



(51) International Patent Classification:

C12N 9/12 (2006.01) C12N 15/10 (2006.01)

(21) International Application Number:

PCT/US2018/046989

(22) International Filing Date:

17 August 2018 (17.08.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/547,677	18 August 2017 (18.08.2017)	US
62/628,484	09 February 2018 (09.02.2018)	US
62/638,684	05 March 2018 (05.03.2018)	US
62/677,527	29 May 2018 (29.05.2018)	US

(71) Applicant: MODERNATX, INC. [US/US]; 200 Technology Square, Cambridge, MA 02139 (US).

(72) Inventors: RABIDEAU, Amy, E.; 121 Pond Street, Unit 1, Waltham, MA 02451 (US). DOUSIS, Athanasios; 348 Franklin Street, #4B, Cambridge, MA 02139 (US). RAVICHANDRAN, Kanchana; 950 Massachusetts Av-

enue, Apt. 420, Cambridge, MA 02139 (US). **HOBERT, Elissa**; 205B Rindge Avenue, Cambridge, MA 02140 (US).

(74) Agent: LOCKHART, Helen, C. et al.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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,

(54) Title: RNA POLYMERASE VARIANTS

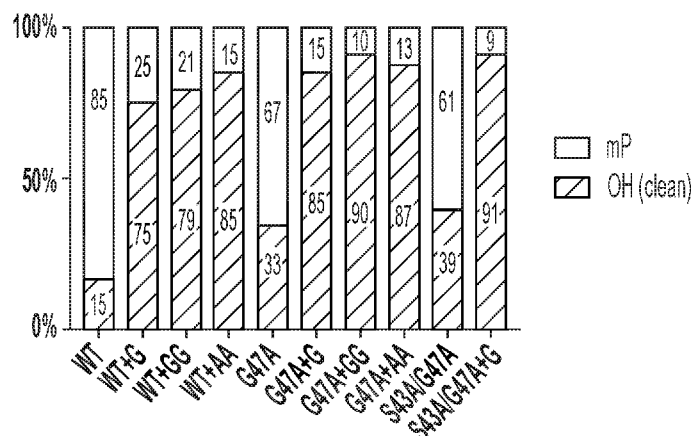


Figure 24C

(57) **Abstract:** The present disclosure provides, in some aspects, variant RNA polymerases, the use of which increases transcription efficiency while reducing the number of double-stranded RNA contaminates and run-on transcripts produced during an *in vitro* transcription reaction. In some embodiments, the RNA polymerase variants (e.g. T7 RNA polymerase variants, include mutations to residues that have a high-helix propensity (e.g., alanine) or that constrain the backbone flexibility to specifically match that of the elongation complex. Amino acid substitutions within the C-helix, including S43 A and G47A, and C-linker regions of T7 RNA polymerase, as well as the addition of a C-terminal glycine residue are described. Co-transcription methods using cap analogs are also described.



MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
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))*

RNA POLYMERASE VARIANTS

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application number 62/547,677, filed August 18, 2017, U.S. provisional application number 62/628,484, filed February 9, 2018, and U.S. provisional application number 62/638,684, filed March 5, 2018, U.S. provisional application number 62/677,527, filed May 29, 2018, each of which is incorporated by reference herein in its entirety.

BACKGROUND

In vitro transcription (IVT) uses bacteriophage DNA-dependent ribonucleic acid (RNA) polymerases (e.g., SP6, T3 and T7) to synthesize template-directed mRNA transcripts. Problems in the IVT reaction can result in complete failure (e.g., no transcript generated) or in transcripts that are the incorrect size (e.g., shorter or longer than expected). Specific problems associated with IVT reactions include, for example, abortive (truncated) transcripts, run-on transcripts, polyA tail variants/3' heterogeneity, mutated transcripts, and/or double-stranded contaminants produced during the reactions.

RNA polymerases exhibit three phases of transcription – initiation, elongation and termination. During the initiation phase, the RNA polymerase binds to a specific promoter DNA sequence, opens the DNA duplex and feeds the template strand into the active site. T7 RNA polymerase, for example, forms a structure referred to as initiation complex, which includes a six-helix bundle sub-domain (the promoter binding domain) that interacts with the promoter to initiate DNA duplex melting. While bound to the promoter, the polymerase produces many short (truncated) transcripts from 2-12 nucleotides (nt) in length, a process often referred to as abortive synthesis/initiation. The truncated RNA transcripts cannot be converted to full-length transcripts by RNA polymerase and become by-products that accumulate during transcription. After the transition to the elongation phase and release of the promoter, the polymerase proceeds down the DNA template producing a full-length RNA transcript.

During the elongation phase, RNA polymerase often continues to transcribe DNA beyond the point at which termination should be initiated, generating longer than expected RNA transcripts (“run-on transcripts”). T7 RNA polymerase, for example, adds nucleotides to the end of a transcript before ‘falling off’ the template. Studies suggest that more than 70% of transcripts generated by T7 RNA polymerase in vitro may be run-on transcripts. In

some cases, these aberrant RNA products are twice the length of the encoded sequence. Because run-on transcription is stochastic, there is often great 3' heterogeneity among products in a given IVT reaction. This 3' heterogeneity is problematic for downstream applications, such as ligation reactions, which are dependent on RNA transcripts of a defined length and/or nucleotide composition.

SUMMARY

During the initiation of transcription, RNA polymerase balances two opposing phases. The polymerase must first associate with the promoter tightly enough to permit dissociation of the two DNA strands (one of which is the template strand) and commence transcription. The polymerase must then release the promoter and enter a highly processive elongation phase. Competition between these two phases leads to the production of abortive transcripts. The polymerase repeatedly tries to clear the promoter but is unable to overcome the transition barrier and releases short (abortive) RNA products. Conversely, the polymerase often fails to ‘run-off’ the DNA template, generating a population of RNA transcripts with 3' heterogeneity. Provided herein are variant RNA polymerases, which increase transcription efficiency and 3' homogeneity, run-on transcripts, double-stranded contaminants, or any combination thereof, produced during an in vitro transcription (IVT) reaction, for example.

During the transition from initiation to elongation, RNA polymerase undergoes a conformational change requiring a large rearrangement of the amino-terminal domain (N-terminal domain) (see, e.g., Bandwar, RP et al. *Journal of Biological Chemistry* 282, 22879–22886 (2009); Guillerez, J. et al. *Proc National Acad Sci* 102, 5958–5963 (2005); Durniak, K. et al. *Science (New York, N.Y.)* 322, 553 (2008); and Tahirov, T. H. et al. *Nature* 420, 43–50 (2002), each of which is incorporated herein by reference. Within this N-terminal domain is a “C-helix” (e.g., amino acids 28-71 of T7 RNA polymerase) and a “C-linker” (e.g., amino acids 258-266 of T7 RNA polymerase), each of which contains subregions (amino acids 42-47 and 257-262, respectively) that undergo a conformational change from a loop structure to a helix structure as the RNA polymerase transitions from an initiation complex to an elongation complex (see, e.g., FIG. 10), abolishing the promoter binding site, enlarging the active site and creating an exit tunnel for the RNA transcript. Mutations to subregions in the C-helix structure and/or the C-linker linker region may drive the conformation equilibrium toward the elongation complex by increasing the thermodynamic stability of the elongation complex relative to the initiation complex. Without being bound by theory, it is also thought that mutations in select regions, such as the C-helix structure and/or the C-linker linker

region, can change how the exit tunnel of the polymerase interacts with the transcript during elongation, causing a significant reduction in transcription errors (e.g., run-on transcripts). Thus, the variant polymerases of the present disclosure include a (at least one) mutation in the C-helix and/or C-linker to drive the conformation equilibrium toward the elongation complex.

Other regions of the N-terminal domain that contain loop structures that transition to helix structures as the polymerase proceeds from initiation to elongation may be mutated (point mutation, referred to as a substitution), as provided herein. Non-limiting examples of such loop-to-helix regions include regions that span amino acids 55-73, 164-169, or 176-187 of T7 RNA polymerase (e.g., SEQ ID NO:1), or regions of other single subunit RNA polymerases (see, e.g., Cermakian, N. et al. *J Mol Evol* 45, 671–681 (1997), incorporated herein by reference) that are homologous (e.g., at least 80%, at least 90%, at least 95%, or at least 98% identical to) to the foregoing regions. Non-limiting examples of other single subunit RNA polymerases include T3 RNA polymerase, K11 RNA polymerase, and SP6 RNA polymerase.

In some embodiments, the RNA polymerase variants (e.g. T7 RNA polymerase variants, include mutations to residues that have a high-helix propensity (e.g., alanine) or that constrain the backbone flexibility to specifically match that of the elongation complex.

Further provided herein, are RNA polymerase variants that include at least one additional amino acid in the C terminus of the polymerase. For example, T7 RNA polymerase variants may include at least one additional amino acid, such as an additional glycine (G) in the C terminus of the polymerase. Surprisingly, the population of RNA transcripts produced using a T7 RNA polymerase modified to include an additional G in the C-terminal “foot” region (e.g., the region comprising “FAFA” amino acids at positions 880-883 of wild-type T7 RNAP) exhibits less 3' heterogeneity. For example, as shown in FIG. 23, a T7 RNA polymerase variant that includes a C-terminal glycine (i.e., ...FAFAG) produces a RNA transcript population in which at least 85% of the transcripts are homogeneous at the 3' end. This data is particularly unexpected, given that previous studies have shown that C-terminal additions/insertions resulted in loss of function of T7 polymerase (Gardner LP, et al. *Biochemistry* 36, 2908–2918 (1997) and Gross L, et al. *Journal of Molecular Biology* 228, 488–505 (1992)).

Also provided herein are methods of capping a ssRNA (e.g., mRNA) with cap analogs (e.g., trinucleotides) co-transcriptionally (co-transcriptional capping methods) in an in vitro

transcription assay using a T7 RNA polymerase variant described herein (e.g., T7 RNA polymerase variant G47A or G47A* C-terminal variant). Efficient co-transcriptional capping typically includes a double-stranded DNA (dsDNA) template that initiates transcription with 5'ATP and equimolar concentrations of NTPs and trinucleotides. Under these conditions, T7 RNA polymerase exhibits severely reduced initiation activity with 5'ATP. Unexpectedly, the data provided herein shows that, in the presence of GAG trinucleotides (e.g., m⁷GpppA_{2'}OMe pG), for example, the limited initiation activity of T7 RNA polymerase with 5'ATP drives initiation with the trinucleotide rather than 5'ATP, generating capped RNA products co-transcriptionally. Surprisingly, in some embodiments, greater than 90% of the RNA produced comprises single-stranded full-length transcripts, at least 90% of the RNA produced includes a functional cap, and the RNA produced does not elicit a substantial cytokine response, even in the absence of post-IVT purification.

Thus, some aspects of the present disclosure provide RNA polymerase variants comprising at least one amino acid substitution, relative to wild-type polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex. In some embodiments, at least one amino acid substitution is estimated to have a more negative change in folding free energy in the elongation complex than in the initiation complex. In some embodiments, at least one amino acid substitution has a high-helix propensity, relative to wild-type amino acid. In some embodiments, the RNA polymerase variant comprises an (at least one) additional amino acid residue at the C terminus. For example, an RNA polymerase variant (e.g., a T7 RNA polymerase variant) may comprise a glycine (G) at the C terminus.

In some embodiments, the RNA polymerase is a T7 RNA polymerase. In some embodiments, the RNA polymerase is a T3 RNA polymerase. In some embodiments, the RNA polymerase is a SP6 RNA polymerase. In some embodiments, the RNA polymerase is a K11 RNA polymerase.

In some embodiments, at least one loop structure is in the C-helix structure. In some embodiments, at least one loop structure is in the C-linker structure. In some embodiments, at least one loop structure is present in a region within amino acids 55-73, 164-169, or 176-187 of T7 RNA polymerase, or a RNA polymerase homologous to T7 RNA polymerase.

