1001

```
Met Gln Pro Gln Arg Leu Gly Pro Arg Ala Gly Met Gly Pro Phe Cys
1               5                   10                  15


Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro Ile Arg Asn Leu Ile
            20                  25                  30


Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn Leu Gly Tyr Ala Lys
        35                  40                  45


Gly Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val Thr Arg Lys Asp Cys
    50                  55                  60


Asp Ser Pro Val Ser Leu His His Gly Val Phe Lys Asn Lys Asp Asn
65                  70                  75                  80


Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe His Asp Lys Val Leu
                85                  90                  95


Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile Thr Trp Tyr Met Ser
            100                 105                 110
```

```
Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile Val Arg Phe Leu Ala
        115                 120                 125

Thr His His Asn Leu Ser Leu Asp Ile Phe Ser Ser Arg Leu Tyr Asn
    130                 135                 140

Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys Arg Leu Val Gln Glu
145                 150                 155                 160

Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu Phe Lys Lys Cys Trp
                165                 170                 175

Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe Arg Pro Trp Lys Arg
            180                 185                 190

Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys Leu Gln Glu Ile Leu
        195                 200                 205

Arg Arg Met Asp Pro Leu Ser Glu Glu Glu Phe Tyr Ser Gln Phe Tyr
    210                 215                 220

Asn Gln Arg Val Lys His Leu Cys Tyr Tyr His Arg Met Lys Pro Tyr
225                 230                 235                 240

Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly Gln Ala Pro Leu Lys Gly
                245                 250                 255

Cys Leu Leu Ser Glu Lys Gly Lys Gln His Ala Glu Ile Leu Phe Leu
            260                 265                 270

Asp Lys Ile Arg Ser Met Glu Leu Ser Gln Val Thr Ile Thr Cys Tyr
        275                 280                 285

Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala Trp Gln Leu Ala Ala Phe
    290                 295                 300
```

```
Lys Arg Asp Arg Pro Asp Leu Ile Leu His Ile Tyr Thr Ser Arg Leu
305                 310                 315                 320


Tyr Phe His Trp Lys Arg Pro Phe Gln Lys Gly Leu Cys Ser Leu Trp
                325                 330                 335


Gln Ser Gly Ile Leu Val Asp Val Met Asp Leu Pro Gln Phe Thr Asp
            340                 345                 350


Cys Trp Thr Asn Phe Val Asn Pro Lys Arg Pro Phe Trp Pro Trp Lys
        355                 360                 365


Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln Arg Arg Leu Arg Arg Ile
    370                 375                 380


Lys Glu Ser Arg Ser His Ala Ser
385                 390


<210>  1002
<211>  407
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  1002

Met Gln Pro Gln Arg Leu Gly Pro Arg Ala Gly Met Gly Pro Phe Cys
1               5                   10                  15


Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro Ile Arg Asn Leu Ile
            20                  25                  30


Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn Leu Gly Tyr Ala Lys
        35                  40                  45


Gly Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val Thr Arg Lys Asp Cys
    50                  55                  60
```

```
Asp Ser Pro Val Ser Leu His His Gly Val Phe Lys Asn Lys Asp Asn
65                  70                  75                  80

Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe His Asp Lys Val Leu
                85                  90                  95

Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile Thr Trp Tyr Met Ser
            100                 105                 110

Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile Val Arg Phe Leu Ala
        115                 120                 125

Thr His His Asn Leu Ser Leu Asp Ile Phe Ser Ser Arg Leu Tyr Asn
    130                 135                 140

Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys Arg Leu Val Gln Glu
145                 150                 155                 160

Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu Phe Lys Lys Cys Trp
                165                 170                 175

Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe Arg Pro Trp Lys Arg
            180                 185                 190

Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys Leu Gln Glu Ile Leu
        195                 200                 205

Arg Arg Met Asp Pro Leu Ser Glu Glu Glu Phe Tyr Ser Gln Phe Tyr
    210                 215                 220

Asn Gln Arg Val Lys His Leu Cys Tyr Tyr His Arg Met Lys Pro Tyr
225                 230                 235                 240

Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly Gln Ala Pro Leu Lys Gly
                245                 250                 255

Cys Leu Leu Ser Glu Lys Gly Lys Gln His Ala Glu Ile Leu Phe Leu
            260                 265                 270
```

```
Asp Lys Ile Arg Ser Met Glu Leu Ser Gln Val Thr Ile Thr Cys Tyr
        275                 280                 285

Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala Trp Gln Leu Ala Ala Phe
    290                 295                 300

Lys Arg Asp Arg Pro Asp Leu Ile Leu His Ile Tyr Thr Ser Arg Leu
305                 310                 315                 320

Tyr Phe His Trp Lys Arg Pro Phe Gln Lys Gly Leu Cys Ser Leu Trp
                325                 330                 335

Gln Ser Gly Ile Leu Val Asp Val Met Asp Leu Pro Gln Phe Thr Asp
            340                 345                 350

Cys Trp Thr Asn Phe Val Asn Pro Lys Arg Pro Phe Trp Pro Trp Lys
        355                 360                 365

Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln Arg Arg Leu Arg Arg Ile
    370                 375                 380

Lys Glu Ser Trp Gly Leu Gln Asp Leu Val Asn Asp Phe Gly Asn Leu
385                 390                 395                 400

Gln Leu Gly Pro Pro Met Ser
                405
```

<210> 1003
<211> 429
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1003

```
Met Gly Pro Phe Cys Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro
1               5                   10                  15
```

```
Ile Arg Asn Leu Ile Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn
            20                  25                  30

Leu Gly Tyr Ala Lys Gly Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val
        35                  40                  45

Thr Arg Lys Asp Cys Asp Ser Pro Val Ser Leu His His Gly Val Phe
    50                  55                  60

Lys Asn Lys Asp Asn Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe
65                  70                  75                  80

His Asp Lys Val Leu Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile
                85                  90                  95

Thr Trp Tyr Met Ser Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile
            100                 105                 110

Val Arg Phe Leu Ala Thr His His Asn Leu Ser Leu Asp Ile Phe Ser
        115                 120                 125

Ser Arg Leu Tyr Asn Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys
    130                 135                 140

Arg Leu Val Gln Glu Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu
145                 150                 155                 160

Phe Lys Lys Cys Trp Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe
                165                 170                 175

Arg Pro Trp Lys Arg Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys
            180                 185                 190

Leu Gln Glu Ile Leu Arg Pro Cys Tyr Ile Ser Val Pro Ser Ser Ser
        195                 200                 205
```

```
Ser Ser Thr Leu Ser Asn Ile Cys Leu Thr Lys Gly Leu Pro Glu Thr
    210                 215                 220

Arg Phe Trp Val Glu Gly Arg Arg Met Asp Pro Leu Ser Glu Glu Glu
225                 230                 235                 240

Phe Tyr Ser Gln Phe Tyr Asn Gln Arg Val Lys His Leu Cys Tyr Tyr
                245                 250                 255

His Arg Met Lys Pro Tyr Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly
            260                 265                 270

Gln Ala Pro Leu Lys Gly Cys Leu Leu Ser Glu Lys Gly Lys Gln His
        275                 280                 285

Ala Glu Ile Leu Phe Leu Asp Lys Ile Arg Ser Met Glu Leu Ser Gln
    290                 295                 300

Val Thr Ile Thr Cys Tyr Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala
305                 310                 315                 320

Trp Gln Leu Ala Ala Phe Lys Arg Asp Arg Pro Asp Leu Ile Leu His
                325                 330                 335

Ile Tyr Thr Ser Arg Leu Tyr Phe His Trp Lys Arg Pro Phe Gln Lys
            340                 345                 350

Gly Leu Cys Ser Leu Trp Gln Ser Gly Ile Leu Val Asp Val Met Asp
        355                 360                 365

Leu Pro Gln Phe Thr Asp Cys Trp Thr Asn Phe Val Asn Pro Lys Arg
    370                 375                 380

Pro Phe Trp Pro Trp Lys Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln
385                 390                 395                 400

Arg Arg Leu Arg Arg Ile Lys Glu Ser Trp Gly Leu Gln Asp Leu Val
                405                 410                 415
```

Asn Asp Phe Gly Asn Leu Gln Leu Gly Pro Pro Met Ser
420 425

<210> 1004
<211> 429
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1004

Met Gly Pro Phe Cys Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro
1 5 10 15

Ile Arg Asn Leu Ile Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn
20 25 30

Leu Gly Tyr Ala Lys Gly Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val
35 40 45

Thr Arg Lys Asp Cys Asp Ser Pro Val Ser Leu His His Gly Val Phe
50 55 60

Lys Asn Lys Asp Asn Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe
65 70 75 80

His Asp Lys Val Leu Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile
85 90 95

Thr Trp Tyr Met Ser Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile
100 105 110

Val Arg Phe Leu Ala Thr His His Asn Leu Ser Leu Asp Ile Phe Ser
115 120 125

Ser Arg Leu Tyr Asn Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys
130 135 140

```
Arg Leu Val Gln Glu Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu
145                 150                 155                 160

Phe Lys Lys Cys Trp Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe
                165                 170                 175

Arg Pro Trp Lys Arg Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys
            180                 185                 190

Leu Gln Glu Ile Leu Arg Pro Cys Tyr Ile Pro Val Pro Ser Ser Ser
        195                 200                 205

Ser Ser Thr Leu Ser Asn Ile Cys Leu Thr Lys Gly Leu Pro Glu Thr
    210                 215                 220

Arg Phe Cys Val Glu Gly Arg Arg Met Asp Pro Leu Ser Glu Glu Glu
225                 230                 235                 240

Phe Tyr Ser Gln Phe Tyr Asn Gln Arg Val Lys His Leu Cys Tyr Tyr
                245                 250                 255

His Arg Met Lys Pro Tyr Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly
            260                 265                 270

Gln Ala Pro Leu Lys Gly Cys Leu Leu Ser Glu Lys Gly Lys Gln His
        275                 280                 285

Ala Glu Ile Leu Phe Leu Asp Lys Ile Arg Ser Met Glu Leu Ser Gln
    290                 295                 300

Val Thr Ile Thr Cys Tyr Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala
305                 310                 315                 320

Trp Gln Leu Ala Ala Phe Lys Arg Asp Arg Pro Asp Leu Ile Leu His
                325                 330                 335
```

```
Ile Tyr Thr Ser Arg Leu Tyr Phe His Trp Lys Arg Pro Phe Gln Lys
            340                 345                 350

Gly Leu Cys Ser Leu Trp Gln Ser Gly Ile Leu Val Asp Val Met Asp
        355                 360                 365

Leu Pro Gln Phe Thr Asp Cys Trp Thr Asn Phe Val Asn Pro Lys Arg
    370                 375                 380

Pro Phe Trp Pro Trp Lys Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln
385                 390                 395                 400

Arg Arg Leu Arg Arg Ile Lys Glu Ser Trp Gly Leu Gln Asp Leu Val
                405                 410                 415

Asn Asp Phe Gly Asn Leu Gln Leu Gly Pro Pro Met Ser
            420                 425
```