In some embodiments, at least one amino acid substitution is at least one high-helix propensity amino acid substitution. For example, at least one high-helix propensity amino

acid substitution may be selected from alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate. In some embodiments, at least one high-helix propensity amino acid substitution is alanine. In some embodiments, at least one high-helix propensity amino acid substitution is isoleucine. In some embodiments, at least one high-helix propensity amino acid substitution is leucine. In some embodiments, at least one high-helix propensity amino acid substitution is arginine. In some embodiments, at least one high-helix propensity amino acid substitution is methionine. In some embodiments, at least one high-helix propensity amino acid substitution is lysine. In some embodiments, at least one high-helix propensity amino acid substitution is glutamine. In some embodiments, at least one high-helix propensity amino acid substitution is glutamate.

In some embodiment, a T7 RNA polymerase is modified to include at least one amino acid substitution of a high-helix propensity amino acid in at least one position selected from E42 (e.g., E42R), S43 (e.g., S43A), Y44 (e.g., Y44A), E45 (e.g., E45R/L), M46 (e.g., M46A), G47 (e.g., G47A), A255 (e.g., A255K/Q/Y/I), R257 (e.g., R257A), A258 (e.g., A258R/E/L), G259 (e.g., G259A), A260 (e.g., A260R/E/L), L261 (e.g., L261A) and A262 (e.g., A262R/E/L). The T7 RNA polymerase may further comprise, in some embodiments, one or more additional amino acid substitutions (in addition to at least one high-helix propensity amino acid substitution). Thus, the present disclosure encompasses the further modification of existing (e.g., currently-available and/or commercially-available) T7 RNA polymerase variants with one or more high-helix propensity amino acid substitutions as provided herein.

In some embodiments, a T7 RNA polymerase comprises an amino acid sequence of SEQ ID NO:1 modified to include at least one amino acid substitution of a high-helix propensity amino acid at a position selected from E42 (e.g., E42R), S43 (e.g., S43A), Y44 (e.g., Y44A), E45 (e.g., E45R/L), M46 (e.g., M46A) and G47 (e.g., G47A). In some embodiments, at least one amino acid substitution comprises S43A. In some embodiments, at least one amino acid substitution comprises G47A.

In some embodiments, a T7 RNA polymerase comprises an amino acid sequence of SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:294, SEQ ID NO:295, or SEQ ID NO:296 modified to include at least one amino acid substitution of a high-helix propensity amino acid at a position selected from E42 (e.g., E42R), S43 (e.g., S43A), Y44 (e.g., Y44A), E45 (e.g., E45R/L), M46 (e.g., M46A) and G47 (e.g., G47A). In some embodiments, at least one amino acid substitution comprises S43A. In some embodiments, at least one amino acid substitution comprises G47A.

In some embodiments, a T7 RNA polymerase comprises an amino acid sequence of SEQ ID NO:1 modified to include at least one amino acid substitution of a high-helix propensity amino acid at a position selected from A255 (e.g., A255K/Q/Y/I), R257 (e.g., R257A), A258 (e.g., A258R/E/L), G259 (e.g., G259A), A260 (e.g., A260R/E/L), L261 (e.g., L261A) and A262 (e.g., A262R/E/L). In some embodiments, at least one amino acid substitution comprises R257A. In some embodiments, at least one amino acid substitution comprises G259A.

In some embodiments, a T7 RNA polymerase comprises an amino acid sequence of SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:294, SEQ ID NO:295, or SEQ ID NO:296 modified to include at least one amino acid substitution of a high-helix propensity amino acid at a position selected from A255 (e.g., A255K/Q/Y/I), R257 (e.g., R257A), A258 (e.g., A258R/E/L), G259 (e.g., G259A), A260 (e.g., A260R/E/L), L261 (e.g., L261A) and A262 (e.g., A262R/E/L). In some embodiments, at least one amino acid substitution comprises R257A. In some embodiments, at least one amino acid substitution comprises G259A.

Also provided herein, in some aspects, are T7 RNA polymerases comprising an amino acid sequence of SEQ ID NO:1 modified to include an amino acid substitution of a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at position G47, S43, R257, or G259. Further provided herein, in some aspects, are T7 RNA polymerases comprising an amino acid sequence of SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:294, SEQ ID NO:295, or SEQ ID NO:296 modified to include an amino acid substitution of a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at position G47, S43, R257, or G259.

In some embodiments, T7 RNA polymerases of the present disclosure comprise an amino acid sequence of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5. In some embodiments, T7 RNA polymerases of the present disclosure comprise an amino acid sequence of SEQ ID NO:107 or 108, SEQ ID NO:109 or 110, SEQ ID NO:111 or 112, or SEQ ID NO:113 or 114.

In some embodiments, a T7 RNA polymerase comprises at least one additional C-terminal amino acid. In some embodiments, a T7 RNA polymerase comprises at least two additional C-terminal amino acids.

In some embodiments, the at least two additional C-terminal amino acids comprise the same type of amino acid (e.g., all Gly, all Ala). In some embodiments, the at least two

additional C-terminal amino acids comprise at least two different types of amino acids (e.g., GlyAla, AlaGly).

In some embodiments, a T7 RNA polymerase comprises at least three additional C-terminal amino acids. In some embodiments, the at least three additional C-terminal amino acids comprise at least two or at least three of the same type or different types of amino acids (e.g., GlyGlyGly, AlaAlaAla).

In some embodiments, a T7 RNA polymerase comprises 1 to 10 (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10) additional C-terminal amino acids. In some embodiments, a T7 RNA polymerase comprises 1 to 5 additional C-terminal amino acids.

In some embodiments, a T7 RNA polymerase comprises a C terminus that comprises a FAFAX_n motif, wherein X is any amino acid and n is any integer greater than zero. In some embodiments, X is glycine (G). In some embodiments, n is 1, 2, 3, 4, or 5. It should be understood that in embodiments, where n is greater than 1, such that the C terminal motif is FAFXX, FAFAXXX, FAFAXXXX, or FAFAXXXXX, for example, the X's may be the same amino acid or they may be different amino acids. For example, the C-terminal motif may be FAFAGG or FAFAGGG, or the C-terminal motif may be FAFAGA, FAFAGC, FAFAGAA, FAFAGAG, FAFAGAC, etc. Other C-terminal amino acid combinations may be used.

In some embodiments, a T7 RNA polymerase comprises a C terminus that comprises a FAFAG motif.

In some embodiments, a T7 RNA polymerase comprises a XAFAX_n motif, a FXFAX_n motif, FAXAX_n motif, or a FAFXX_n motif, wherein each X is any amino acid and n is any integer greater than zero. Thus, the present disclosure comprises various C-terminal F⁸⁸⁰A⁸⁸¹F⁸⁸²A⁸⁸³ motifs, wherein the amino acid at one or more of positions 880, 881, 882, or 883 (e.g., relative to wild-type T7 RNAP, e.g., SEQ ID NO:1) is modified to include at least one amino acid substitution, with or without an additional C-terminal amino acid (X_n).

In some aspects, the present disclosure provides RNA polymerases comprising at least one additional C-terminal amino acid relative to a corresponding wild-type RNA polymerase. In some embodiments, the RNA polymerase is selected from T7 RNA polymerases, T3 RNA polymerases, and SP6 RNA polymerases.

In some embodiments, the RNA polymerase further comprises at least one additional amino acid substitution. In some embodiments, the RNA polymerase further comprises an amino acid substitution corresponding to an amino acid substitution of SEQ ID NO:1 selected from G47A, S43A, R257A, and G259A.

In some embodiments, the RNA polymerase is a T7 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, at least 98%, or at least 99% identical to SEQ ID NO:1 modified to include an amino acid substitution at position 43±1 (e.g., 42, 43, or 44), 47±1 (e.g., 46, 47, or 48), 257±1 (e.g., 256, 257, or 258), and/or 259±1 (e.g., 258, 259, or 260), optionally wherein the amino acid substitution is alanine (A). In some embodiments, the RNA polymerase is a T3 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:6 modified to include an amino acid substitution at a position corresponding to wild-type T7 RNA polymerase position 43±1, 47±1, 257±1, and/or 259±1 (based on a sequence or structural alignment), optionally wherein the amino acid substitution is alanine (A). In some embodiments, the RNA polymerase is a SP6 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, at least 98%, or at least 99% identical to SEQ ID NO:7 modified to include an amino acid substitution at a position corresponding to wild-type T7 RNA polymerase position 43±1, 47±1, 257±1, and/or 259±1 (based on a sequence or structural alignment), optionally wherein the amino acid substitution is alanine (A).

Some aspects of the present disclosure provide T7 RNA polymerases comprising an amino acid sequence at least 90%, at least 95%, at least 98%, or at least 99% identical to SEQ ID NO:1 modified to include an amino acid substitution at position 43, 47, 257, and/or 259, optionally wherein the amino acid substitution is alanine (A).

Other aspects of the present disclosure provide T3 RNA polymerases comprising an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:6 modified to include an amino acid substitution at a position corresponding to wild-type T7 RNA polymerase position 43, 47, 257, and/or 259, optionally wherein the amino acid substitution is alanine (A). Thus, in some embodiments, T3 RNA polymerases comprise an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:6 modified to include an amino acid substitution at position 44, 48, 258, and/or 260, optionally wherein the amino acid substitution is alanine (A).

Yet other aspects of the present disclosure provide SP6 RNA polymerases comprising an amino acid sequence at least 90%, at least 95%, or at least 98% % identical to SEQ ID NO:7 modified to include an amino acid substitution at a position corresponding to wild-type T7 RNA polymerase position 43, 47, 257, and/or 259, optionally wherein the amino acid substitution is alanine (A). Thus, in some embodiments, SP6 TNA polymerases comprise an amino acid sequence at least 90%, at least 95%, or at least 98% % identical to SEQ ID NO:7

modified to include an amino acid substitution at position 15, 19, 230, and/or 232, optionally wherein the amino acid substitution is alanine (A).

The present disclosure also provides methods of producing RNA comprising contacting a DNA template with a RNA polymerase variant as described herein under conditions that result in the production of RNA transcript (e.g., under IVT condition).

The present disclosure further provides methods of performing an IVT reaction, comprising contacting a DNA template with a RNA polymerase variant as provided herein in the presence of nucleoside triphosphates (NTPs) and buffer under conditions that result in the production of RNA transcripts.

In some embodiments, the RNA transcript produced, when delivered to cells, optionally in unpurified form, stimulates cytokine response that is at least 50% (e.g., at least 60%, at least 70%, at least 80%, at least 95% or at least 98%) lower relative to RNA produced using wild-type RNA polymerase under the same IVT conditions.

In some embodiments, the concentration of dsRNA transcript produced by IVT is at least 50% (e.g., at least 60%, at least 70%, at least 80%, at least 95% or at least 98%) lower relative to dsRNA transcript produced using wild-type polymerase.

In some embodiments, less than 20% (e.g., less than 15%, less than 10%, less than 5%) of the RNA transcripts produced exhibit 3' heterogeneity.

In some embodiments, less than 50% (e.g., less than 40%, less than 30%, less than 20%, less than 10%) of the RNA transcript produced is truncated RNA transcript.

In some embodiments, less than 50% (e.g., less than 40%, less than 30%, less than 20%, less than 10%) of the RNA transcript produced is run-on RNA transcript.

In some embodiments, the amount of full-length RNA transcript produced is at least 15 times greater than the amount of the DNA template

In some embodiments, the ratio of truncated RNA transcript:full-length RNA transcript produced is less than 1:1.

In some embodiments, the RNA transcript produced has less than 1 mutation per 100 nucleotides relative to the DNA template.

The present disclosure, in some aspects provides nucleic acids encoding the RNA polymerase variants and, in some embodiments, vectors (e.g., plasmids) and/or host cells (e.g., mammalian cells, e.g., human cells) comprising the nucleic acids.

The RNA transcripts produced by the methods of the present disclosure are also provided. In some embodiments, the RNA transcripts (e.g., mRNA) are formulated in a lipid nanoparticle. The lipid nanoparticle may comprise, for example, a molar ratio of 20-60%

ionizable amino lipid, 5-25% non-cationic lipid, 25-55% sterol, and 0.5-15% PEG-modified lipid. See, e.g., WO 2017/070624, published 27 April 2017, incorporated herein by reference.

Other compositions and kits comprising the RNA polymerase variants are encompassed herein.