<210> 1005
<211> 430
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1005

```
Met Gln Pro Gln Arg Leu Gly Pro Arg Ala Gly Met Gly Pro Phe Cys
1               5                   10                  15

Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro Ile Arg Asn Leu Ile
            20                  25                  30

Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn Leu Gly Tyr Ala Lys
        35                  40                  45

Gly Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val Thr Arg Lys Asp Cys
    50                  55                  60
```

```
Asp Ser Pro Val Ser Leu His His Gly Val Phe Lys Asn Lys Asp Asn
65                  70                  75                  80

Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe His Asp Lys Val Leu
                85                  90                  95

Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile Thr Trp Tyr Met Ser
            100                 105                 110

Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile Val Arg Phe Leu Ala
        115                 120                 125

Thr His His Asn Leu Ser Leu Asp Ile Phe Ser Ser Arg Leu Tyr Asn
    130                 135                 140

Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys Arg Leu Val Gln Glu
145                 150                 155                 160

Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu Phe Lys Lys Cys Trp
                165                 170                 175

Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe Arg Pro Trp Lys Arg
            180                 185                 190

Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys Leu Gln Glu Ile Leu
        195                 200                 205

Arg Pro Cys Tyr Ile Ser Val Pro Ser Ser Ser Ser Ser Thr Leu Ser
    210                 215                 220

Asn Ile Cys Leu Thr Lys Gly Leu Pro Glu Thr Arg Phe Trp Val Glu
225                 230                 235                 240

Gly Arg Arg Met Asp Pro Leu Ser Glu Glu Glu Phe Tyr Ser Gln Phe
                245                 250                 255

Tyr Asn Gln Arg Val Lys His Leu Cys Tyr Tyr His Arg Met Lys Pro
            260                 265                 270
```

```
Tyr Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly Gln Ala Pro Leu Lys
        275                 280                 285

Gly Cys Leu Leu Ser Glu Lys Gly Lys Gln His Ala Glu Ile Leu Phe
    290                 295                 300

Leu Asp Lys Ile Arg Ser Met Glu Leu Ser Gln Val Thr Ile Thr Cys
305                 310                 315                 320

Tyr Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala Trp Gln Leu Ala Ala
                325                 330                 335

Phe Lys Arg Asp Arg Pro Asp Leu Ile Leu His Ile Tyr Thr Ser Arg
            340                 345                 350

Leu Tyr Phe His Trp Lys Arg Pro Phe Gln Lys Gly Leu Cys Ser Leu
        355                 360                 365

Trp Gln Ser Gly Ile Leu Val Asp Val Met Asp Leu Pro Gln Phe Thr
    370                 375                 380

Asp Cys Trp Thr Asn Phe Val Asn Pro Lys Arg Pro Phe Trp Pro Trp
385                 390                 395                 400

Lys Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln Arg Arg Leu Arg Arg
                405                 410                 415

Ile Lys Glu Val Arg Thr Thr Leu Leu Gln Gly Pro Ala Ser
            420                 425                 430
```

<210> 1006
<211> 440
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1006

Met Gln Pro Gln Arg Leu Gly Pro Arg Ala Gly Met Gly Pro Phe Cys
1 5 10 15

Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro Ile Arg Asn Leu Ile
20 25 30

Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn Leu Gly Tyr Ala Lys
35 40 45

Gly Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val Thr Arg Lys Asp Cys
50 55 60

Asp Ser Pro Val Ser Leu His His Gly Val Phe Lys Asn Lys Asp Asn
65 70 75 80

Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe His Asp Lys Val Leu
85 90 95

Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile Thr Trp Tyr Met Ser
100 105 110

Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile Val Arg Phe Leu Ala
115 120 125

Thr His His Asn Leu Ser Leu Asp Ile Phe Ser Ser Arg Leu Tyr Asn
130 135 140

Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys Arg Leu Val Gln Glu
145 150 155 160

Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu Phe Lys Lys Cys Trp
165 170 175

Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe Arg Pro Trp Lys Arg
180 185 190

```
Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys Leu Gln Glu Ile Leu
        195                 200                 205

Arg Pro Cys Tyr Ile Ser Val Pro Ser Ser Ser Ser Ser Thr Leu Ser
    210                 215                 220

Asn Ile Cys Leu Thr Lys Gly Leu Pro Glu Thr Arg Phe Trp Val Glu
225                 230                 235                 240

Gly Arg Arg Met Asp Pro Leu Ser Glu Glu Glu Phe Tyr Ser Gln Phe
                245                 250                 255

Tyr Asn Gln Arg Val Lys His Leu Cys Tyr Tyr His Arg Met Lys Pro
            260                 265                 270

Tyr Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly Gln Ala Pro Leu Lys
        275                 280                 285

Gly Cys Leu Leu Ser Glu Lys Gly Lys Gln His Ala Glu Ile Leu Phe
    290                 295                 300

Leu Asp Lys Ile Arg Ser Met Glu Leu Ser Gln Val Thr Ile Thr Cys
305                 310                 315                 320

Tyr Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala Trp Gln Leu Ala Ala
                325                 330                 335

Phe Lys Arg Asp Arg Pro Asp Leu Ile Leu His Ile Tyr Thr Ser Arg
            340                 345                 350

Leu Tyr Phe His Trp Lys Arg Pro Phe Gln Lys Gly Leu Cys Ser Leu
        355                 360                 365

Trp Gln Ser Gly Ile Leu Val Asp Val Met Asp Leu Pro Gln Phe Thr
    370                 375                 380

Asp Cys Trp Thr Asn Phe Val Asn Pro Lys Arg Pro Phe Trp Pro Trp
385                 390                 395                 400
```

Lys Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln Arg Arg Leu Arg Arg
405 410 415

Ile Lys Glu Ser Trp Gly Leu Gln Asp Leu Val Asn Asp Phe Gly Asn
420 425 430

Leu Gln Leu Gly Pro Pro Met Ser
435 440

<210> 1007
<211> 525
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1007

Met Lys Glu Thr Asp Gln Met Gln Ser Leu Glu Gly Ser Gly Ala Glu
1 5 10 15

Arg Ser Val Gly Thr Gln Thr Gly Ser Met Thr Gly Gln Ile Pro Arg
20 25 30

Leu Ser Lys Val Asn Leu Phe Thr Leu Leu Ser Leu Trp Met Glu Leu
35 40 45

Phe Pro Gly Val Glu Ala Gln Gly Gln Lys Ser Gln Lys Thr Glu Glu
50 55 60

Glu Ser Arg Gly Pro Leu Gly Asp Asn Glu Glu Leu Thr Arg Val Ser
65 70 75 80

Thr Glu Lys Lys Gln Val Lys Lys Thr Gly Leu Val Val Val Lys Asn
85 90 95

Met Lys Ile Ile Gly Leu His Cys Ser Ser Glu Asp Leu His Thr Gly
100 105 110

```
Gln Ile Ala Leu Ile Lys His Gly Ser Arg Leu Lys Asn Cys Asp Leu
        115                 120                 125

Tyr Phe Ser Arg Lys Pro Cys Ser Ala Cys Leu Lys Met Ile Val Asn
    130                 135                 140

Ala Gly Val Asn Arg Ile Ser Tyr Trp Pro Ser Asp Pro Glu Ile Ser
145                 150                 155                 160

Leu Leu Thr Glu Ala Ser Ser Ser Glu Asp Ala Lys Leu Asp Ala Lys
                165                 170                 175

Ala Ala Glu Arg Leu Lys Ser Asn Ser Arg Ala His Val Cys Val Leu
            180                 185                 190

Leu Gln Pro Leu Val Cys Tyr Met Val Gln Phe Val Glu Glu Thr Ser
        195                 200                 205

Tyr Lys Cys Asp Phe Ile Gln Lys Thr Ala Lys Ala Leu Pro Gly Ala
    210                 215                 220

Asp Thr Asp Phe Tyr Ser Glu Cys Lys Gln Glu Arg Ile Lys Glu Tyr
225                 230                 235                 240

Glu Met Leu Phe Leu Val Ser Asn Glu Glu Arg His Lys Gln Ile Leu
                245                 250                 255

Met Thr Ile Gly Leu Glu Ser Leu Cys Glu Asp Pro Tyr Phe Ser Asn
            260                 265                 270

Leu Arg Gln Asn Met Lys Asp Leu Ile Leu Leu Leu Ala Thr Val Ala
        275                 280                 285

Ser Ser Val Pro Asn Leu Lys His Phe Gly Phe Tyr Cys Ser Ser Pro
    290                 295                 300
```

```
Glu Gln Ile Asn Glu Ile His Asn Gln Ser Leu Pro Gln Glu Val Ala
305                 310                 315                 320

Arg His Cys Met Val Gln Ala Arg Leu Leu Ala Tyr Arg Thr Glu Asp
                325                 330                 335

His Lys Thr Gly Val Gly Ala Val Ile Trp Ala Glu Ala Lys Ser Arg
            340                 345                 350

Ser Cys Asp Gly Thr Gly Ala Met Tyr Phe Ile Gly Cys Gly Tyr Asn
        355                 360                 365

Ala Phe Pro Val Gly Ser Glu Tyr Ala Asp Phe Pro His Met Asp Asp
    370                 375                 380

Lys His Lys Asp Arg Glu Ile Arg Lys Phe Arg Tyr Ile Ile His Ala
385                 390                 395                 400

Glu Gln Asn Ala Leu Thr Phe Arg Cys Gln Asp Ile Lys Pro Glu Glu
                405                 410                 415

Arg Ser Met Ile Phe Val Thr Lys Cys Pro Cys Asp Glu Cys Val Pro
            420                 425                 430

Leu Ile Lys Gly Ala Gly Ile Lys Gln Ile Tyr Ala Gly Asp Val Asp
        435                 440                 445

Val Gly Lys Lys Lys Ala Asp Ile Ser Tyr Met Lys Phe Gly Glu Leu
    450                 455                 460

Glu Gly Val Arg Lys Phe Thr Trp Gln Leu Asn Pro Ser Glu Ala Tyr
465                 470                 475                 480

Ser Leu Asp Pro Asn Glu Pro Glu Arg Arg Glu Asn Gly Val Leu Arg
                485                 490                 495

Arg Arg Ser Ala Lys Asp Glu Gln Arg Ser Ser Lys Arg Pro Arg Leu
            500                 505                 510
```

```
Glu Thr Arg Ser Ala Gly Ser Ala Thr Thr Ala Cys Phe
        515                 520                 525


<210>  1008
<211>  427
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  1008

Met Ser Asn Asn Ala Leu Gln Thr Ile Ile Asn Ala Arg Leu Pro Gly
1               5                   10                  15


Glu Glu Gly Leu Trp Gln Ile His Leu Gln Asp Gly Lys Ile Ser Ala
            20                  25                  30


Ile Asp Ala Gln Ser Gly Val Met Pro Ile Thr Glu Asn Ser Leu Asp
        35                  40                  45


Ala Glu Gln Gly Leu Val Ile Pro Pro Phe Val Glu Pro His Ile His
    50                  55                  60


Leu Asp Thr Thr Gln Thr Ala Gly Gln Pro Asn Trp Asn Gln Ser Gly
65                  70                  75                  80


Thr Leu Phe Glu Gly Ile Glu Arg Trp Ala Glu Arg Lys Ala Leu Leu
                85                  90                  95


Thr His Asp Asp Val Lys Gln Arg Ala Trp Gln Thr Leu Lys Trp Gln
            100                 105                 110


Ile Ala Asn Gly Ile Gln His Val Arg Thr His Val Asp Val Ser Asp
        115                 120                 125


Ala Thr Leu Thr Ala Leu Lys Ala Met Leu Glu Val Lys Gln Glu Val
    130                 135                 140
```

```
Ala Pro Trp Ile Asp Leu Gln Ile Val Ala Phe Pro Gln Glu Gly Ile
145                 150                 155                 160

Leu Ser Tyr Pro Asn Gly Glu Ala Leu Leu Glu Glu Ala Leu Arg Leu
                165                 170                 175

Gly Ala Asp Val Val Gly Ala Ile Pro His Phe Glu Phe Thr Arg Glu
            180                 185                 190

Tyr Gly Val Glu Ser Leu His Lys Thr Phe Ala Leu Ala Gln Lys Tyr
        195                 200                 205

Asp Arg Leu Ile Asp Val His Cys Asp Glu Ile Asp Asp Glu Gln Ser
    210                 215                 220

Arg Phe Val Glu Thr Val Ala Ala Leu Ala His His Glu Gly Met Gly
225                 230                 235                 240

Ala Arg Val Thr Ala Ser His Thr Thr Ala Met His Ser Tyr Asn Gly
                245                 250                 255

Ala Tyr Thr Ser Arg Leu Phe Arg Leu Leu Lys Met Ser Gly Ile Asn
            260                 265                 270

Phe Val Ala Asn Pro Leu Val Asn Ile His Leu Gln Gly Arg Phe Asp
        275                 280                 285

Thr Tyr Pro Lys Arg Arg Gly Ile Thr Arg Val Lys Glu Met Leu Glu
    290                 295                 300

Ser Gly Ile Asn Val Cys Phe Gly His Asp Asp Val Phe Asp Pro Trp
305                 310                 315                 320

Tyr Pro Leu Gly Thr Ala Asn Met Leu Gln Val Leu His Met Gly Leu
                325                 330                 335
```

```
His Val Cys Gln Leu Met Gly Tyr Gly Gln Ile Asn Asp Gly Leu Asn
            340                 345                 350

Leu Ile Thr His His Ser Ala Arg Thr Leu Asn Leu Gln Asp Tyr Gly
        355                 360                 365

Ile Ala Ala Gly Asn Ser Ala Asn Leu Ile Ile Leu Pro Ala Glu Asn
    370                 375                 380

Gly Phe Asp Ala Leu Arg Arg Gln Val Pro Val Arg Tyr Ser Val Arg
385                 390                 395                 400

Gly Gly Lys Val Ile Ala Ser Thr Gln Pro Ala Gln Thr Thr Val Tyr
                405                 410                 415

Leu Glu Gln Pro Glu Ala Ile Asp Tyr Lys Arg
            420                 425
```

<210> 1009
<211> 210
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1009

```
Met Ala Leu Leu Thr Ala Lys Thr Phe Ser Leu Gln Phe Asn Asn Lys
1               5                   10                  15

Arg Arg Val Asn Lys Pro Tyr Tyr Pro Arg Lys Ala Leu Leu Cys Tyr
            20                  25                  30

Gln Leu Thr Pro Gln Asn Gly Ser Thr Pro Thr Arg Gly His Leu Lys
        35                  40                  45

Asn Lys Lys Lys Asp His Ala Glu Ile Arg Phe Ile Asn Lys Ile Lys
    50                  55                  60
```

```
Ser Met Gly Leu Asp Glu Thr Gln Cys Tyr Gln Val Thr Cys Tyr Leu
65                  70                  75                  80

Thr Trp Ser Pro Cys Pro Ser Cys Ala Gly Glu Leu Val Asp Phe Ile
                85                  90                  95

Lys Ala His Arg His Leu Asn Leu Arg Ile Phe Ala Ser Arg Leu Tyr
            100                 105                 110

Tyr His Trp Arg Pro Asn Tyr Gln Glu Gly Leu Leu Leu Leu Cys Gly
        115                 120                 125

Ser Gln Val Pro Val Glu Val Met Gly Leu Pro Glu Phe Thr Asp Cys
    130                 135                 140

Trp Glu Asn Phe Val Asp His Lys Glu Pro Pro Ser Phe Asn Pro Ser
145                 150                 155                 160

Glu Lys Leu Glu Glu Leu Asp Lys Asn Ser Gln Ala Ile Lys Arg Arg
                165                 170                 175

Leu Glu Arg Ile Lys Ser Arg Ser Val Asp Val Leu Glu Asn Gly Leu
            180                 185                 190

Arg Ser Leu Gln Leu Gly Pro Val Thr Pro Ser Ser Ser Ile Arg Asn
        195                 200                 205

Ser Arg
    210
```

<210> 1010
<211> 428
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1010

```
Met Glu Lys Asp Ile Asn Leu Lys Ile Phe Lys Gly Asn Leu Ile Phe
1               5                  10                  15

Thr Lys Thr Ser Asp Lys Phe Thr Ile Met Lys Asp Ser Tyr Ile Val
            20                  25                  30

Val Ile Asp Gly Lys Ile Ala Ser Val Ser Ser Asn Leu Pro Asp Lys
        35                  40                  45

Tyr Lys Gly Asn Pro Ile Ile Asp Phe Arg Asn Asn Ile Ile Ile Pro
    50                  55                  60

Gly Met Asn Asp Leu His Ala His Ala Ser Gln Tyr Lys Asn Leu Gly
65                  70                  75                  80

Ile Gly Met Asp Lys Glu Leu Leu Pro Trp Leu Asn Asn Tyr Thr Phe
                85                  90                  95

Pro Glu Glu Ala Lys Phe Leu Asn Val Asp Tyr Ala Lys Lys Thr Tyr
            100                 105                 110

Gly Arg Leu Ile Lys Asp Leu Ile Lys Asn Gly Thr Thr Arg Val Ala
        115                 120                 125

Leu Phe Ala Thr Leu His Lys Asp Ser Thr Ile Glu Leu Phe Asn Met
    130                 135                 140

Leu Ile Lys Ser Gly Ile Gly Ala Tyr Val Gly Lys Val Asn Met Asp
145                 150                 155                 160

Tyr Asn Cys Pro Asp Tyr Leu Thr Glu Asn Tyr Ile Thr Ser Leu Asn
                165                 170                 175

Asp Thr Glu Glu Ile Ile Leu Lys Tyr Lys Asp Lys Ser Asn Ile Val
            180                 185                 190

Lys Pro Ile Ile Thr Pro Arg Phe Val Pro Ser Cys Ser Asn Glu Leu
        195                 200                 205
```

```
Met Asp Gly Leu Gly Lys Leu Ser Tyr Lys Tyr Arg Leu Pro Val Gln
    210                 215                 220

Ser His Leu Ser Glu Asn Leu Asp Glu Ile Ala Val Val Lys Ser Leu
225                 230                 235                 240

His Lys Lys Ser Asn Phe Tyr Gly Glu Val Tyr Asp Lys Phe Gly Leu
                245                 250                 255

Phe Gly Asn Thr Pro Thr Leu Met Ala His Cys Ile His Ser Ser Lys
            260                 265                 270

Glu Glu Ile Asn Leu Ile Lys Arg Asn Asn Val Thr Ile Val His Cys
        275                 280                 285

Pro Thr Ser Asn Phe Asn Leu Gly Ser Gly Met Met Pro Val Arg Lys
    290                 295                 300

Tyr Leu Asn Leu Gly Ile Asn Val Val Leu Gly Ser Asp Ile Ser Ala
305                 310                 315                 320

Gly His Thr Cys Ser Leu Phe Lys Val Ile Ala Tyr Ala Ile Gln Asn
                325                 330                 335

Ser Lys Ile Lys Trp Gln Glu Ser Gly Lys Lys Asp Met Phe Leu Ser
            340                 345                 350

Thr Ser Glu Ala Phe Tyr Met Ala Thr Lys Lys Gly Gly Ser Phe Phe
        355                 360                 365

Gly Lys Val Gly Ser Phe Glu Glu Gly Tyr Asp Phe Asp Ala Leu Val
    370                 375                 380

Ile Asn Asp Ser Asn Leu Tyr Pro Glu Asp Tyr Asp Leu Thr Glu Arg
385                 390                 395                 400
```