Also provided herein are co-transcriptional capping methods for ribonucleic acid (RNA) synthesis, the methods comprising reacting a polynucleotide template with a T7 RNA polymerase variant, nucleoside triphosphates, and a cap analog under in vitro transcription reaction conditions to produce RNA transcript.

In some embodiments, greater than 80%, greater than 85%, or greater than 90% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 95% of the RNA transcript produced includes a functional cap.

In some embodiments, the nucleoside triphosphates comprise unmodified or modified ATP, modified or unmodified UTP, modified or unmodified GTP, and/or modified or unmodified CTP.

In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:1 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1) modified to include at least one amino acid substitution of a high-propensity amino acid at a position selected from E42, S43, Y44, E45, M46, G47, A255, R257, and G259. In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:1 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1) modified to include amino acid substitution of G47A. In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:1 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1) modified to include amino acid substitution of S43A.

In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:99 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99) modified to include at least one amino acid substitution of a high-propensity amino acid at a position selected from E42, S43, Y44, E45, M46, G47, A255, R257, and G259. In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:99 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at

least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99) modified to include amino acid substitution of G47A. In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:99 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99) modified to include amino acid substitution of S43A.

In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:100 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:100) modified to include at least one amino acid substitution of a high-propensity amino acid at a position selected from E42, S43, Y44, E45, M46, G47, A255, R257, and G259. In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:100 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:100) modified to include amino acid substitution of G47A. In some embodiments, the T7 polymerase variant comprises an amino acid sequence of SEQ ID NO:100 (or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:100) modified to include amino acid substitution of S43A.

In some embodiments, the nucleoside triphosphates and cap analog are present in the reaction at equimolar concentrations. In some embodiments, a molar ratio of cap analog to nucleoside triphosphates in the reaction is greater than 1:1. In some embodiments, a molar ratio of cap analog to nucleoside triphosphates in the reaction is less than 1:1.

In some embodiments, the cap analog is a dinucleotide cap, a trinucleotide cap, or a tetranucleotide cap. In some embodiments, the cap analog is a trinucleotide cap.

In some embodiments, the trinucleotide cap comprises a sequence selected from the following sequences: GAA, GAC, GAG, GAU, GCA, GCC, GCG, GCU, GGA, GGC, GGG, GGU, GUA, GUC, GUG, and GUU.

In some embodiments, the trinucleotide cap comprises a sequence selected from the following sequences: m⁷GpppApA, m⁷GpppApC, m⁷GpppApG, m⁷GpppApU, m⁷GpppCpA, m⁷GpppCpC, m⁷GpppCpG, m⁷GpppCpU, m⁷GpppGpA, m⁷GpppGpC, m⁷GpppGpG, m⁷GpppGpU, m⁷GpppUpA, m⁷GpppUpC, m⁷GpppUpG, and m⁷GpppUpU.

In some embodiments, the trinucleotide cap comprises a sequence selected from the following sequences: m⁷G_{3'OMe}pppApA, m⁷G_{3'OMe}pppApC, m⁷G_{3'OMe}pppApG,

$m^7G_{3'OMe}pppApU$, $m^7G_{3'OMe}pppCpA$, $m^7G_{3'OMe}pppCpC$, $m^7G_{3'OMe}pppCpG$,
 $m^7G_{3'OMe}pppCpU$, $m^7G_{3'OMe}pppGpA$, $m^7G_{3'OMe}pppGpC$, $m^7G_{3'OMe}pppGpG$,
 $m^7G_{3'OMe}pppGpU$, $m^7G_{3'OMe}pppUpA$, $m^7G_{3'OMe}pppUpC$, $m^7G_{3'OMe}pppUpG$, and
 $m^7G_{3'OMe}pppUpU$.

In some embodiments, the trinucleotide cap comprises a sequence selected from the following sequences: $m^7G_{3'OMe}pppA_{2'OMe}pA$, $m^7G_{3'OMe}pppA_{2'OMe}pC$, $m^7G_{3'OMe}pppA_{2'OMe}pG$,
 $m^7G_{3'OMe}pppA_{2'OMe}pU$, $m^7G_{3'OMe}pppC_{2'OMe}pA$, $m^7G_{3'OMe}pppC_{2'OMe}pC$, $m^7G_{3'OMe}pppC_{2'OMe}pG$,
 $m^7G_{3'OMe}pppC_{2'OMe}pU$, $m^7G_{3'OMe}pppG_{2'OMe}pA$, $m^7G_{3'OMe}pppG_{2'OMe}pC$,
 $m^7G_{3'OMe}pppG_{2'OMe}pG$, $m^7G_{3'OMe}pppG_{2'OMe}pU$, $m^7G_{3'OMe}pppU_{2'OMe}pA$,
 $m^7G_{3'OMe}pppU_{2'OMe}pC$, $m^7G_{3'OMe}pppU_{2'OMe}pG$, and $m^7G_{3'OMe}pppU_{2'OMe}pU$.

In some embodiments, the trinucleotide cap comprises a sequence selected from the following sequences: $m^7GpppA_{2'OMe}pA$, $m^7GpppA_{2'OMe}pC$, $m^7GpppA_{2'OMe}pG$,
 $m^7GpppA_{2'OMe}pU$, $m^7GpppC_{2'OMe}pA$, $m^7GpppC_{2'OMe}pC$, $m^7GpppC_{2'OMe}pG$,
 $m^7GpppC_{2'OMe}pU$, $m^7GpppG_{2'OMe}pA$, $m^7GpppG_{2'OMe}pC$, $m^7GpppG_{2'OMe}pG$,
 $m^7GpppG_{2'OMe}pU$, $m^7GpppU_{2'OMe}pA$, $m^7GpppU_{2'OMe}pC$, $m^7GpppU_{2'OMe}pG$, and
 $m^7GpppU_{2'OMe}pU$.

In some embodiments, the trinucleotide cap comprises a sequence selected from the following sequences: GAG, GCG, GUG, and GGG. In some embodiments, the trinucleotide cap comprises sequence GAG. In some embodiments, the trinucleotide cap comprises $m^7GpppA_{2'OMe}pG$.

In some embodiments, the polynucleotide template includes a 2'-deoxythymidine residue at template position +1. In some embodiments, the polynucleotide template includes a 2'-deoxycytidine residue at template position +1. In some embodiments, the polynucleotide template includes a 2'-deoxyadenosine residue at template position +1. In some embodiments, the polynucleotide template includes a 2'-deoxyguanosine residue at template position +1.

Also provided herein are co-transcriptional capping methods for RNA synthesis, the methods comprising reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a

trinucleotide cap comprising sequence GpppA_{2'}OmepG, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.

In some embodiments, the RNA transcript produced, when delivered to cells, optionally in unpurified form, does not stimulate a detectable cytokine response.

Further provided herein are compositions comprising an in vitro-transcribed (IVT) RNA and a pharmaceutically acceptable excipient, wherein the composition is substantially free of cytokine-inducing RNA contaminant in the absence of post-IVT purification.

A composition, in some embodiments, comprises an IVT RNA and a pharmaceutically acceptable excipient, wherein the composition has less than 5% uncapped RNA species.

In some embodiments, greater than 80%, 85%, or 90% of the IVT RNA include a functional cap. In some embodiments, greater than 95% of the IVT RNA include a functional cap.

In some embodiments, the IVT RNA is not chemically modified. In other embodiments, the IVT RNA is chemically modified.

In some embodiments, greater than 95% of the IVT RNA comprises single-stranded full-length transcripts.

In some embodiments, the RNA is produced by a process comprising: reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a trinucleotide cap comprising sequence GpppA_{2'}OmepG, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.

In some aspects, the disclosure provides T7 RNAP variants comprising an amino acid sequence of SEQ ID NOs: 294-313, wherein x is any amino acid and n is any integer, e.g., between 1 and 5 (e.g., 1, 2, 3, 4, or 5).

In some aspects, the disclosure provides a method of performing an IVT reaction, comprising contacting a DNA template with an RNA polymerase variant of the disclosure in the presence of nucleoside triphosphate and buffer under conditions that result in the

production of RNA transcripts. In some embodiments, the RNA produced, when delivered to cells, optionally in unpurified form, stimulates a cytokine response that is at least 50% lower relative to the dsRNA transcript produced using a WT. In some embodiments, less than 30% of the RNA transcripts produced by the IVT reaction exhibit 3' heterogeneity.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows HPLC chromatograms at 260 nm of human erythropoietin (hEPO) mRNA generated using wild-type (WT) T7 polymerase or T7 polymerase variants, G47A* (with C-terminal G) and S43A* (with C-terminal G).

FIG. 2, left panel, shows a graph depicting IFN β response in BJ fibroblasts transfected with chemically-unmodified hEPO RNA transcript produced using WT T7 RNA polymerase or one of the T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G), with or without reverse phase (RP) purification. FIG. 2, right panel, shows a graph depicting hEPO expression in transfected cells.

FIG. 3, top panel, shows a graph depicting IFN β response in BJ fibroblasts transfected with chemically-modified hEPO RNA transcripts (N1-methylpseudouracil ($m^1\psi$) modifications) produced using WT T7 RNA polymerase or one of the T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G), with or without reverse phase (RP) purification. FIG. 3, bottom panel, shows a graph depicting IFN β response in cells transfected with chemically-modified hEPO RNA transcripts (5-methoxy-uridine (mo^5U) modifications) produced using WT T7 RNA polymerase or one of the T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G), with or without reverse phase (RP) purification.

FIG. 4, top and bottom panels, show graphs depicting hEPO expression in cells transfected with the hEPO RNA transcripts used for FIG. 3. N1-methylpseudouracil ($m^1\psi$) modified hEPO expression is shown in the top graph, and 5-methoxy-uridine (mo^5U) modified hEPO expression is shown in the bottom graph.

FIG. 5 shows a graph of IP10 response in monocyte-derived macrophages transfected with chemically-unmodified hEPO RNA transcript produced using WT T7 RNA polymerase or T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G).

FIG. 6, top panel, shows a graph of IP10 response in monocyte-derived macrophages transfected with chemically-modified (N1-methylpseudouracil ($m^1\psi$)) hEPO RNA transcript produced using WT T7 RNA polymerase or T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G). FIG. 6, bottom panel, shows a graph of IP10

response in monocyte-derived macrophages transfected with chemically-modified (5-methoxy-uridine (mo⁵U)) hEPO RNA transcript produced using WT T7 RNA polymerase or T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G).

FIG. 7 shows graphs depicting concentration of contaminating double-stranded (ds) RNA detected using a dsRNA ELISA with 1 µg of chemically-unmodified hEPO RNA transcript (left) or 5 µg chemically-unmodified hEPO RNA transcript produced via IVT reaction using WT T7 RNA polymerase or T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G).

FIG. 8 shows graphs depicting concentration of contaminating dsRNA detected using a dsRNA ELISA with chemically-modified hEPO RNA transcript (N1-methylpseudouracil (m¹ψ) modification, top; 5-methoxy-uridine (mo⁵U) modification, bottom).

FIGS. 9A-9B show mass chromatogram results from a RNase T1 tail digest of hEPO RNA transcript produced using WT T7 RNA polymerase or T7 RNA polymerase variants, S43A* (with C-terminal G) or G47A* (with C-terminal G). FIG. 9A shows results from a LCMS analysis. FIG. 9B shows quantification of the 3' end population distribution from FIG. 9A.

FIG. 10 shows a schematic of the C-helix and C-linker loop structures changing conformation into C-helix and C-loop helix structures as the T7 RNA polymerase transitions from the initiation complex to the elongation complex. FIG. 10 was generated from PDB crystal structures 1MSW and 1QLN and rendered in Molecular Operating Environment [Chemical Computing Group ULC].

FIG. 11A shows a schematic of using a DNA splint in the presence of ligase to ligate a “leftmer” RNA transcript produced by IVT to a “rightmer” fluorescently-labeled polyA signal. Even a single nucleotide overhang at the ligation site abolishes ligation efficiently. FIG. 11B is a PAGE-D gel showing ligation efficiency for rightmers produced with WT T7 RNA polymerase and T7 RNA polymerase variants, G47A* (with C-terminal G) or S43A* (with C-terminal G).

FIG. 12 shows a radioactive gel using ³²P-GTP to label abortive transcripts or ³²P-CTP to label reverse complements. The IVTs were performed using a short model transcript. This data demonstrates that T7 RNA polymerase variants S43A* (with C-terminal G) and G47A* (with C-terminal G) reduce reverse complement formation.

FIG. 13 is a schematic of a conventional method of producing a capped mRNA via in vitro transcription.

FIGs. 14A-14E are graphs showing results of a co-transcriptional capping assay using either wild-type T7 RNA polymerase or a T7 RNA polymerase variant of the present disclosure (G47A*). FIG. 14A shows examples of dinucleotide- (Vaccinia cap1) or trinucleotide- (GAG or GmAG) caps. Trinucleotide caps were used in the co-transcriptional capping assay. FIG. 14B shows the mRNA yields of the co-transcriptional capping assay. FIG. 14C shows that the mRNAs produced in the co-transcriptional capping assay were of high integrity. FIG. 14D shows the capped mRNAs produced by WT T7 RNA polymerase induced cytokine production in BJ fibroblasts, while the capped mRNAs produced by T7 RNA polymerase variant G47A* (with C-terminal G), surprisingly, did not induce cytokine production. FIG. 14E shows the capped mRNAs produced by WT T7 RNA polymerase did not express the encoded protein (hEPO) in BJ fibroblasts, while the capped mRNAs produced by T7 RNA polymerase variant G47A* (with C-terminal G) lead to hEPO expression level comparable to the control. The control is a mRNA produced by WT T7 RNA polymerase and capped with Vaccinia cap1.

FIG. 15 shows the results of LC-mass spectrometry analysis of the mRNA produced in the co-transcriptional capping assay. The result shows unexpectedly that T7 RNA polymerase variant G47A* (with C-terminal G) produced cleaner mRNA than WT T7 RNA polymerase. The trinucleotide incorporation rates appeared equivalent for both enzymes.

FIGs. 16A and 16B are graphs comparing the cytokine response (FIG. 16A) or expression (FIG. 16B) of mRNAs produced and capped with GAG trinucleotide in the co-transcriptional capping assay or capped in a standard capping assay with Vaccinia cap1.

FIG. 17 is a schematic showing the co-transcriptionally capping assay described herein using T7 RNA polymerase variant G47A* and a trinucleotide cap $m^7GpppA_{2'}Ome$ pG.

FIGs. 18A and 18B are liquid chromatography–mass spectrometry (LCMS) results showing the comparison of capping efficiency for mRNAs that start with 5'ATP or 5'GTP. FIG. 18A shows the analysis of intact mRNAs. FIG. 18B shows the analysis of the 5' ends of the mRNAs cleaved by RNase H.

FIGs. 19A-19C are graphs showing the analysis of mRNA chemically modified with 1-methyl-pseudouridine produced from PCR fragment templates for three model constructs (hEPO, luciferase, and eGFP). FIG. 19A shows that the mRNAs have the same sequence, as analyzed by RNase T1 fingerprinting assay. FIG. 19B shows that the mRNAs have high degree of integrity. FIG. 19C shows that the mRNAs did not induce cytokine response in BJ fibroblasts.

FIGs. 20A-20F are graphs showing the analysis of mRNAs chemically modified with 1-methyl-pseudouridine produced from plasmid templates for three model constructs (hEPO, Luc, and eGFP). The NTP consumption in the co-transcriptional capping assays using hEPO plasmids (FIG. 20A) or eGFP plasmids (FIG. 20B) as templates are shown. FIG. 20C shows that the mRNA products are of high integrity. FIG. 20D shows the cytokine response of the mRNA products. FIG. 20E shows the expression of mRNA encoding hEPO in BJ fibroblasts. FIG. 20F shows the expression of mRNA encoding luciferase (Luc) in BJ fibroblasts. FIG. 20G shows the expression of mRNA encoding eGFP in BJ fibroblasts.

FIG. 21 is a graph showing that greater than 85% (e.g., ~90%) of the mRNA transcripts produced using T7 RNAP variant G47A* comprising a C-terminal glycine (G) have a hydroxyl group at the 3' end.

FIG. 22 shows unmodified hEPO mRNA production from IVT reactions containing WT or G47A T7 RNAP variants, some of which have additional amino acids at the C-terminus (e.g., one or two glycines, or two alanines).

FIG. 23 shows 1-methyl-pseudouridine-modified hEPO mRNA production from IVT reactions containing WT, G47A, or S43A/G47A T7 RNAP variants as indicated, some of which have additional amino acids at the C-terminus (e.g., one or two glycines or two alanines).

FIGs. 24A-24D show results from a RNase T1 tail digest of hEPO mRNA transcript produced using WT, G47A, or S43A/G47A T7 RNAP variants as indicated, some of which have additional amino acids at the C-terminus (e.g., one or two glycines or two alanines). FIG. 24A shows results from a LCMS analysis of hEPO mRNA produced from IVT reactions using WT T7 RNAP with equimolar concentrations of NTPs (WT EQ), WT T7 RNAP with excess of GTP and ATP (WT alpha), or G47A* T7 RNAP with an additional C-terminal glycine with equimolar concentrations of NTPs (G47A* EQ). FIG. 24B shows quantification of the 3' end population distribution from FIG. 24A. FIG. 24C shows the percentage of clean 3' end populations produced in IVT reactions using unmodified mRNA. FIG. 24D shows the percentage of clean 3' end populations produced in IVT reactions using 1-methyl-pseudouridine-modified mRNA.

FIGs. 25A-25B are graphs comparing the cytokine response of BJ fibroblasts to unmodified (FIG. 25A) and 1-methyl-pseudouridine-modified (FIG. 25B) mRNAs produced from IVT reactions containing WT, G47A, or S43A/G47A T7 RNAP variants as indicated, some of which have substitutions at the C-terminus.

FIGs. 26A-26B show radiolabeled mRNAs produced from IVT reactions containing WT or G47A T7 RNAP variants as indicated, some of which have additional amino acids at the C-terminus. The IVT mRNA products are separated by size on a polyacrylamide denaturing gel. In FIG. 26A, either unmodified or 1-methyl-pseudouridine-modified mRNA is radiolabeled with ^{32}P -CTP. In FIG. 26B, the reverse complement mRNA is radiolabeled with ^{32}P -CTP.

FIGs. 27A and 27B show graphs of data demonstrating that trinuc-capped G47A* mRNAs encoding firefly luciferase (ffLuc) induce IP-10 serum cytokine levels in vivo at Day 1 (FIG. 27A) and Day 29 (FIG. 27B) that are similar to mRNA controls.

FIGs. 28A and 28B show graphs of data demonstrating that trinuc-capped G47A* mRNAs encoding ffLuc induce baseline cytokine levels in vitro in BJ fibroblasts (BJFs) (FIG. 28A) and in monocyte-derived macrophages (MDMs) (FIG. 28B) that are similar to mRNA controls.

FIG. 29 shows a graph of data demonstrating that trinuc-capped G47A* mRNAs encoding ffLuc maintain high expression in vivo after six (6) weekly doses.

FIG. 30 shows a graph of data demonstrating that trinuc-capped G47A* mRNAs encoding ffLuc induce low anti-PEG IgM levels after six (6) weekly doses.

FIG. 31 shows a graph of data demonstrating that trinuc-capped G47A* mRNAs encoding ffLuc exhibit low B cell activation, similar to mRNA controls.

FIGs. 32A and 32B show graphs of data demonstrating that trinuc-capped G47A* mRNAs encoding human erythropoietin (hEPO) induce IP-10 serum cytokine levels in vivo at Day 1 (FIG. 32A) and Day 22 (FIG. 32B) that are similar to mRNA controls.

FIGs. 33A and 33B show graphs of data demonstrating that trinuc-capped G47A* mRNAs encoding hEPO induce baseline cytokine levels in vitro in BJ fibroblasts (BJFs) (FIG. 33A) and in monocyte-derived macrophages (MDMs) (FIG. 33B) that are similar to mRNA controls.

FIG. 34 shows a graph of data demonstrating that trinuc-capped G47A* mRNAs encoding hEPO maintain high expression in vivo after six (6) weekly doses.

FIG. 35 shows a graph of data demonstrating that trinuc-capped G47A* mRNAs encoding hEPO induce low anti-PEG IgM levels after six (6) weekly doses.

FIG. 36 shows a graph of data demonstrating that trinuc-capped G47A* mRNAs encoding hEPO exhibit low B cell activation, similar to mRNA controls.

FIGs. 37A and 37B show that the G47A* T7 RNA polymerase variant does not affect indel frequency (FIG. 37A) or point mutation frequency (FIG. 37B).

DETAILED DESCRIPTION

The present disclosure provides RNA polymerase (RNAP) variants that increase transcription efficiency and 3' homogeneity while reducing the number of run-on transcripts, and/or double-stranded contaminants produced during an in vitro transcription (IVT) reaction, for example. These RNAP variants, which include, in some embodiments, a single amino acid substitution, facilitate the RNAP conformational transition from initiation complex to elongation complex, thereby reducing many of the problems associated with the transcription initiation phase.

Provided herein are unexpected experimental results demonstrating that modification(s) of the N-terminal C-helix and/or C-linker structure(s) of DNA-dependent RNA polymerase (e.g., T7 RNA polymerase) drives the conformation equilibrium of the polymerase toward the elongation complex, facilitating release of the DNA template promoter and commencement of a highly processive elongation phase. Surprisingly, use of the polymerase variants provided herein in IVT reactions reduces 3' heterogeneity among the transcripts produced, and also reduces (or eliminates) the production of double-stranded contaminants. Further, results show that the purity and expression levels of transcripts produced using the polymerase variants are comparable to those of transcripts produced using wild-type polymerase.

Thus, use of the RNA polymerase variants of the present disclosure reduces many of the problems associated with IVT reaction products. Although wild-type T7 RNA polymerase (RNAP) is used commonly in both industry and academia, several of its activities significantly compromise the purity of resulting RNA transcripts. In particular, use of this wild-type T7 RNAP enzyme results in the non-templated addition of nucleotides to the 3'-end of RNA transcripts. For example, wild-type T7 RNAP installs at least 1, and often 2 or more, non-templated nucleotides at the 3'-end with seemingly little preference for nucleobase identity. Surprisingly, the T7 RNAP variants as provided herein reduce the occurrence of 3' heterogeneity. In some embodiments, less than 30% of the RNA transcripts produced by IVT using a RNA polymerase variant of the present disclosure exhibit 3' heterogeneity. In some embodiments, less than 20% (e.g., less than 15%, 10%, or 5%) of the RNA transcripts produced by IVT using a RNA polymerase variant of the present disclosure exhibit 3' heterogeneity. In some embodiments, 1-20%, 1-15%, 1-10%, or 1-5% of the RNA transcripts produced by IVT exhibit 3' heterogeneity. In some embodiments, less than 1% of the RNA transcripts produced by IVT using a RNA polymerase variant of the present disclosure exhibit 3' heterogeneity.

Thus, use of the RNA polymerase variants of the present disclosure, for example, in an IVT reaction, results in the production of 'cleaner' RNA transcript (a population of RNA transcripts with reduced heterogeneity/increased homogeneity at the 3' end). In some embodiments, at least 70% of the RNA transcripts produced using the RNA polymerase variants of the present disclosure exhibit 3' homogeneity. In some embodiments, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% of the RNA transcripts produced using the RNA polymerase variants of the present disclosure exhibit 3' homogeneity. In some embodiments, at least 90% of the RNA transcripts produced using the RNA polymerase variants of the present disclosure exhibit 3' homogeneity. In some embodiments, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% of the RNA transcripts produced using the RNA polymerase variants of the present disclosure exhibit 3' homogeneity.