Leu Glu Arg Phe Ile Tyr Leu Gly Asp Asp Arg Asn Ile Met Lys Arg
405 410 415

Tyr Val Cys Gly Asn Glu Ile Phe Gly Pro Lys Phe
420 425

<210> 1011
<211> 412
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1011

Met Lys Ile Ile Asn Ala Arg Leu Arg Arg Gln Glu Ala Leu Phe Thr
1 5 10 15

Leu Asp Leu Gln Asp Gly Ile Ile His Arg Ile Thr Ala Gln Ala Ala
20 25 30

Met Gln Thr Ala Asp Ala Gly Ala Ile Asp Ala Gln Gly Arg Leu Ala
35 40 45

Ile Pro Pro Phe Val Glu Pro His Ile His Leu Asp Ala Thr Leu Thr
50 55 60

Ala Gly Glu Pro Glu Trp Asn Arg Ser Gly Thr Leu Phe Glu Gly Ile
65 70 75 80

Thr Arg Trp Ser Gln Arg Lys Ala Ser Ile Thr Pro Glu Asp Thr Arg
85 90 95

Gln Arg Ala Leu Lys Thr Ile Gly Met Leu Arg Asp Phe Gly Val Gln
100 105 110

His Val Arg Thr His Val Asp Val Thr Asp Pro Ser Leu Ala Ala Leu
115 120 125

```
Gln Ala Leu Leu Ala Val Lys Gln Glu Ala Ala Asp Leu Ile Asp Leu
    130                 135                 140

Gln Ile Val Ala Phe Pro Gln Glu Gly Ile Glu Ser Tyr Pro Asn Gly
145                 150                 155                 160

Arg Glu Leu Met Thr Arg Ala Ile Glu Met Gly Ala Asp Val Val Gly
                165                 170                 175

Gly Ile Pro His Tyr Glu Asn Thr Arg Asp Lys Gly Val Ser Ser Val
            180                 185                 190

Met Phe Leu Met Asp Leu Ala Gln Arg Tyr Gly Arg Leu Val Asp Val
        195                 200                 205

His Cys Asp Glu Ile Asp Asp Pro Gln Ser Arg Phe Leu Glu Val Leu
    210                 215                 220

Ala Glu Glu Ala Arg Val Arg Gly Met Gly Ala Gln Val Thr Ala Ser
225                 230                 235                 240

His Thr Cys Ala Met Gly Ser Tyr Asp Asn Ala Tyr Cys Ser Lys Leu
                245                 250                 255

Phe Arg Leu Leu Lys Ala Ser Gly Ile Asn Phe Ile Ser Cys Pro Thr
            260                 265                 270

Glu Ser Ile His Leu Gln Gly Arg Phe Asp Ser Trp Pro Lys Arg Arg
        275                 280                 285

Gly Val Thr Arg Val Ala Glu Leu Asp Arg Ala Gly Ile Asn Val Cys
    290                 295                 300

Phe Ala Gln Asp Ser Ile Gln Asp Pro Trp Tyr Pro Leu Gly Asn Gly
305                 310                 315                 320

Asn Ile Leu Arg Ile Leu Asp Ala Gly Leu His Ile Cys His Met Leu
                325                 330                 335
```

```
Gly Tyr Asp Asp Leu Gln Arg Cys Leu Asp Phe Val Thr Asp Asn Ser
            340                 345                 350


Ala Arg Ala Leu Cys Leu Gly Asp Asn Tyr Gly Leu Ala Glu Gly Arg
        355                 360                 365


Pro Ala Asn Leu Leu Ile Leu Asp Ala Glu Asn Asp Tyr Glu Ala Val
    370                 375                 380


Arg Arg Gln Ala Arg Val Leu Thr Ser Ile Arg His Gly Lys Val Ile
385                 390                 395                 400


Leu Gln Arg Glu Val Glu His Ile Arg Tyr Pro Ala
                405                 410


<210>  1012
<211>  812
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  1012

Met Gly Arg Lys Leu Asp Pro Thr Lys Glu Lys Arg Gly Pro Gly Arg
1               5                   10                  15


Lys Ala Arg Lys Gln Lys Gly Ala Glu Thr Glu Leu Val Arg Phe Leu
            20                  25                  30


Pro Ala Val Ser Asp Glu Asn Ser Lys Arg Leu Ser Ser Arg Ala Arg
        35                  40                  45


Lys Arg Ala Ala Lys Arg Arg Leu Gly Ser Val Glu Ala Pro Lys Thr
    50                  55                  60


Asn Lys Ser Pro Glu Ala Lys Pro Leu Pro Gly Lys Leu Pro Lys Gly
65                  70                  75                  80
```

```
Ile Ser Ala Gly Ala Val Gln Thr Ala Gly Lys Lys Gly Pro Gln Ser
                85                  90                  95

Leu Phe Asn Ala Pro Arg Gly Lys Lys Arg Pro Ala Pro Gly Ser Asp
            100                 105                 110

Glu Glu Glu Glu Glu Glu Asp Ser Glu Glu Asp Gly Met Val Asn His
        115                 120                 125

Gly Asp Leu Trp Gly Ser Glu Asp Asp Ala Asp Thr Val Asp Asp Tyr
    130                 135                 140

Gly Ala Asp Ser Asn Ser Glu Asp Glu Glu Glu Gly Glu Ala Leu Leu
145                 150                 155                 160

Pro Ile Glu Arg Ala Ala Arg Lys Gln Lys Ala Arg Glu Ala Ala Ala
                165                 170                 175

Gly Ile Gln Trp Ser Glu Glu Glu Thr Glu Asp Glu Glu Glu Glu Lys
            180                 185                 190

Glu Val Thr Pro Glu Ser Gly Pro Pro Lys Val Glu Glu Ala Asp Gly
        195                 200                 205

Gly Leu Gln Ile Asn Val Asp Glu Glu Pro Phe Val Leu Pro Pro Ala
    210                 215                 220

Gly Glu Met Glu Gln Asp Ala Gln Ala Pro Asp Leu Gln Arg Val His
225                 230                 235                 240

Lys Arg Ile Gln Asp Ile Val Gly Ile Leu Arg Asp Phe Gly Ala Gln
                245                 250                 255

Arg Glu Glu Gly Arg Ser Arg Ser Glu Tyr Leu Asn Arg Leu Lys Lys
            260                 265                 270
```

```
Asp Leu Ala Ile Tyr Tyr Ser Tyr Gly Asp Phe Leu Leu Gly Lys Leu
        275                 280                 285

Met Asp Leu Phe Pro Leu Ser Glu Leu Val Glu Phe Leu Glu Ala Asn
    290                 295                 300

Glu Val Pro Arg Pro Val Thr Leu Arg Thr Asn Thr Leu Lys Thr Arg
305                 310                 315                 320

Arg Arg Asp Leu Ala Gln Ala Leu Ile Asn Arg Gly Val Asn Leu Asp
                325                 330                 335

Pro Leu Gly Lys Trp Ser Lys Thr Gly Leu Val Val Tyr Asp Ser Ser
            340                 345                 350

Val Pro Ile Gly Ala Thr Pro Glu Tyr Leu Ala Gly His Tyr Met Leu
        355                 360                 365

Gln Gly Ala Ser Ser Met Leu Pro Val Met Ala Leu Ala Pro Gln Glu
    370                 375                 380

His Glu Arg Ile Leu Asp Met Cys Cys Ala Pro Gly Gly Lys Thr Ser
385                 390                 395                 400

Tyr Met Ala Gln Leu Met Lys Asn Thr Gly Val Ile Leu Ala Asn Asp
                405                 410                 415

Ala Asn Ala Glu Arg Leu Lys Ser Val Val Gly Asn Leu His Arg Leu
            420                 425                 430

Gly Val Thr Asn Thr Ile Ile Ser His Tyr Asp Gly Arg Gln Phe Pro
        435                 440                 445

Lys Val Val Gly Gly Phe Asp Arg Val Leu Leu Asp Ala Pro Cys Ser
    450                 455                 460

Gly Thr Gly Val Ile Ser Lys Asp Pro Ala Val Lys Thr Asn Lys Asp
465                 470                 475                 480
```

```
Glu Lys Asp Ile Leu Arg Cys Ala His Leu Gln Lys Glu Leu Leu Leu
                485                 490                 495

Ser Ala Ile Asp Ser Val Asn Ala Thr Ser Lys Thr Gly Gly Tyr Leu
            500                 505                 510

Val Tyr Cys Thr Cys Ser Ile Thr Val Glu Glu Asn Glu Trp Val Val
        515                 520                 525

Asp Tyr Ala Leu Lys Lys Arg Asn Val Arg Leu Val Pro Thr Gly Leu
    530                 535                 540

Asp Phe Gly Gln Glu Gly Phe Thr Arg Phe Arg Glu Arg Arg Phe His
545                 550                 555                 560

Pro Ser Leu Arg Ser Thr Arg Arg Phe Tyr Pro His Thr His Asn Met
                565                 570                 575

Asp Gly Phe Phe Ile Ala Lys Phe Lys Lys Phe Ser Asn Ser Ile Pro
            580                 585                 590

Gln Ser Gln Thr Gly Asn Ser Glu Thr Ala Thr Pro Thr Asn Val Asp
        595                 600                 605

Leu Pro Gln Val Ile Pro Lys Ser Glu Asn Ser Ser Gln Pro Ala Lys
    610                 615                 620

Lys Ala Lys Gly Ala Ala Lys Thr Lys Gln Gln Leu Gln Lys Gln Gln
625                 630                 635                 640

His Pro Lys Lys Ala Ser Phe Gln Lys Leu Asn Gly Ile Ser Lys Gly
                645                 650                 655

Ala Asp Ser Glu Leu Ser Thr Val Pro Ser Val Thr Lys Thr Gln Ala
            660                 665                 670
```

```
Ser Ser Ser Phe Gln Asp Ser Ser Gln Pro Ala Gly Lys Ala Glu Gly
        675                 680                 685

Ile Arg Glu Pro Lys Val Thr Gly Lys Leu Lys Gln Arg Ser Pro Lys
    690                 695                 700

Leu Gln Ser Ser Lys Lys Val Ala Phe Leu Arg Gln Asn Ala Pro Pro
705                 710                 715                 720

Lys Gly Thr Asp Thr Gln Thr Pro Ala Val Leu Ser Pro Ser Lys Thr
                725                 730                 735

Gln Ala Thr Leu Lys Pro Lys Asp His His Gln Pro Leu Gly Arg Ala
            740                 745                 750

Lys Gly Val Glu Lys Gln Gln Leu Pro Glu Gln Pro Phe Glu Lys Ala
        755                 760                 765

Ala Phe Gln Lys Gln Asn Asp Thr Pro Lys Gly Pro Gln Pro Pro Thr
    770                 775                 780

Val Ser Pro Ile Arg Ser Ser Arg Pro Pro Pro Ala Lys Arg Lys Lys
785                 790                 795                 800

Ser Gln Ser Arg Gly Asn Ser Gln Leu Leu Leu Ser
                805                 810
```