Also, surprising herein were results demonstrating that the RNA transcripts produced using the amino acid substitution (e.g., S43A and/or G47A of SEQ ID NO:1) or C-terminal additions (e.g., WT, S43A*, and/or G47A* with one or more amino acid additions) T7 RNA polymerase variants, when delivered to cells, optionally in unpurified form (e.g., not purified by reverse phase chromatography), stimulates a cytokine response that is lower relative to RNA produced using wild-type RNA polymerase. In some embodiments, there is no detectable cytokine response from cells that have received an RNA transcript produced using a T7 RNA variant polymerase of the present disclosure. In some embodiments, RNA produced using a RNA polymerase variant (e.g., T7 RNAP S43A* variant and/or T7 RNAP G47A* variant) stimulates a cytokine response that is at least 50% (e.g., 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 100%) lower relative to RNA produced using wild-type RNA polymerase. In some embodiments, RNA produced using a RNA polymerase variant (e.g., T7 RNAP S43A* variant and/or T7 RNAP G47A* variant) stimulates a cytokine response that is 50-60%, 50-70%, 50-80%, 50-90%, 50-100%, 60-70%, 60-80%, 60-90%, 60-100%, 70-80%, 70-90%, 70-100%, 80-90%, 80-100% or 90-100% lower relative to RNA produced using wild-type RNA polymerase. In some embodiments, RNA produced using a RNA polymerase variant (e.g., T7 RNAP S43A* variant and/or T7 RNAP G47A* variant) stimulates a cytokine response that is at least 2-fold, 3-fold, 4-fold or 5-fold lower relative to RNA produced using wild-type RNA polymerase. In some embodiments, the cells used to test a cytokine response are human fibroblast cells (e.g., BJ (ATCC® CRL-2522™) cells). In some embodiments, the cells used to test a cytokine response are monocyte-derived macrophages (MDMs).

In some embodiments, the concentration of dsRNA transcript produced using a RNA polymerase variant is at least 50% (e.g., 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 100%) lower relative to dsRNA transcript produced using wild-type polymerase. In some embodiments, the concentration of dsRNA transcript produced is 50-60%, 50-70%, 50-80%, 50-90%, 50-100%, 60-70%, 60-80%, 60-90%, 60-100%, 70-80%, 70-90%, 70-100%, 80-90%, 80-100% or 90-100% lower relative to dsRNA transcript produced using wild-type polymerase. In some embodiments, the concentration of dsRNA transcript produced is at least 2-fold, 3-fold, 4-fold or 5-fold lower relative to dsRNA transcript produced using wild-type polymerase.

Use of the RNA polymerase variants of the present disclosure also resulted unexpectedly in the production of double stranded contaminants and fewer run-on transcripts.

In some embodiments, less than 50% (e.g., less than 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 5% or 1%) of the RNA transcript produced, e.g., by IVT, using a RNA polymerase variant is double-stranded contaminant. In some embodiments, 1-50%, 1-40%, 1-30%, 1-20%, 1-10%, 1-5%, 5-50%, 5-40%, 5-30%, 5-20%, 5-10%, 10-50%, 10-40%, 10-30% or 10-20% of the RNA transcript produced is double-stranded contaminant.

In some embodiments, less than 50% (e.g., less than 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 5% or 1%) of the RNA transcript produced using a RNA polymerase variant is run-on RNA transcript. In some embodiments, 1-50%, 1-40%, 1-30%, 1-20%, 1-10%, 1-5%, 5-50%, 5-40%, 5-30%, 5-20%, 5-10%, 10-50%, 10-40%, 10-30% or 10-20% of the RNA transcript produced is run-on RNA transcript.

In some embodiments, the amount of full-length RNA transcript produced using a RNA polymerase variant is at least 15 times greater than the amount of the DNA template. For example, the amount of full-length RNA transcript produced may be at least 20, 30, 40, 45, 50, 60, 70, 80, 90 or 100 times greater than the amount of the DNA template. In some embodiments, the amount of full-length RNA transcript produced is 15-100, 15-90, 15-80, 15-70, 15-60, 15-50, 15-40, 15-30, 15-20, 20-100, 20-90, 20-80, 20-70, 20-60, 20-50, 20-40 or 20-30 times greater than the amount of the DNA template. In some embodiments, the amount of full-length RNA transcript produced is 2-fold, 3-fold, 4-fold or 5-fold greater than the amount of the DNA template.

In some embodiments, the ratio of double-stranded contaminant:full-length RNA transcript produced using a RNA polymerase variant is less than 1:1. For example, the ratio of double-stranded contaminant:full-length RNA transcript produced may be 0.9:1, 0.8:1, 0.7:1, 0.6:1, 0.5:1, 0.4:1, 0.3:1, 0.2:1 or 0.1:1.

In some embodiments, the RNA transcript produced using a RNA polymerase variant has less than 1 mutation per 100 nucleotides relative to the DNA template. For example, the RNA transcript produced may have less than 1 mutation per 200, 300, 400, 500, 600, 700, 800, 900 or 1000 nucleotides relative to the DNA template.

RNA Polymerases

RNA polymerase (DNA-dependent RNA polymerase) is an enzyme that catalyzes the sequential addition of a ribonucleotide to the 3' end of a growing RNA chain (transcription of RNA in the 5' → 3' direction), with nucleoside triphosphates (NTPs) acting as substrates for the enzyme and with the sequence of nucleotides specified by a DNA template.

Transcription relies on the complementary pairing of bases. The two strands of a double helix separate locally, and one of the separated strands serves as a template (DNA template). RNA polymerase then catalyzes the alignment of free nucleotides on the DNA template by their complementary bases in the template. Thus, a RNA polymerase is considered to have RNA polymerase activity if the polymerase catalyzes the sequential addition of a ribonucleotide to the 3' end of a growing RNA chain.

T7 RNA polymerase (T7 RNAP) is a 99 kDa DNA-dependent RNA polymerase encoded by the genome of bacteriophage T7 and is highly specific for T7 phage promoters. Structural studies of T7 RNAP have shown that the conformation of the N-terminal domain changes substantially between the initiation phase and elongation phase of transcription. The N-terminal domain comprises a C-helix subdomain and the promoter binding domain, which includes two segments separated by subdomain H. The promoter binding domain and the bound promoter rotate by approximately 45 degrees upon synthesis of an 8-nt RNA transcript, allowing the promoter contacts to be maintained while the active site is expanded to accommodate a growing heteroduplex. The C-helix subdomain moves modestly toward its elongation conformation, whereas subdomain H remains in its initiation- rather than its elongation-phase location, more than 70 angstroms away. Comparison of the structures of the T7 RNAP initiation and elongation complexes reveal extensive conformational changes within the N-terminal 267 residues (N-terminal domain) and little change in the rest of the RNAP. A rigid body rotation of the promoter binding domain as well as the refolding of the N-terminal C-helix (residues 28-71) and H (residues 151-190) subdomains are responsible for abolishing the promoter binding site, enlarging the active site and creating an exit tunnel for the RNA transcript. The structural changes within the N-terminal domain account for the

increased stability and the processivity of the elongation complex (see, e.g., Durniak, K.J. et al., Science 322(5901): 553–557, 2008, incorporated herein by reference).

Provided herein, in some aspects, are RNA polymerase variants (e.g., T7 RNAP variants) that facilitate the conformational change from the RNAP initiation complex to the RNAP elongation complex. A RNA polymerase variant is an enzyme having RNA polymerase activity and at least one substitution relative to the counterpart wild-type RNA polymerase. As indicated above, a RNA polymerase is considered to have RNA polymerase activity if the polymerase catalyzes the sequential addition of a ribonucleotide to the 3' end of a growing RNA chain. For example, an enzyme that includes the amino acid sequence SEQ ID NO:1 with an amino acid substitution at position S43 (e.g., S43A) or G47 (e.g., G47A) and maintains RNA polymerase activity is considered a T7 RNAP variant of wild-type T7 RNAP (SEQ ID NO:1).

In some embodiments, a RNA polymerase variant comprises at least one amino acid modification, relative to wild-type RNA polymerase, that causes at least one three-dimensional loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex. Thus, in some embodiments, at least one amino acid modification has a high-helix propensity, relative to wild-type amino acid.

Examples of loop structures include but are not limited to amino acid (aa) 42-47 in the C-helix structure (e.g., aa 28-71 of SEQ ID NO:1) of the T7 RNA polymerase initiation complex (IC) conformation and aa 257-262 in the C-linker structure (e.g., aa 258-266 of SEQ ID NO:1) of the IC.

Also provided herein are RNA polymerase variants (e.g., T7 RNAP variants) that include at least one additional amino acid at the C terminus. The at least one additional amino acid, in some embodiments, is selected from alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine. In some embodiments, the at least one additional amino acid is a polar amino acid. In some embodiments, the at least one additional amino acid is a non-polar amino acid. In some embodiments, the at least one additional amino acid is glycine. In some embodiments, the at least one additional amino acid is alanine. In some embodiments, the at least one additional amino acid is serine.

In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAX_n (SEQ ID NO:294), wherein x is any amino acid and

n is any integer, e.g., between 1 and 5 (e.g., 1, 2, 3, 4, or 5). In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAG_n (SEQ ID NO:295), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAA_n (SEQ ID NO:296), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAR_n (SEQ ID NO:297), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAN_n (SEQ ID NO:298), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAD_n (SEQ ID NO:299), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAC_n (SEQ ID NO:300), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAE_n (SEQ ID NO:301), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAQ_n (SEQ ID NO:302), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAH_n (SEQ ID NO:303), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAI_n (SEQ ID NO:304), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAL_n (SEQ ID NO:305), n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAK_n (SEQ ID NO:306), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAM_n (SEQ ID NO:307), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAF_n (SEQ ID NO:308), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAP_n (SEQ ID NO:309), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAS_n (SEQ ID NO:310), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following

consensus sequence: FAFAT_n (SEQ ID NO:311), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAW_n (SEQ ID NO:312), wherein n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAY_n (SEQ ID NO:313), n is any integer, e.g., between 1 and 5. In some embodiments, the C terminus of an RNA polymerase comprises the following consensus sequence: FAFAV_n (SEQ ID NO:314), wherein n is any integer, e.g., between 1 and 5.

In some embodiments, the C-terminal motif (FAFA or FAFAX_n) comprises an amino acid substitution at one or more of positions 880, 881, 882, or 883, relative to wild-type T7 RNAP (e.g., SEQ ID NO:1). Thus, the present disclosure comprises various C-terminal F⁸⁸⁰A⁸⁸¹F⁸⁸²A⁸⁸³ motifs, wherein the amino acid at one or more of positions 880, 881, 882, or 883 relative to wild-type T7 RNAP (e.g., SEQ ID NO:1) is modified to include at least one amino acid substitution, with or without an additional C-terminal amino acid (X_n).

Amino Acid Substitutions

RNA polymerase variants of the present disclosure include at least one amino acid substitution, relative to the WT RNA polymerase. For example, with reference to WT T7 RNA polymerase having an amino acid sequence of SEQ ID NO:1, the serine at position 43 is considered a “wild-type amino acid,” whereas a substitution of the serine for alanine at position 43 is considered an “amino acid substitution” that has a high-helix propensity.

The average globular protein contains 30% α -helix, the most common type of secondary structure. Some amino acids occur more frequently in α -helices than others; this tendency is known as helix propensity. See, e.g., Pace, N.C. and Scholtz, J.M. *Biophysical Journal*, 75:422-427 (1998), incorporated herein by reference. In some embodiments, at least one amino acid substitution has a high-helix propensity, relative to wild-type amino acid. In general, high-helix propensity amino acid substitutions are selected to thermodynamically bias the population of polymerase conformers toward the elongation complex. The relative $\Delta\Delta G$ s (free energies) is calculated for substitutions in the IC and the elongation complex (EC) structures using publicly-available software (e.g., University of Washington’s Rosetta and Schrödinger’s Maestro). Substitutions are then selected based on the calculations and on additional knowledge, including for example, amino acid helix propensity and/or polypeptide backbone phi-psi compatibility.

In some embodiments, the RNA polymerase variant is a T7 RNA polymerase variant comprising at least one (one or more) amino acid substitution relative to WT T7 RNA polymerase (e.g., WT T7 RNA polymerase having an amino acid sequence of SEQ ID NO:1). In some embodiments, an amino acid substitution is a high-helix propensity amino acid substitution. Examples of high-helix propensity amino acids include alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate.

In some embodiments, the RNA polymerase variant is a T7 RNA polymerase variant comprising at least one additional amino acid at the C terminus (e.g., SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:294, SEQ ID NO:295, or SEQ ID NO:296, or an amino acid sequence that shares 90%-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:294, SEQ ID NO:295, or SEQ ID NO:296). In some embodiments, the RNA polymerase variant is a T7 RNA polymerase variant comprising at least one additional amino acid at the C terminus and an amino acid substitution relative to WT T7 RNA polymerase. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution.

Provided herein are two approaches that may be used to identify high or higher helix propensity amino acid substitutions that thermodynamically bias/favor the conformational equilibrium toward the EC. In both approaches, care is taken to avoid making amino acid substitutions to any positions that are directly involved in promoter binding or catalysis (based on review of the structures, literature searches). In Method A, specialized software is used to compute the change in free energy of protein folding due to mutation ($\Delta\Delta G_{mut}$) in the EC and IC. All possible amino acid substitutions at the desired sequence positions are evaluated by this method. Desired amino acid substitutions (i.e., those that favor the EC over the IC) are those for which the $\Delta\Delta G_{mut}$ difference ($\Delta\Delta G_{mut}^{EC} - \Delta\Delta G_{mut}^{IC}$) is negative. In Method B, the IC and EC secondary structure annotations (e.g., as determined from 3D models of IC and EC tertiary structures using the DSSP software (Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Kabsch W, Sander C, Biopolymers. 1983 22 2577-2637)) are inspected to identify regions with a loop in the IC and a helix in the EC. Sequence positions with low helix propensity residues (e.g., Gly) are then substituted with high helix propensity amino acids (e.g., Ala). For T7 RNAP, for example, mutational (e.g., substitutional) overlap from Methods A and B were identified as well as the top variant from Method B.

In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitution at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46 and/or G47) of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitution at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46 and/or G47) of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitution at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46 and/or G47) of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitution at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46 and/or G47) of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 45 of SEQ ID NO:99, 100, 294,

295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is an alanine at position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is an alanine at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is an isoleucine at position 42 of SEQ ID NO:1. In some

embodiments, the amino acid substitution is an isoleucine at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is an isoleucine at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a leucine at position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 42-47 of an amino acid

sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a leucine at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is an arginine at position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is an arginine at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a methionine at position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a methionine at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a lysine at position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a

lysine at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a lysine at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a glutamine at position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a glutamine at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a

glutamine at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 42-47 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a glutamate at position 42 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 43 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 44 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 45 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 46 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 47 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 42-47 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 42-47 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a glutamate at position 42 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 43 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 44 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 45 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 46 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 47 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or

glutamate) substitution at any one of amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitution at any one of amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitution at any one of amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitution at any one of amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix

propensity amino acid substitution at amino acid position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a high-helix propensity amino acid substitution at amino acid position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is an alanine at position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an alanine at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is an alanine at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an alanine at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is an isoleucine at position 257 of SEQ ID NO:1. In some

embodiments, the amino acid substitution is an isoleucine at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an isoleucine at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is an isoleucine at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an isoleucine at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a leucine at position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a leucine at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at any one of amino acid positions 257-262 of an amino

acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a leucine at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a leucine at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is an arginine at position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is an arginine at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is an arginine at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is an arginine at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the

amino acid substitution is an arginine at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a methionine at position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a methionine at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a methionine at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a methionine at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a lysine at position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is

a lysine at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a lysine at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a lysine at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a lysine at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a glutamine at position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamine at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at

least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a glutamine at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamine at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 257-262 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. Thus, in some embodiments, the amino acid substitution is a glutamate at position 257 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 258 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 259 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 260 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 261 of SEQ ID NO:1. In some embodiments, the amino acid substitution is a glutamate at position 262 of SEQ ID NO:1.

In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 257-262 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at any one of amino acid positions 257-262 of an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. Thus, in some embodiments, the amino acid substitution is a glutamate at position 257 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 258 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 259 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 260 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, the amino acid substitution is a glutamate at position 261 of SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments,

the amino acid substitution is a glutamate at position 262 of SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, the T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:1 modified to include an amino acid substitution of a high-helix propensity amino acid at position G47, S43, R257, and/or G259. In some embodiments, the T7 RNA polymerase variant comprises an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1 modified to include an amino acid substitution of a high-helix propensity amino acid at position G47, S43, R257, and/or G259.

In some embodiments, the T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:99, 100, 294, 295, or 296 modified to include an amino acid substitution of a high-helix propensity amino acid at position G47, S43, R257, and/or G259. In some embodiments, the T7 RNA polymerase variant comprises an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296 modified to include an amino acid substitution of a high-helix propensity amino acid at position G47, S43, R257, and/or G259.

In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:2. In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:3. In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:4. In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:5.

In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:107 or 108. In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:109 or 110. In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:111 or 112. In some embodiments, a T7 RNA polymerase variant comprises an amino acid sequence of SEQ ID NO:113 or 114.

Also provided herein are RNA polymerase variants with at least 2 (2 or more) substitutions. In some embodiments, a RNA polymerase comprise at least 2 high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitutions at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46 and/or G47) and/or amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of SEQ ID NO:1 or an amino acid sequence that shares 90-99%,

e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1.

For example, a RNA polymerase variant may comprise an amino acid substitution (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at positions E42 and S43, E42 and Y44, E42 and E45, E42 and M46, E42 and G47, S43 and Y44, S43 and E45, S43 and M46, S43 and G47, Y44 and E45, Y44 and M46, Y44 and G47, E45 and M46, E45 and G47, or M46 and G47 of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1.

In some embodiments, a RNA polymerase variant comprises an amino acid substitution (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at positions R257 and A258, R257 and G259, R257 and A260, R257 and L261, R257 and A262, A258 and G259, A258 and A260, A258 and L261, A258 and A262, G259 and A260, G259 and L261, G259 and A262, A260 and L261, A260 and A262, or L261 and A262 of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1.

In some embodiments, a RNA polymerase variant comprises an amino acid substitution (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at positions E42 and R257, E42 and A258, E42 and G259, E42 and A260, E42 and L261, E42 and A262, S43 and R257, S43 and A258, S43 and G259, S43 and A260, S43 and L261, S43 and A262, Y44 and R257, Y44 and A258, Y44 and G259, Y44 and A260, Y44 and L261, Y44 and A262, E45 and R257, E45 and A258, E45 and G259, E45 and A260, E45 and L261, E45 and A262, M46 and R257, M46 and A258, M46 and G259, M46 and A260, M46 and L261, M46 and A262, G47 and R257, G47 and A258, G47 and G259, G47 and A260, G47 and L261, or G47 and A262 of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1.

In some embodiments, a RNA polymerase comprise amino acid substitutions S43A and G47A of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. In some embodiments, a RNA polymerase comprise amino acid substitutions S43A and R257A of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ

ID NO:1. In some embodiments, a RNA polymerase comprise amino acid substitutions S43A and G259A of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. In some embodiments, a RNA polymerase comprise amino acid substitutions G47A and R257A of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. In some embodiments, a RNA polymerase comprise amino acid substitutions G47A and R257A of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. In some embodiments, a RNA polymerase comprise amino acid substitutions R257A and G259A of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1. In some embodiments, a RNA polymerase comprise amino acid substitutions G47A and G259A of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1.

In some embodiments, a RNA polymerase variant comprises at least 3 (or at least 4 or at least 5) high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitutions at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46 and/or G47) and/or amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1.

In some embodiments, a RNA polymerase variant comprises at least 1 (or at least 2 or at least 3 or at least 4 or at least 5) amino acid substitution(s) in SEQ ID NO:1 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:1, selected from the following: E42R, S43A, S43E, S43L, S43R, E45R, E45L, M46A, G47A, G47E, G47L, G47R, N165W, E167M, E167N, E168I, E168T, E168V, A181F, A181W, G184M, E187F, A255Q, A255K, A255I, A255Y, R257A, R257E, R257L, R257W, G259A, G259E, G259L, G259R, A260W, and A260R.

In some embodiments, a RNA polymerase comprise at least 2 high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitutions at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46

and/or G47) and/or amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296.

For example, a RNA polymerase variant may comprise an amino acid substitution (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at positions E42 and S43, E42 and Y44, E42 and E45, E42 and M46, E42 and G47, S43 and Y44, S43 and E45, S43 and M46, S43 and G47, Y44 and E45, Y44 and M46, Y44 and G47, E45 and M46, E45 and G47, or M46 and G47 of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, a RNA polymerase variant comprises an amino acid substitution (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at positions R257 and A258, R257 and G259, R257 and A260, R257 and L261, R257 and A262, A258 and G259, A258 and A260, A258 and L261, A258 and A262, G259 and A260, G259 and L261, G259 and A262, A260 and L261, A260 and A262, or L261 and A262 of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, a RNA polymerase variant comprises an amino acid substitution (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) at positions E42 and R257, E42 and A258, E42 and G259, E42 and A260, E42 and L261, E42 and A262, S43 and R257, S43 and A258, S43 and G259, S43 and A260, S43 and L261, S43 and A262, Y44 and R257, Y44 and A258, Y44 and G259, Y44 and A260, Y44 and L261, Y44 and A262, E45 and R257, E45 and A258, E45 and G259, E45 and A260, E45 and L261, E45 and A262, M46 and R257, M46 and A258, M46 and G259, M46 and A260, M46 and L261, M46 and A262, G47 and R257, G47 and A258, G47 and G259, G47 and A260, G47 and L261, or G47 and A262 of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, a RNA polymerase comprises amino acid substitutions S43A and G47A of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, a RNA

polymerase comprises amino acid substitutions S43A and R257A of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, a RNA polymerase comprises amino acid substitutions S43A and G259A of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, a RNA polymerase comprises amino acid substitutions G47A and R257A of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, a RNA polymerase comprises amino acid substitutions G47A and R257A of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, a RNA polymerase comprises amino acid substitutions R257A and G259A of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296. In some embodiments, a RNA polymerase comprises amino acid substitutions G47A and G259A of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, a RNA polymerase variant comprises at least 3 (or at least 4 or at least 5) high-helix propensity amino acid (e.g., alanine, isoleucine, leucine, arginine, methionine, lysine, glutamine, and/or glutamate) substitutions at any one of amino acid positions 42-47 (E42, S43, Y44, E45, M46 and/or G47) and/or amino acid positions 257-262 (R257, A258, G259, A260, L261 and/or A262) of SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296.

In some embodiments, a RNA polymerase variant comprises at least 1 (or at least 2 or at least 3 or at least 4 or at least 5) amino acid substitution(s) in SEQ ID NO:99, 100, 294, 295, or 296 or an amino acid sequence that shares 90-99%, e.g., at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity with SEQ ID NO:99, 100, 294, 295, or 296, selected from the following: E42R, S43A, S43E, S43L, S43R, E45R, E45L,

M46A, G47A, G47E, G47L, G47R, N165W, E167M, E167N, E168I, E168T, E168V, A181F, A181W, G184M, E187F, R257A, R257E, R257L, R257W, G259A, G259E, G259L, G259R, A260W, and A260R.

In some embodiments, a RNA polymerase variant comprises at least 1 (or at least 2 or at least 3 or at least 4 or at least 5) amino acid substitution(s) selected from the following: E42R, S43A, S43E, S43L, S43R, E45R, E45L, M46A, G47A, G47E, G47L, G47R, N165W, E167M, E167N, E168I, E168T, E168V, A181F, A181W, G184M, E187F, R257A, R257E, R257L, R257W, G259A, G259E, G259L, G259R, A260W, and A260R, further comprises at least 1 (or at least 2 or at least 3 or at least 4 or at least 5) other amino acid substitution (e.g., an amino acid substitution not provided herein).

Thus, the present disclosure encompasses RNA polymerase variants that comprise at least 1 (or at least 2 or at least 3 or at least 4 or at least 5) amino acid substitution(s) selected from the following: E42R, S43A, S43E, S43L, S43R, E45R, E45L, M46A, G47A, G47E, G47L, G47R, N165W, E167M, E167N, E168I, E168T, E168V, A181F, A181W, G184M, E187F, A255K, A255Q, A255Y, A255I, R257A, R257E, R257L, R257W, G259A, G259E, G259L, G259R, A260W, and A260R and are 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% identical to the amino acid sequence of SEQ ID NO: 1.