<210> 1013
<211> 767
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1013

```
Met Gly Arg Arg Ser Arg Gly Arg Arg Leu Gln Gln Gln Gln Arg Pro
1               5                   10                  15
```

```
Glu Asp Ala Glu Asp Gly Ala Glu Gly Gly Gly Lys Arg Gly Glu Ala
            20                  25                  30

Gly Trp Glu Gly Gly Tyr Pro Glu Ile Val Lys Glu Asn Lys Leu Phe
        35                  40                  45

Glu His Tyr Tyr Gln Glu Leu Lys Ile Val Pro Glu Gly Glu Trp Gly
    50                  55                  60

Gln Phe Met Asp Ala Leu Arg Glu Pro Leu Pro Ala Thr Leu Arg Ile
65                  70                  75                  80

Thr Gly Tyr Lys Ser His Ala Lys Glu Ile Leu His Cys Leu Lys Asn
                85                  90                  95

Lys Tyr Phe Lys Glu Leu Glu Asp Leu Glu Val Asp Gly Gln Lys Val
            100                 105                 110

Glu Val Pro Gln Pro Leu Ser Trp Tyr Pro Glu Glu Leu Ala Trp His
        115                 120                 125

Thr Asn Leu Ser Arg Lys Ile Leu Arg Lys Ser Pro His Leu Glu Lys
    130                 135                 140

Phe His Gln Phe Leu Val Ser Glu Thr Glu Ser Gly Asn Ile Ser Arg
145                 150                 155                 160

Gln Glu Ala Val Ser Met Ile Pro Pro Leu Leu Leu Asn Val Arg Pro
                165                 170                 175

His His Lys Ile Leu Asp Met Cys Ala Ala Pro Gly Ser Lys Thr Thr
            180                 185                 190

Gln Leu Ile Glu Met Leu His Ala Asp Met Asn Val Pro Phe Pro Glu
        195                 200                 205

Gly Phe Val Ile Ala Asn Asp Val Asp Asn Lys Arg Cys Tyr Leu Leu
    210                 215                 220
```

```
Val His Gln Ala Lys Arg Leu Ser Ser Pro Cys Ile Met Val Val Asn
225                 230                 235                 240

His Asp Ala Ser Ser Ile Pro Arg Leu Gln Ile Asp Val Asp Gly Arg
                245                 250                 255

Lys Glu Ile Leu Phe Tyr Asp Arg Ile Leu Cys Asp Val Pro Cys Ser
            260                 265                 270

Gly Asp Gly Thr Met Arg Lys Asn Ile Asp Val Trp Lys Lys Trp Thr
        275                 280                 285

Thr Leu Asn Ser Leu Gln Leu His Gly Leu Gln Leu Arg Ile Ala Thr
    290                 295                 300

Arg Gly Ala Glu Gln Leu Ala Glu Gly Gly Arg Met Val Tyr Ser Thr
305                 310                 315                 320

Cys Ser Leu Asn Pro Ile Glu Asp Glu Ala Val Ile Ala Ser Leu Leu
                325                 330                 335

Glu Lys Ser Glu Gly Ala Leu Glu Leu Ala Asp Val Ser Asn Glu Leu
            340                 345                 350

Pro Gly Leu Lys Trp Met Pro Gly Ile Thr Gln Trp Lys Val Met Thr
        355                 360                 365

Lys Asp Gly Gln Trp Phe Thr Asp Trp Asp Ala Val Pro His Ser Arg
    370                 375                 380

His Thr Gln Ile Arg Pro Thr Met Phe Pro Pro Lys Asp Pro Glu Lys
385                 390                 395                 400

Leu Gln Ala Met His Leu Glu Arg Cys Leu Arg Ile Leu Pro His His
                405                 410                 415
```

```
Gln Asn Thr Gly Gly Phe Phe Val Ala Val Leu Val Lys Lys Ser Ser
            420                 425                 430

Met Pro Trp Asn Lys Arg Gln Pro Lys Leu Gln Gly Lys Ser Ala Glu
        435                 440                 445

Thr Arg Glu Ser Thr Gln Leu Ser Pro Ala Asp Leu Thr Glu Gly Lys
    450                 455                 460

Pro Thr Asp Pro Ser Lys Leu Glu Ser Pro Ser Phe Thr Gly Thr Gly
465                 470                 475                 480

Asp Thr Glu Ile Ala His Ala Thr Glu Asp Leu Glu Asn Asn Gly Ser
                485                 490                 495

Lys Lys Asp Gly Val Cys Gly Pro Pro Pro Ser Lys Lys Met Lys Leu
            500                 505                 510

Phe Gly Phe Lys Glu Asp Pro Phe Val Phe Ile Pro Glu Asp Asp Pro
        515                 520                 525

Leu Phe Pro Pro Ile Glu Lys Phe Tyr Ala Leu Asp Pro Ser Phe Pro
    530                 535                 540

Arg Met Asn Leu Leu Thr Arg Thr Thr Glu Gly Lys Lys Arg Gln Leu
545                 550                 555                 560

Tyr Met Val Ser Lys Glu Leu Arg Asn Val Leu Leu Asn Asn Ser Glu
                565                 570                 575

Lys Met Lys Val Ile Asn Thr Gly Ile Lys Val Trp Cys Arg Asn Asn
            580                 585                 590

Ser Gly Glu Glu Phe Asp Cys Ala Phe Arg Leu Ala Gln Glu Gly Ile
        595                 600                 605

Tyr Thr Leu Tyr Pro Phe Ile Asn Ser Arg Ile Ile Thr Val Ser Met
    610                 615                 620
```

```
Glu Asp Val Lys Ile Leu Leu Thr Gln Glu Asn Pro Phe Phe Arg Lys
625                 630                 635                 640


Leu Ser Ser Glu Thr Tyr Ser Gln Ala Lys Asp Leu Ala Lys Gly Ser
                645                 650                 655


Ile Val Leu Lys Tyr Glu Pro Asp Ser Ala Asn Pro Asp Ala Leu Gln
            660                 665                 670


Cys Pro Ile Val Leu Cys Gly Trp Arg Gly Lys Ala Ser Ile Arg Thr
        675                 680                 685


Phe Val Pro Lys Asn Glu Arg Leu His Tyr Leu Arg Met Met Gly Leu
    690                 695                 700


Glu Val Leu Gly Glu Lys Lys Lys Glu Gly Val Ile Leu Thr Asn Glu
705                 710                 715                 720


Ser Ala Ala Ser Thr Gly Gln Pro Asp Asn Asp Val Thr Glu Gly Gln
                725                 730                 735


Arg Ala Gly Glu Pro Asn Ser Pro Asp Ala Glu Glu Ala Asn Ser Pro
            740                 745                 750


Asp Val Thr Ala Gly Cys Asp Pro Ala Gly Val His Pro Pro Arg
        755                 760                 765


<210>  1014
<211>  429
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  1014

Met Gly Pro Phe Cys Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro
1               5                   10                  15
```

```
Ile Arg Asn Leu Ile Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn
            20                  25                  30

Leu Gly Tyr Ala Lys Gly Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val
        35                  40                  45

Thr Arg Lys Asp Cys Asp Ser Pro Val Ser Leu His His Gly Val Phe
    50                  55                  60

Lys Asn Lys Asp Asn Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe
65                  70                  75                  80

His Asp Lys Val Leu Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile
                85                  90                  95

Thr Trp Tyr Met Ser Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Ile
            100                 105                 110

Val Arg Phe Leu Ala Thr His His Asn Leu Ser Leu Asp Ile Phe Ser
        115                 120                 125

Ser Arg Leu Tyr Asn Val Gln Asp Pro Glu Thr Gln Gln Asn Leu Cys
    130                 135                 140

Arg Leu Val Gln Glu Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu
145                 150                 155                 160

Phe Lys Lys Cys Trp Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe
                165                 170                 175

Arg Pro Trp Lys Arg Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys
            180                 185                 190

Leu Gln Glu Ile Leu Arg Pro Cys Tyr Ile Pro Val Pro Ser Ser Ser
        195                 200                 205
```

```
Ser Ser Thr Leu Ser Asn Ile Cys Leu Thr Lys Gly Leu Pro Glu Thr
    210                 215                 220

Arg Phe Cys Val Glu Gly Arg Arg Met Asp Pro Leu Ser Glu Glu Glu
225                 230                 235                 240

Phe Tyr Ser Gln Phe Tyr Asn Gln Arg Val Lys His Leu Cys Tyr Tyr
                245                 250                 255

His Arg Met Lys Pro Tyr Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly
            260                 265                 270

Gln Ala Pro Leu Lys Gly Cys Leu Leu Ser Glu Lys Gly Lys Gln His
        275                 280                 285

Ala Glu Ile Leu Phe Leu Asp Lys Ile Arg Ser Met Glu Leu Ser Gln
    290                 295                 300

Val Thr Ile Thr Cys Tyr Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala
305                 310                 315                 320

Trp Gln Leu Ala Ala Phe Lys Arg Asp Arg Pro Asp Leu Ile Leu His
                325                 330                 335

Ile Tyr Thr Ser Arg Leu Tyr Phe His Trp Lys Arg Pro Phe Gln Lys
            340                 345                 350

Gly Leu Cys Ser Leu Trp Gln Ser Gly Ile Leu Val Asp Val Met Asp
        355                 360                 365

Leu Pro Gln Phe Thr Asp Cys Trp Thr Asn Phe Val Asn Pro Lys Arg
    370                 375                 380

Pro Phe Trp Pro Trp Lys Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln
385                 390                 395                 400

Arg Arg Leu Arg Arg Ile Lys Glu Ser Trp Gly Leu Gln Asp Leu Val
                405                 410                 415
```

Asn Asp Phe Gly Asn Leu Gln Leu Gly Pro Pro Met Ser
            420   425

<210> 1015
<211> 429
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1015

Met Gly Pro Phe Cys Leu Gly Cys Ser His Arg Lys Cys Tyr Ser Pro
1   5   10   15

Ile Arg Asn Leu Ile Ser Gln Glu Thr Phe Lys Phe His Phe Lys Asn
            20   25   30

Leu Arg Tyr Ala Ile Asp Arg Lys Asp Thr Phe Leu Cys Tyr Glu Val
        35   40   45

Thr Arg Lys Asp Cys Asp Ser Pro Val Ser Leu His His Gly Val Phe
    50   55   60

Lys Asn Lys Asp Asn Ile His Ala Glu Ile Cys Phe Leu Tyr Trp Phe
65   70   75   80

His Asp Lys Val Leu Lys Val Leu Ser Pro Arg Glu Glu Phe Lys Ile
                85   90   95

Thr Trp Tyr Met Ser Trp Ser Pro Cys Phe Glu Cys Ala Glu Gln Val
            100   105   110

Leu Arg Phe Leu Ala Thr His His Asn Leu Ser Leu Asp Ile Phe Ser
        115   120   125