The term “identity” refers to a relationship between the sequences of two or more polypeptides (e.g. enzymes) or polynucleotides (nucleic acids), as determined by comparing the sequences. Identity also refers to the degree of sequence relatedness between or among sequences as determined by the number of matches between strings of two or more amino acid residues or nucleic acid residues. Identity measures the percent of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program (e.g., “algorithms”). Identity of related proteins or nucleic acids can be readily calculated by known methods. “Percent (%) identity” as it applies to polypeptide or polynucleotide sequences is defined as the percentage of residues (amino acid residues or nucleic acid residues) in the candidate amino acid or nucleic acid sequence that are identical with the residues in the amino acid sequence or nucleic acid sequence of a second sequence after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent identity. Methods and computer programs for the alignment are well known in the art. It is understood that identity depends on a calculation of percent identity but may differ in value due to gaps and penalties introduced in the calculation. Generally, variants of a particular polynucleotide or polypeptide (e.g.,

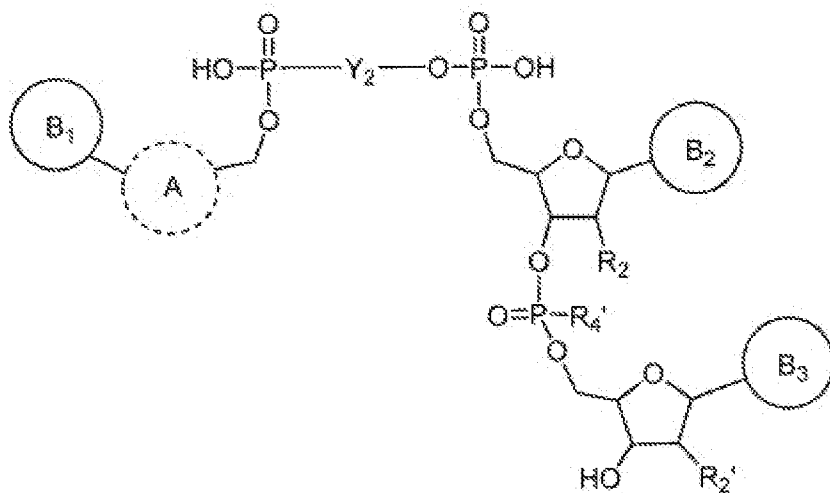
antigen) have at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% but less than 100% sequence identity to that particular reference polynucleotide or polypeptide as determined by sequence alignment programs and parameters described herein and known to those skilled in the art. Such tools for alignment include those of the BLAST suite (Stephen F. Altschul, et al (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402). Another popular local alignment technique is based on the Smith-Waterman algorithm (Smith, T.F. & Waterman, M.S. (1981) "Identification of common molecular subsequences." *J. Mol. Biol.* 147:195-197). A general global alignment technique based on dynamic programming is the Needleman–Wunsch algorithm (Needleman, S.B. & Wunsch, C.D. (1970) "A general method applicable to the search for similarities in the amino acid sequences of two proteins." *J. Mol. Biol.* 48:443-453). More recently a Fast Optimal Global Sequence Alignment Algorithm (FOGSAA) has been developed that purportedly produces global alignment of nucleotide and protein sequences faster than other optimal global alignment methods, including the Needleman–Wunsch algorithm.

Trinucleotide Caps

Also provided herein are co-transcriptional capping methods for ribonucleic acid (RNA) synthesis. That is, RNA is produced in a "one-pot" reaction, without the need for a separate capping reaction. Thus, the methods, in some embodiments, comprise reacting a polynucleotide template with a T7 RNA polymerase variant, nucleoside triphosphates, and a cap analog under in vitro transcription reaction conditions to produce RNA transcript.

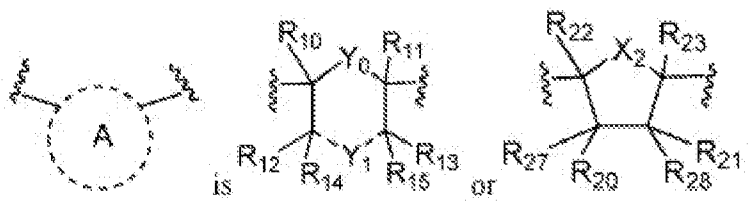
A cap analog may be, for example, a dinucleotide cap, a trinucleotide cap, or a tetranucleotide cap. In some embodiments, a cap analog is a dinucleotide cap. In some embodiments, a cap analog is a trinucleotide cap. In some embodiments, a cap analog is a tetranucleotide cap.

A trinucleotide cap, in some embodiments, comprises a compound of formula (I)



(I), or a stereoisomer,

tautomer or salt thereof, wherein



ring B₁ is a modified or unmodified Guanine;

ring B₂ and ring B₃ each independently is a nucleobase or a modified nucleobase;

X₂ is O, S(O)_p, NR₂₄ or CR₂₅R₂₆ in which p is 0, 1, or 2;

Y₀ is O or CR₆R₇;

Y₁ is O, S(O)_n, CR₆R₇, or NR₈, in which n is 0, 1, or 2;

each --- is a single bond or absent, wherein when each --- is a single bond, Y_i is O, S(O)_n, CR₆R₇, or NR₈; and when each --- is absent, Y₁ is void;

Y₂ is (OP(O)R₄)_m in which m is 0, 1, or 2, or -O-(CR₄₀R₄₁)_u-Q₀-(CR₄₂R₄₃)_v-, in which Q₀ is a bond, O, S(O)_r, NR₄₄, or CR₄₅R₄₆, r is 0, 1, or 2, and each of u and v independently is 1, 2, 3 or 4;

each R₂ and R₂' independently is halo, LNA, or OR₃;

each R₃ independently is H, C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl and R₃, when being C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl, is optionally substituted with one or more of halo, OH and C₁-C₆ alkoxy that is optionally substituted with one or more OH or OC(O)-C₁-C₆ alkyl;

each R₄ and R₄' independently is H, halo, C₁-C₆ alkyl, OH, SH, SeH, or BH₃⁻;

each of R₆, R₇, and R₈, independently, is -Q₁-T₁, in which Q₁ is a bond or C₁-C₃ alkyl linker optionally substituted with one or more of halo, cyano, OH and C₁-C₆ alkoxy, and T₁ is

H, halo, OH, COOH, cyano, or R_{s1} , in which R_{s1} is C₁-C₃ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₁-C₆ alkoxy, C(O)O-C₁-C₆ alkyl, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, $NR_{31}R_{32}$, $(NR_{31}R_{32}R_{33})^+$, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl, and R_{s1} is optionally substituted with one or more substituents selected from the group consisting of halo, OH, oxo, C₁-C₆ alkyl, COOH, C(O)O-C₁-C₆ alkyl, cyano, C₁-C₆ alkoxy, $NR_{31}R_{32}$, $(NR_{31}R_{32}R_{33})^+$, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, 4 to 12-membered heterocycloalkyl, and 5- or 6-membered heteroaryl;

each of R_{10} , R_{11} , R_{12} , R_{13} , R_{14} , and R_{15} , independently, is -Q₂-T₂, in which Q₂ is a bond or C₁-C₃ alkyl linker optionally substituted with one or more of halo, cyano, OH and C₁-C₆ alkoxy, and T₂ is H, halo, OH, NH₂, cyano, NO₂, N₃, R_{s2} , or OR_{s2} , in which R_{s2} is C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, $NHC(O)$ -C₁-C₆ alkyl, $NR_{31}R_{32}$, $(NR_{31}R_{32}R_{33})^+$, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl, and R_{s2} is optionally substituted with one or more substituents selected from the group consisting of halo, OH, oxo, C₁-C₆ alkyl, COOH, C(O)O-C₁-C₆ alkyl, cyano, C₁-C₆ alkoxy, $NR_{31}R_{32}$, $(NR_{31}R_{32}R_{33})^+$, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, 4 to 12-membered heterocycloalkyl, and 5- or 6-membered heteroaryl; or alternatively R_{12} together with R_{14} is oxo, or R_{13} together with R_{15} is oxo,

each of R_{20} , R_{21} , R_{22} , and R_{23} independently is -Q₃-T₃, in which Q₃ is a bond or C₁-C₃ alkyl linker optionally substituted with one or more of halo, cyano, OH and C₁-C₆ alkoxy, and T₃ is H, halo, OH, NH₂, cyano, NO₂, N₃, R_{s3} , or OR_{s3} , in which R_{s3} is C₁-C₆ alkyl, C₂-C₆ alkenyl, C₂-C₆ alkynyl, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, $NHC(O)$ -C₁-C₆ alkyl, mono-C₁-C₆ alkylamino, di-C₁-C₆ alkylamino, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl, and R_{s3} is optionally substituted with one or more substituents selected from the group consisting of halo, OH, oxo, C₁-C₆ alkyl, COOH, C(O)O-C₁-C₆ alkyl, cyano, C₁-C₆ alkoxy, amino, mono-C₁-C₆ alkylamino, di-C₁-C₆ alkylamino, C₃-C₈ cycloalkyl, C₆-C₁₀ aryl, 4 to 12-membered heterocycloalkyl, and 5- or 6-membered heteroaryl;

each of R_{24} , R_{25} , and R_{26} independently is H or C₁-C₆ alkyl;

each of R_{27} and R_{28} independently is H or OR_{29} ; or R_{27} and R_{28} together form O- R_{30} -O; each R_{29} independently is H, C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl and R_{29} , when being C₁-C₆ alkyl, C₂-C₆ alkenyl, or C₂-C₆ alkynyl, is optionally substituted with one or more of halo, OH and C₁-C₆ alkoxy that is optionally substituted with one or more OH or $OC(O)$ -C₁-C₆ alkyl;

R_{30} is C₁-C₆ alkylene optionally substituted with one or more of halo, OH and C₁-C₆ alkoxy;

each of R_{31} , R_{32} , and R_{33} , independently is H, C_1 - C_6 alkyl, C_3 - C_8 cycloalkyl, C_6 - C_{10} aryl, 4 to 12-membered heterocycloalkyl, or 5- or 6-membered heteroaryl;

each of R_{40} , R_{41} , R_{42} , and R_{43} independently is H, halo, OH, cyano, N_3 , $OP(O)R_{47}R_{48}$, or C_1 - C_6 alkyl optionally substituted with one or more $OP(O)R_{47}R_{48}$, or one R_{41} and one R_{43} , together with the carbon atoms to which they are attached and Q_0 , form C_4 - C_{10} cycloalkyl, 4- to 14-membered heterocycloalkyl, C_6 - C_{10} aryl, or 5- to 14-membered heteroaryl, and each of the cycloalkyl, heterocycloalkyl, phenyl, or 5- to 6-membered heteroaryl is optionally substituted with one or more of OH, halo, cyano, N_3 , oxo, $OP(O)R_{47}R_{48}$, C_1 - C_6 alkyl, C_1 - C_6 haloalkyl, $COOH$, $C(O)O$ - C_1 - C_6 alkyl, C_1 - C_6 alkoxy, C_1 - C_6 haloalkoxy, amino, mono- C_1 - C_6 alkylamino, and di- C_1 - C_6 alkylamino;

R_{44} is H, C_1 - C_6 alkyl, or an amine protecting group;

each of R_{45} and R_{46} independently is H, $OP(O)R_{47}R_{48}$, or C_1 - C_6 alkyl optionally substituted with one or more $OP(O)R_{47}R_{48}$, and

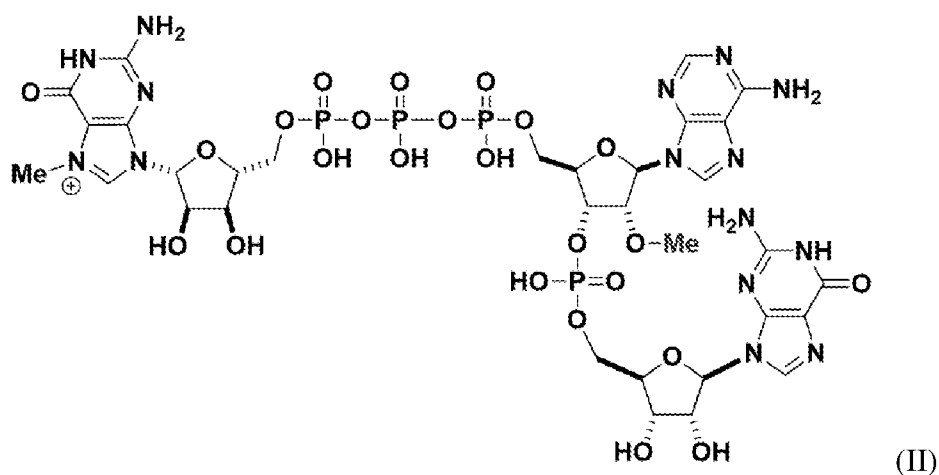
each of R_{47} and R_{48} , independently is H, halo, C_1 - C_6 alkyl, OH, SH, SeH, or BH_3^- .