Ser Arg Leu Tyr Asn Ile Arg Asp Pro Glu Asn Gln Gln Asn Leu Cys
    130   135   140

```
Arg Leu Val Gln Glu Gly Ala Gln Val Ala Ala Met Asp Leu Tyr Glu
145                 150                 155                 160

Phe Lys Lys Cys Trp Lys Lys Phe Val Asp Asn Gly Gly Arg Arg Phe
                165                 170                 175

Arg Pro Trp Lys Lys Leu Leu Thr Asn Phe Arg Tyr Gln Asp Ser Lys
            180                 185                 190

Leu Gln Glu Ile Leu Arg Pro Cys Tyr Ile Pro Val Pro Ser Ser Ser
        195                 200                 205

Ser Ser Thr Leu Ser Asn Ile Cys Leu Thr Lys Gly Leu Pro Glu Thr
    210                 215                 220

Arg Phe Cys Val Glu Arg Arg Arg Val His Leu Leu Ser Glu Glu Glu
225                 230                 235                 240

Phe Tyr Ser Gln Phe Tyr Asn Gln Arg Val Lys His Leu Cys Tyr Tyr
                245                 250                 255

His Gly Val Lys Pro Tyr Leu Cys Tyr Gln Leu Glu Gln Phe Asn Gly
            260                 265                 270

Gln Ala Pro Leu Lys Gly Cys Leu Leu Ser Glu Lys Gly Lys Gln His
        275                 280                 285

Ala Glu Ile Leu Phe Leu Asp Lys Ile Arg Ser Met Glu Leu Ser Gln
    290                 295                 300

Val Ile Ile Thr Cys Tyr Leu Thr Trp Ser Pro Cys Pro Asn Cys Ala
305                 310                 315                 320

Trp Gln Leu Ala Ala Phe Lys Arg Asp Arg Pro Asp Leu Ile Leu His
                325                 330                 335
```

```
Ile Tyr Thr Ser Arg Leu Tyr Phe His Trp Lys Arg Pro Phe Gln Lys
            340                 345                 350

Gly Leu Cys Ser Leu Trp Gln Ser Gly Ile Leu Val Asp Val Met Asp
        355                 360                 365

Leu Pro Gln Phe Thr Asp Cys Trp Thr Asn Phe Val Asn Pro Lys Arg
    370                 375                 380

Pro Phe Trp Pro Trp Lys Gly Leu Glu Ile Ile Ser Arg Arg Thr Gln
385                 390                 395                 400

Arg Arg Leu His Arg Ile Lys Glu Ser Trp Gly Leu Gln Asp Leu Val
                405                 410                 415

Asn Asp Phe Gly Asn Leu Gln Leu Gly Pro Pro Met Ser
            420                 425
```

<210> 1016
<211> 370
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1016

```
Met Val Glu Pro Met Asp Pro Arg Thr Phe Val Ser Asn Phe Asn Asn
1               5                   10                  15

Arg Pro Ile Leu Ser Gly Leu Asn Thr Val Trp Leu Cys Cys Glu Val
            20                  25                  30

Lys Thr Lys Asp Pro Ser Gly Pro Pro Leu Asp Ala Lys Ile Phe Gln
        35                  40                  45

Gly Lys Val Tyr Ser Lys Ala Lys Tyr His Pro Glu Met Arg Phe Leu
    50                  55                  60
```

```
Arg Trp Phe His Lys Trp Arg Gln Leu His His Asp Gln Glu Tyr Lys
65                  70                  75                  80

Val Thr Trp Tyr Val Ser Trp Ser Pro Cys Thr Arg Cys Ala Asn Ser
                85                  90                  95

Val Ala Thr Phe Leu Ala Lys Asp Pro Lys Val Thr Leu Thr Ile Phe
            100                 105                 110

Val Ala Arg Leu Tyr Tyr Phe Trp Lys Pro Asp Tyr Gln Gln Ala Leu
        115                 120                 125

Arg Ile Leu Cys Gln Lys Arg Gly Gly Pro His Ala Thr Met Lys Ile
    130                 135                 140

Met Asn Tyr Asn Glu Phe Gln Asp Cys Trp Asn Lys Phe Val Asp Gly
145                 150                 155                 160

Arg Gly Lys Pro Phe Lys Pro Arg Asn Asn Leu Pro Lys His Tyr Thr
                165                 170                 175

Leu Leu Gln Ala Thr Leu Gly Glu Leu Leu Arg His Leu Met Asp Pro
            180                 185                 190

Gly Thr Phe Thr Ser Asn Phe Asn Asn Lys Pro Trp Val Ser Gly Gln
        195                 200                 205

His Glu Thr Tyr Leu Cys Tyr Lys Val Glu Arg Leu His Asn Asp Thr
    210                 215                 220

Trp Val Pro Leu Asn Gln His Arg Gly Phe Leu Arg Asn Gln Ala Pro
225                 230                 235                 240

Asn Ile His Gly Phe Pro Lys Gly Arg His Ala Glu Leu Cys Phe Leu
                245                 250                 255

Asp Leu Ile Pro Phe Trp Lys Leu Asp Gly Gln Gln Tyr Arg Val Thr
            260                 265                 270
```

```
Cys Phe Thr Ser Trp Ser Pro Cys Phe Ser Cys Ala Gln Glu Met Ala
        275                 280                 285

Lys Phe Ile Ser Asn Asn Glu His Val Ser Leu Cys Ile Phe Ala Ala
    290                 295                 300

Arg Ile Tyr Asp Asp Gln Gly Arg Tyr Gln Glu Gly Leu Arg Ala Leu
305                 310                 315                 320

His Arg Asp Gly Ala Lys Ile Ala Met Met Asn Tyr Ser Glu Phe Glu
                325                 330                 335

Tyr Cys Trp Asp Thr Phe Val Asp Arg Gln Gly Arg Pro Phe Gln Pro
            340                 345                 350

Trp Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg
        355                 360                 365

Ala Ile
    370
```

<210> 1017
<211> 384
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1017

```
Met Lys Pro His Phe Arg Asn Pro Val Glu Arg Met Tyr Gln Asp Thr
1               5                   10                  15

Phe Ser Asp Asn Phe Tyr Asn Arg Pro Ile Leu Ser His Arg Asn Thr
            20                  25                  30

Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Gly Pro Ser Arg Pro Pro
        35                  40                  45
```

```
Leu Asp Ala Lys Ile Phe Arg Gly Gln Val Tyr Ser Lys Leu Lys Tyr
    50                  55                  60

His Pro Glu Met Arg Phe Phe His Trp Phe Ser Lys Trp Arg Lys Leu
65                  70                  75                  80

His Arg Asp Gln Glu Tyr Glu Val Thr Trp Tyr Ile Ser Trp Ser Pro
                85                  90                  95

Cys Thr Lys Cys Thr Arg Asp Val Ala Thr Phe Leu Ala Glu Asp Pro
            100                 105                 110

Lys Val Thr Leu Thr Ile Phe Val Ala Arg Leu Tyr Tyr Phe Trp Asp
        115                 120                 125

Pro Asp Tyr Gln Glu Ala Leu Arg Ser Leu Cys Gln Lys Arg Asp Gly
    130                 135                 140

Pro Arg Ala Thr Met Lys Ile Met Asn Tyr Asp Glu Phe Gln His Cys
145                 150                 155                 160

Trp Ser Lys Phe Val Tyr Ser Gln Arg Glu Leu Phe Glu Pro Trp Asn
                165                 170                 175

Asn Leu Pro Lys Tyr Tyr Ile Leu Leu His Ile Met Leu Gly Glu Ile
            180                 185                 190

Leu Arg His Ser Met Asp Pro Pro Thr Phe Thr Ser Asn Phe Asn Asn
        195                 200                 205

Glu Leu Trp Val Arg Gly Arg His Glu Thr Tyr Leu Cys Tyr Glu Val
    210                 215                 220

Glu Arg Leu His Asn Asp Thr Trp Val Leu Leu Asn Gln Arg Arg Gly
225                 230                 235                 240
```

```
Phe Leu Cys Asn Gln Ala Pro His Lys His Gly Phe Leu Glu Gly Arg
                245                 250                 255

His Ala Glu Leu Cys Phe Leu Asp Val Ile Pro Phe Trp Lys Leu Asp
            260                 265                 270

Leu His Gln Asp Tyr Arg Val Thr Cys Phe Thr Ser Trp Ser Pro Cys
        275                 280                 285

Phe Ser Cys Ala Gln Glu Met Ala Lys Phe Ile Ser Asn Asn Lys His
    290                 295                 300

Val Ser Leu Cys Ile Phe Ala Ala Arg Ile Tyr Asp Asp Gln Gly Arg
305                 310                 315                 320

Cys Gln Glu Gly Leu Arg Thr Leu Ala Lys Ala Gly Ala Lys Ile Ser
                325                 330                 335

Ile Met Thr Tyr Ser Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp
            340                 345                 350

His Gln Gly Cys Pro Phe Gln Pro Trp Asp Gly Leu Glu Glu His Ser
        355                 360                 365

Gln Ala Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn
    370                 375                 380
```

<210> 1018
<211> 377
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1018

```
Met Asn Pro Gln Ile Arg Asn Met Val Glu Gln Met Glu Pro Asp Ile
1               5                   10                  15
```

```
Phe Val Tyr Tyr Phe Asn Asn Arg Pro Ile Leu Ser Gly Arg Asn Thr
            20                  25                  30

Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Asp Pro Ser Gly Pro Pro
        35                  40                  45

Leu Asp Ala Asn Ile Phe Gln Gly Lys Leu Tyr Pro Glu Ala Lys Asp
    50                  55                  60

His Pro Glu Met Lys Phe Leu His Trp Phe Arg Lys Trp Arg Gln Leu
65                  70                  75                  80

His Arg Asp Gln Glu Tyr Glu Val Thr Trp Tyr Val Ser Trp Ser Pro
                85                  90                  95

Cys Thr Arg Cys Ala Asn Ser Val Ala Thr Phe Leu Ala Glu Asp Pro
            100                 105                 110

Lys Val Thr Leu Thr Ile Phe Val Ala Arg Leu Tyr Tyr Phe Trp Lys
        115                 120                 125

Pro Asp Tyr Gln Gln Ala Leu Arg Ile Leu Cys Gln Glu Arg Gly Gly
    130                 135                 140

Pro His Ala Thr Met Lys Ile Met Asn Tyr Asn Glu Phe Gln His Cys
145                 150                 155                 160

Trp Asn Glu Phe Val Asp Gly Gln Gly Lys Pro Phe Lys Pro Arg Lys
                165                 170                 175

Asn Leu Pro Lys His Tyr Thr Leu Leu His Ala Thr Leu Gly Glu Leu
            180                 185                 190

Leu Arg His Val Met Asp Pro Gly Thr Phe Thr Ser Asn Phe Asn Asn
        195                 200                 205

Lys Pro Trp Val Ser Gly Gln Arg Glu Thr Tyr Leu Cys Tyr Lys Val
    210                 215                 220
```

```
Glu Arg Ser His Asn Asp Thr Trp Val Leu Leu Asn Gln His Arg Gly
225                 230                 235                 240

Phe Leu Arg Asn Gln Ala Pro Asp Arg His Gly Phe Pro Lys Gly Arg
                245                 250                 255

His Ala Glu Leu Cys Phe Leu Asp Leu Ile Pro Phe Trp Lys Leu Asp
            260                 265                 270

Asp Gln Gln Tyr Arg Val Thr Cys Phe Thr Ser Trp Ser Pro Cys Phe
        275                 280                 285

Ser Cys Ala Gln Lys Met Ala Lys Phe Ile Ser Asn Asn Lys His Val
    290                 295                 300

Ser Leu Cys Ile Phe Ala Ala Arg Ile Tyr Asp Asp Gln Gly Arg Cys
305                 310                 315                 320

Gln Glu Gly Leu Arg Thr Leu His Arg Asp Gly Ala Lys Ile Ala Val
                325                 330                 335

Met Asn Tyr Ser Glu Phe Glu Tyr Cys Trp Asp Thr Phe Val Asp Arg
            340                 345                 350

Gln Gly Arg Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser Gln
        355                 360                 365

Ala Leu Ser Gly Arg Leu Arg Ala Ile
    370                 375
```

<210> 1019
<211> 384
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1019

```
Met Lys Pro His Phe Arg Asn Thr Val Glu Arg Met Tyr Arg Asp Thr
1               5                   10                  15


Phe Ser Tyr Asn Phe Tyr Asn Arg Pro Ile Leu Ser Arg Arg Asn Thr
            20                  25                  30


Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Gly Pro Ser Arg Pro Pro
        35                  40                  45


Leu Asp Ala Lys Ile Phe Arg Gly Gln Val Tyr Ser Glu Leu Lys Tyr
    50                  55                  60


His Pro Glu Met Arg Phe Phe His Trp Phe Ser Lys Trp Arg Lys Leu
65                  70                  75                  80


His Arg Asp Gln Glu Tyr Glu Val Thr Trp Tyr Ile Ser Trp Ser Pro
                85                  90                  95


Cys Thr Lys Cys Thr Arg Asp Met Ala Thr Phe Leu Ala Glu Asp Pro
            100                 105                 110


Lys Val Thr Leu Thr Ile Phe Val Ala Arg Leu Tyr Tyr Phe Trp Asp
        115                 120                 125


Pro Asp Tyr Gln Glu Ala Leu Arg Ser Leu Cys Gln Lys Arg Asp Gly
    130                 135                 140


Pro Arg Ala Thr Met Lys Ile Met Asn Tyr Asp Glu Phe Gln His Cys
145                 150                 155                 160


Trp Ser Lys Phe Val Tyr Ser Gln Arg Glu Leu Phe Glu Pro Trp Asn
                165                 170                 175


Asn Leu Pro Lys Tyr Tyr Ile Leu Leu His Ile Met Leu Gly Glu Ile
            180                 185                 190
```

```
Leu Arg His Ser Met Asp Pro Pro Thr Phe Thr Phe Asn Phe Asn Asn
        195                 200                 205

Glu Pro Trp Val Arg Gly Arg His Glu Thr Tyr Leu Cys Tyr Glu Val
    210                 215                 220

Glu Arg Met His Asn Asp Thr Trp Val Leu Leu Asn Gln Arg Arg Gly
225                 230                 235                 240

Phe Leu Cys Asn Gln Ala Pro His Lys His Gly Phe Leu Glu Gly Arg
                245                 250                 255

His Ala Glu Leu Cys Phe Leu Asp Val Ile Pro Phe Trp Lys Leu Asp
            260                 265                 270

Leu Asp Gln Asp Tyr Arg Val Thr Cys Phe Thr Ser Trp Ser Pro Cys
        275                 280                 285

Phe Ser Cys Ala Gln Glu Met Ala Lys Phe Ile Ser Lys Asn Lys His
    290                 295                 300

Val Ser Leu Cys Ile Phe Thr Ala Arg Ile Tyr Asp Asp Gln Gly Arg
305                 310                 315                 320

Cys Gln Glu Gly Leu Arg Thr Leu Ala Glu Ala Gly Ala Lys Ile Ser
                325                 330                 335

Ile Met Thr Tyr Ser Glu Phe Lys His Cys Trp Asp Thr Phe Val Asp
            340                 345                 350

His Gln Gly Cys Pro Phe Gln Pro Trp Asp Gly Leu Asp Glu His Ser
        355                 360                 365

Gln Asp Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Glu Asn
    370                 375                 380
```

<210> 1020
<211> 373

<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1020

```
Met Lys Pro His Phe Arg Asn Thr Val Glu Arg Met Tyr Arg Asp Thr
1               5                   10                  15


Phe Ser Tyr Asn Phe Tyr Asn Arg Pro Ile Leu Ser Arg Arg Asn Thr
            20                  25                  30


Val Trp Leu Cys Tyr Glu Val Lys Thr Lys Gly Pro Ser Arg Pro Arg
        35                  40                  45


Leu Asp Ala Lys Ile Phe Arg Gly Gln Val Tyr Ser Gln Pro Glu His
    50                  55                  60


His Ala Glu Met Cys Phe Leu Ser Trp Phe Cys Gly Asn Gln Leu Pro
65                  70                  75                  80


Ala Tyr Lys Cys Phe Gln Ile Thr Trp Phe Val Ser Trp Thr Pro Cys
                85                  90                  95


Pro Asp Cys Val Ala Lys Leu Ala Glu Phe Leu Ala Glu His Pro Asn
            100                 105                 110


Val Thr Leu Thr Ile Ser Ala Ala Arg Leu Tyr Tyr Tyr Trp Glu Arg
        115                 120                 125


Asp Tyr Arg Arg Ala Leu Cys Arg Leu Ser Gln Ala Gly Ala Arg Val
    130                 135                 140


Lys Ile Met Asp Asp Glu Glu Phe Ala Tyr Cys Trp Glu Asn Phe Val
145                 150                 155                 160


Tyr Ser Glu Gly Gln Pro Phe Met Pro Trp Tyr Lys Phe Asp Asp Asn
                165                 170                 175
```

```
Tyr Ala Phe Leu His Arg Thr Leu Lys Glu Ile Leu Arg Asn Pro Met
            180                 185                 190

Glu Ala Met Tyr Pro His Ile Phe Tyr Phe His Phe Lys Asn Leu Arg
        195                 200                 205

Lys Ala Tyr Gly Arg Asn Glu Ser Trp Leu Cys Phe Thr Met Glu Val
    210                 215                 220

Val Lys His His Ser Pro Val Ser Trp Lys Arg Gly Val Phe Arg Asn
225                 230                 235                 240

Gln Val Asp Pro Glu Thr His Cys His Ala Glu Arg Cys Phe Leu Ser
                245                 250                 255

Trp Phe Cys Asp Asp Ile Leu Ser Pro Asn Thr Asn Tyr Glu Val Thr
            260                 265                 270

Trp Tyr Thr Ser Trp Ser Pro Cys Pro Glu Cys Ala Gly Glu Val Ala
        275                 280                 285

Glu Phe Leu Ala Arg His Ser Asn Val Asn Leu Thr Ile Phe Thr Ala
    290                 295                 300

Arg Leu Tyr Tyr Phe Trp Asp Thr Asp Tyr Gln Glu Gly Leu Arg Ser
305                 310                 315                 320

Leu Ser Gln Glu Gly Ala Ser Val Glu Ile Met Gly Tyr Lys Asp Phe
                325                 330                 335

Lys Tyr Cys Trp Glu Asn Phe Val Tyr Asn Asp Asp Glu Pro Phe Lys
            340                 345                 350

Pro Trp Lys Gly Leu Lys Tyr Asn Phe Leu Phe Leu Asp Ser Lys Leu
        355                 360                 365
```

Gln Glu Ile Leu Glu
370

<210> 1021
<211> 382
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1021

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Arg Met Tyr Arg Asp Thr
1 5 10 15

Phe Tyr Asp Asn Phe Glu Asn Glu Pro Ile Leu Tyr Gly Arg Ser Tyr
20 25 30

Thr Trp Leu Cys Tyr Glu Val Lys Ile Lys Arg Gly Arg Ser Asn Leu
35 40 45

Leu Trp Asp Thr Gly Val Phe Arg Gly Gln Val Tyr Phe Lys Pro Gln
50 55 60

Tyr His Ala Glu Met Cys Phe Leu Ser Trp Phe Cys Gly Asn Gln Leu
65 70 75 80

Pro Ala Tyr Lys Cys Phe Gln Ile Thr Trp Phe Val Ser Trp Thr Pro
85 90 95

Cys Pro Asp Cys Val Ala Lys Leu Ala Glu Phe Leu Ser Glu His Pro
100 105 110

Asn Val Thr Leu Thr Ile Ser Ala Ala Arg Leu Tyr Tyr Tyr Trp Glu
115 120 125

Arg Asp Tyr Arg Arg Ala Leu Cys Arg Leu Ser Gln Ala Gly Ala Arg
130 135 140

```
Val Thr Ile Met Asp Tyr Glu Glu Phe Ala Tyr Cys Trp Glu Asn Phe
145                 150                 155                 160

Val Tyr Asn Glu Gly Gln Gln Phe Met Pro Trp Tyr Lys Phe Asp Glu
                165                 170                 175

Asn Tyr Ala Phe Leu His Arg Thr Leu Lys Glu Ile Leu Arg Tyr Leu
            180                 185                 190

Met Asp Pro Asp Thr Phe Thr Phe Asn Phe Asn Asn Asp Pro Leu Val
        195                 200                 205

Leu Arg Arg Arg Gln Thr Tyr Leu Cys Tyr Glu Val Glu Arg Leu Asp
    210                 215                 220

Asn Gly Thr Trp Val Leu Met Asp Gln His Met Gly Phe Leu Cys Asn
225                 230                 235                 240

Glu Ala Lys Asn Leu Leu Cys Gly Phe Tyr Gly Arg His Ala Glu Leu
                245                 250                 255

Arg Phe Leu Asp Leu Val Pro Ser Leu Gln Leu Asp Pro Ala Gln Ile
            260                 265                 270

Tyr Arg Val Thr Trp Phe Ile Ser Trp Ser Pro Cys Phe Ser Trp Gly
        275                 280                 285

Cys Ala Gly Glu Val Arg Ala Phe Leu Gln Glu Asn Thr His Val Arg
    290                 295                 300

Leu Arg Ile Phe Ala Ala Arg Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys
305                 310                 315                 320

Glu Ala Leu Gln Met Leu Arg Asp Ala Gly Ala Gln Val Ser Ile Met
                325                 330                 335

Thr Tyr Asp Glu Phe Glu Tyr Cys Trp Asp Thr Phe Val Tyr Arg Gln
            340                 345                 350
```

```
Gly Cys Pro Phe Gln Pro Trp Asp Gly Leu Glu Glu His Ser Gln Ala
        355                 360                 365


Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln Asn Gln Gly Asn
    370                 375                 380


<210>  1022
<211>  386
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic

<400>  1022

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Arg Met Tyr Arg Asp Thr
1               5                   10                  15


Phe Tyr Asp Asn Phe Glu Asn Glu Pro Ile Leu Tyr Gly Arg Ser Tyr
            20                  25                  30


Thr Trp Leu Cys Tyr Glu Val Lys Ile Lys Arg Gly Arg Ser Asn Leu
        35                  40                  45


Leu Trp Asp Thr Gly Val Phe Arg Gly Pro Val Leu Pro Lys Arg Gln
    50                  55                  60


Ser Asn His Arg Gln Glu Val Tyr Phe Arg Phe Glu Asn His Ala Glu
65                  70                  75                  80


Met Cys Phe Leu Ser Trp Phe Cys Gly Asn Arg Leu Pro Ala Asn Arg
                85                  90                  95


Arg Phe Gln Ile Thr Trp Phe Val Ser Trp Asn Pro Cys Leu Pro Cys
            100                 105                 110


Val Val Lys Val Thr Lys Phe Leu Ala Glu His Pro Asn Val Thr Leu
        115                 120                 125
```

```
Thr Ile Ser Ala Ala Arg Leu Tyr Tyr Tyr Arg Asp Arg Asp Trp Arg
    130                 135                 140

Trp Val Leu Leu Arg Leu His Lys Ala Gly Ala Arg Val Lys Ile Met
145                 150                 155                 160

Asp Tyr Glu Asp Phe Ala Tyr Cys Trp Glu Asn Phe Val Cys Asn Glu
                165                 170                 175

Gly Gln Pro Phe Met Pro Trp Tyr Lys Phe Asp Asp Asn Tyr Ala Ser
            180                 185                 190

Leu His Arg Thr Leu Lys Glu Ile Leu Arg Asn Pro Met Glu Ala Met
        195                 200                 205

Tyr Pro His Ile Phe Tyr Phe His Phe Lys Asn Leu Leu Lys Ala Cys
    210                 215                 220

Gly Arg Asn Glu Ser Trp Leu Cys Phe Thr Met Glu Val Thr Lys His
225                 230                 235                 240

His Ser Ala Val Phe Arg Lys Arg Gly Val Phe Arg Asn Gln Val Asp
                245                 250                 255

Pro Glu Thr His Cys His Ala Glu Arg Cys Phe Leu Ser Trp Phe Cys
            260                 265                 270

Asp Asp Ile Leu Ser Pro Asn Thr Asn Tyr Glu Val Thr Trp Tyr Thr
        275                 280                 285

Ser Trp Ser Pro Cys Pro Glu Cys Ala Gly Glu Val Ala Glu Phe Leu
    290                 295                 300

Ala Arg His Ser Asn Val Asn Leu Thr Ile Phe Thr Ala Arg Leu Cys
305                 310                 315                 320
```

Tyr Phe Trp Asp Thr Asp Tyr Gln Glu Gly Leu Cys Ser Leu Ser Gln
325 330 335

Glu Gly Ala Ser Val Lys Ile Met Gly Tyr Lys Asp Phe Val Ser Cys
340 345 350

Trp Lys Asn Phe Val Tyr Ser Asp Asp Glu Pro Phe Lys Pro Trp Lys
355 360 365

Gly Leu Gln Thr Asn Phe Arg Leu Leu Lys Arg Arg Leu Arg Glu Ile
370 375 380

Leu Gln
385

<210> 1023
<211> 229
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1023

Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1 5 10 15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
20 25 30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
35 40 45

Ser Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val
50 55 60

Asn Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr
65 70 75 80

```
Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
                85                  90                  95

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val Thr Leu
            100                 105                 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg Asn Arg
        115                 120                 125

Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met
    130                 135                 140

Thr Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn Tyr Ser
145                 150                 155                 160

Pro Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp Val Arg
                165                 170                 175

Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
            180                 185                 190

Leu Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile
        195                 200                 205

Ala Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp
    210                 215                 220

Ala Thr Gly Leu Lys
225
```

<210> 1024
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1024
ccacacccaa aaggccacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1025
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1025
cccaccagcu cagcuccacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1026
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1026
ucccuagcag gaucucauag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1027
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1027
ggcucaaaca cagcgaccuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1028
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1028
ggcacaccag uguggccuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1029
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1029
agguggccga gacccucagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1030
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1030
accugcucua ccccaggccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1031
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1031
cauagaggau gguggcagac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1032
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1032
ccacguggag cugagcuggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1033
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1033
ggagaaugac gaguggaccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1034
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1034
ccaguguggc cuuuugggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1035
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 1035
caaacacagc gaccucgggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1036
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1036
uaccauggcc aucaacacaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1037
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1037
gcgcugacga ucugggugac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1038
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1038
cacggacccg cagccccuca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1039
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1039
ucaaacacag cgaccucggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1040
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1040
cagggaagaa gccuguggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1041
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1041
caguguggcc uuuugggugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1042
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1042
accacgugga gcugagcugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1043
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1043
gagggcgggc ugcuccuuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1044
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1044
agcucagcuc cacgugguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1045
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1045
gaacaaggug uucccacccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1046
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1046
gcacaccagu guggccuuuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1047
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1047
gagcuggugg gugaauggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1048
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1048
ggcugcuccu ugaggggcug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1049
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1049
cgggugggaa caccuuguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1050
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1050
cagcucagcu ccacgugguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1051
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1051
gacgaucugg gugacggguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1052
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1052
uagcaggauc ucauagagga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1053
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1053
gaccagcaca gcauacaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1054
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1054
cugaccacgu ggagcugagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1055
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1055
ccaacagugu ccuaccagca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1056
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1056
cuggugggug aaugggaagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1057
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1057
cuaugagauc cugcuaggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1058
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1058
caggaucuca uagaggaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1059
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1059
ggccacccug uaugcugugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1060
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<400> 1060
caacaguguc cuaccagcaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1061
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1061
agcugagcug gugggugaau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1062
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1062
aacagugucc uaccagcaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1063
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1063
aggcuucuuc ccugaccacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1064
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1064
ucuucugcag gucaagagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1065
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1065
gagcugagcu gguggggugaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1066
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1066
ggucagcgcc cuuguguuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1067
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1067
agaucgucag cgccgaggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1068
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1068
gcugcuccuu gaggggcugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1069
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1069
guaucuggag ucauugaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1070
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1070
auccucuaug agauccugcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1071
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1071
uccucuauga gauccugcua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1072
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1072
gcaguaucug gagucauuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1073
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1073
gcugaccagc acagcauaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1074
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1074
acaggguggc cuucccuagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1075
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1075
cgcugaccag cacagcauac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1076
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1076
uccuaccugu cagagcccca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1077
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1077
gagccccaag guaaaaaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1078
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1078
agccccaagg uaaaaaggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1079
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1079
uuucccggcc uuuuuaccuu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1080
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1080
cuuucccggc cuuuuuaccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1081
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1081
acacugugag cugccuggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1082
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1082
cacacuguga gcugccuggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1083
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1083
guggcacacu gugagcugcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1084
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1084
gagauugagc ucuacucagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1085
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1085
agauugagcu cuacucaggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1086
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1086
caggugggcc cuccucccuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1087

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1087
cggaagagac cagagggagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1088
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1088
uaccggaaga gaccagaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1089
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1089
agggaggcuu augccaauau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1090
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1090
ggcuuaugcc aauaucggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1091
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1091
ugacugcgcu uuuccuuguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1092
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1092
gacugcgcuu uuccuugucu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1093
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1093
gcuuuuccuu gucugggcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1094
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 1094
ccaguuccgc ugcccagaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1095
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 1095
gaccugaagc acuggaagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1096
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 1096
gcacuggaag ccaggugugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1097
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1097
cacuggaagc caggugugca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1098
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1098
ggaagccagg ugugcagggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1099
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1099
agccaggugu gcagggcagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1100
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1100
gcccaccugc ccugcacacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1101
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1101
gggcucagcu gugaggaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1102
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1102
ggggcucagc ugugaggaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1103
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1103
gaccagauuc ccaguauguu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1104
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1104
ccagauuccc aguauguuag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1105
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1105
cagauuccca guauguuagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1106
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1106
ucccaguaug uuaggggggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1107
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1107
uccaagcccc cuaacauacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1108
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1108
cuccaagccc ccuaacauac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1109
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1109
aaaggcacug caagagacaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1110
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1110
uccaguauau ucaucuacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1111
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1111
uucaucuacc auggugagug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1112
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1112
ucaucuacca uggugagugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1113
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 1113
caucuaccau ggugagugcg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1114
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 1114
accaugguga gugcggggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1115
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred

embodiments

<400> 1115
gccaggcccc gcacucacca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1116
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1116
ccacucucca cccccaaugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1117
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1117
cuccaccccc aauguaggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1118
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1118
caccucaccu acauuggggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1119
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1119
gggcaccuca ccuacauugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1120
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1120
uggggcaccu caccuacauu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1121
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1121
cuggggcacc ucaccuacau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1122
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1122

accucccgag caaacaugac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1123
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1123
ccgagcaaac augacaggua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1124
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1124
guaccucuca caggcccuaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1125
<211> 100
<212> RNA

<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1125
aguaccucuc acaggcccua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1126
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1126
gacgucuugu gcucuggaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1127
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature

<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1127
aacaagcuuc caaaauggcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1128
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1128
gagaucccgc aucacucacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1129
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1129
gccccuggcc uuugcagagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1130
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1130
ucagcugugu cacccguuuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1131
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1131
gcugugucac ccguuucagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1132
<211> 100
<212> RNA
<213> Artificial Sequence

<220>

<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1132
cacccguuuc agguggggug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1133
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1133
gucugaggcc cucccuccac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1134
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1134
ucugaggccc ucccuccaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1135
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1135
caaggcagcc cuguggaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1136
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1136
gcucaaggca gcccugugga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1137

<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1137
cgcucaaggc agcccugugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1138
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1138
uaacauugcc uguucucucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1139
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1139
cuucucgucc uggagagaac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1140
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1140
uuccgaggaa cuucucgucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1141
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1141
acccuugcuc uuugccuccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1142
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1142
uugcucuuug ccuccuaggc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1143
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1143
ugcucuuugc cuccuaggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1144
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 1144
ccugagacag ggcccagccu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1145
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 1145
aucugauucc accugcagcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100

<210> 1146
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed

for detailed description of substitutions and preferred embodiments

<400> 1146
acucagcgca uccaggcugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1147
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1147
aaacccucaa gugagugagc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1148
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1148
aacccucaag ugagugagcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1149
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1149
ggcccagcuc acucacuuga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1150
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1150
aggcccagcu cacucacuug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1151
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1151
cagacugcgg ggacacagug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1152
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1152
cugcaucccu gcucaggcua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1153
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1153
ccugcucagg cuaaggugag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1154
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1154
cugcucaggc uaaggugagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1155
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1155
cagcaccuga ccgguauccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1156
<211> 100

<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1156
guacaagcug ucggaaacag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1157
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1157
ggagacgcug gcguaagucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1158
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>

<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1158
ggcguaaguc caggcaaccc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1159
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1159
cuccacccac caggguugcc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1160
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1160
ugagucccau ccccccuugc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1161
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1161
ccacauccug caaggggga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1162
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed for detailed description of substitutions and preferred embodiments

<400> 1162
gggcguccac auccugcaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu 100

<210> 1163
<211> 100
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<223> contains 2'-O-Me and PS modifications; see specification as filed
       for detailed description of substitutions and preferred
       embodiments

<400> 1163
gugggcgucc acauccugca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc   60

cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu   100