It should be understood that a cap analog, as provided herein, may include any of the cap analogs described in international publication WO 2017/066797, published on 20 April 2017, incorporated by reference herein in its entirety.

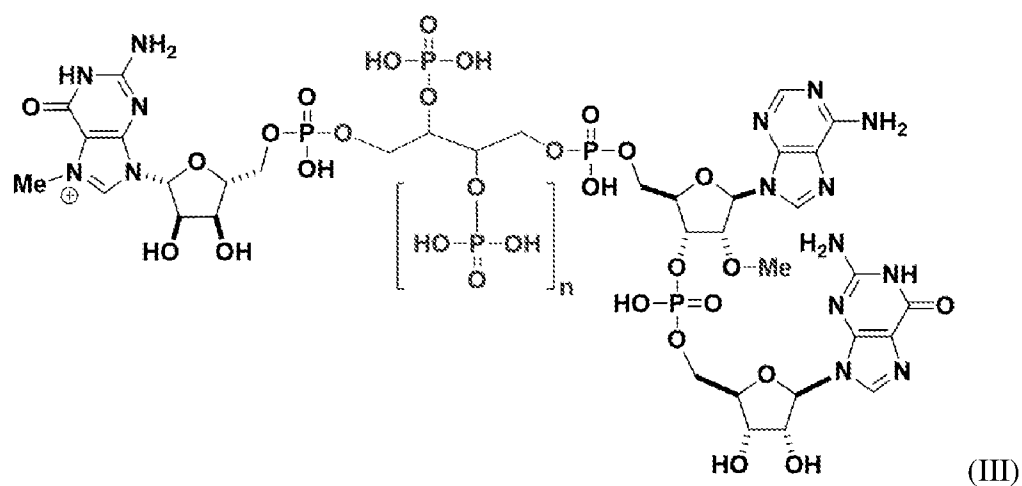
In some embodiments, the B_2 middle position can be a non-ribose molecule, such as arabinose.

In some embodiments R_2 is ethyl-based.

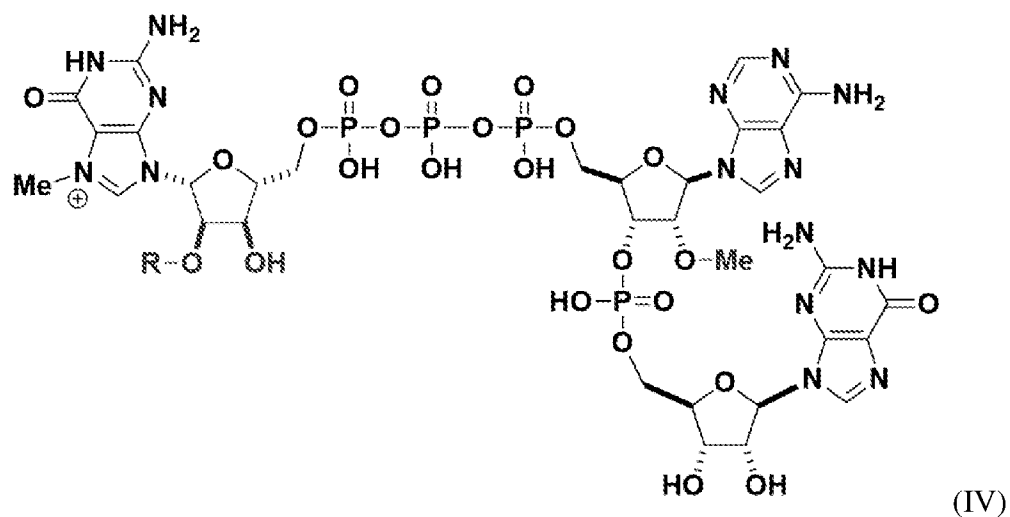
Thus, in some embodiments, a trinucleotide cap comprises the following structure:



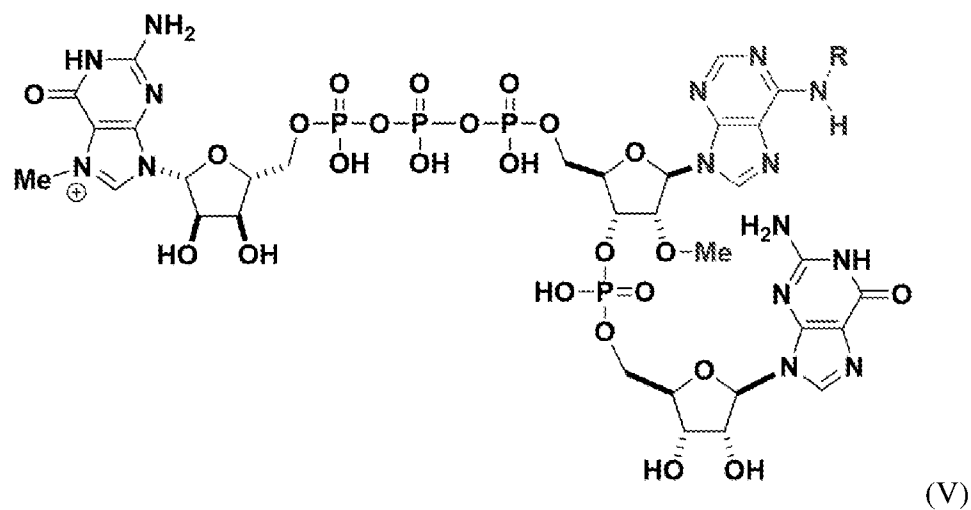
In other embodiments, a trinucleotide cap comprises the following structure:



In yet other embodiments, a trinucleotide cap comprises the following structure:



In still other embodiments, a trinucleotide cap comprises the following structure:



A trinucleotide cap, in some embodiments, comprises a sequence selected from the following sequences: GAA, GAC, GAG, GAU, GCA, GCC, GCG, GCU, GGA, GGC, GGG, GGU, GUA, GUC, GUG, and GUU. In some embodiments, a trinucleotide cap comprises GAA. In some embodiments, a trinucleotide cap comprises GAC. In some embodiments, a trinucleotide cap comprises GAG. In some embodiments, a trinucleotide cap comprises GAU. In some embodiments, a trinucleotide cap comprises GCA. In some embodiments, a trinucleotide cap comprises GCC. In some embodiments, a trinucleotide cap comprises GCG. In some embodiments, a trinucleotide cap comprises GCU. In some embodiments, a trinucleotide cap comprises GGA. In some embodiments, a trinucleotide cap comprises GGC. In some embodiments, a trinucleotide cap comprises GGG. In some embodiments, a trinucleotide cap comprises GGU. In some embodiments, a trinucleotide cap comprises GUA. In some embodiments, a trinucleotide cap comprises GUC. In some embodiments, a trinucleotide cap comprises GUG. In some embodiments, a trinucleotide cap comprises GUU.

In some embodiments, a trinucleotide cap comprises a sequence selected from the following sequences: m⁷GpppApA, m⁷GpppApC, m⁷GpppApG, m⁷GpppApU, m⁷GpppCpA, m⁷GpppCpC, m⁷GpppCpG, m⁷GpppCpU, m⁷GpppGpA, m⁷GpppGpC, m⁷GpppGpG, m⁷GpppGpU, m⁷GpppUpA, m⁷GpppUpC, m⁷GpppUpG, and m⁷GpppUpU.

In some embodiments, a trinucleotide cap comprises m⁷GpppApA. In some embodiments, a trinucleotide cap comprises m⁷GpppApC. In some embodiments, a trinucleotide cap comprises m⁷GpppApG. In some embodiments, a trinucleotide cap comprises m⁷GpppApU. In some embodiments, a trinucleotide cap comprises m⁷GpppCpA. In some embodiments, a trinucleotide cap comprises m⁷GpppCpC. In some embodiments, a trinucleotide cap comprises m⁷GpppCpG. In some embodiments, a trinucleotide cap comprises m⁷GpppCpU. In some embodiments, a trinucleotide cap comprises m⁷GpppGpA. In some embodiments, a trinucleotide cap comprises m⁷GpppGpC. In some embodiments, a trinucleotide cap comprises m⁷GpppGpG. In some embodiments, a trinucleotide cap comprises m⁷GpppGpU. In some embodiments, a trinucleotide cap comprises m⁷GpppUpA. In some embodiments, a trinucleotide cap comprises m⁷GpppUpC. In some embodiments, a trinucleotide cap comprises m⁷GpppUpG. In some embodiments, a trinucleotide cap comprises m⁷GpppUpU.

A trinucleotide cap, in some embodiments, comprises a sequence selected from the following sequences: m⁷G_{3'OMe}pppApA, m⁷G_{3'OMe}pppApC, m⁷G_{3'OMe}pppApG, m⁷G_{3'OMe}pppApU, m⁷G_{3'OMe}pppCpA, m⁷G_{3'OMe}pppCpC, m⁷G_{3'OMe}pppCpG,

$m^7G_{3'OMe}pppCpU$, $m^7G_{3'OMe}pppGpA$, $m^7G_{3'OMe}pppGpC$, $m^7G_{3'OMe}pppGpG$,
 $m^7G_{3'OMe}pppGpU$, $m^7G_{3'OMe}pppUpA$, $m^7G_{3'OMe}pppUpC$, $m^7G_{3'OMe}pppUpG$, and
 $m^7G_{3'OMe}pppUpU$.

In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppApA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppApC$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppApG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppApU$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppCpA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppCpC$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppCpG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppCpU$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppGpA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppGpC$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppGpG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppGpU$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppUpA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppUpC$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppUpG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppUpU$.

A trinucleotide cap, in other embodiments, comprises a sequence selected from the following sequences: $m^7G_{3'OMe}pppA_{2'OMe}pA$, $m^7G_{3'OMe}pppA_{2'OMe}pC$, $m^7G_{3'OMe}pppA_{2'OMe}pG$, $m^7G_{3'OMe}pppA_{2'OMe}pU$, $m^7G_{3'OMe}pppC_{2'OMe}pA$, $m^7G_{3'OMe}pppC_{2'OMe}pC$, $m^7G_{3'OMe}pppC_{2'OMe}pG$, $m^7G_{3'OMe}pppC_{2'OMe}pU$, $m^7G_{3'OMe}pppG_{2'OMe}pA$, $m^7G_{3'OMe}pppG_{2'OMe}pC$, $m^7G_{3'OMe}pppG_{2'OMe}pG$, $m^7G_{3'OMe}pppG_{2'OMe}pU$, $m^7G_{3'OMe}pppU_{2'OMe}pA$, $m^7G_{3'OMe}pppU_{2'OMe}pC$, $m^7G_{3'OMe}pppU_{2'OMe}pG$, and $m^7G_{3'OMe}pppU_{2'OMe}pU$.

In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppA_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppA_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppA_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppA_{2'OMe}pU$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppC_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppC_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppC_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppC_{2'OMe}pU$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppG_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppG_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises

$m^7G_{3'OMe}pppG_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppG_{2'OMe}pU$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppU_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppU_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppU_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7G_{3'OMe}pppU_{2'OMe}pU$.

A trinucleotide cap, in still other embodiments, comprises a sequence selected from the following sequences: $m^7GpppA_{2'OMe}pA$, $m^7GpppA_{2'OMe}pC$, $m^7GpppA_{2'OMe}pG$, $m^7GpppA_{2'OMe}pU$, $m^7GpppC_{2'OMe}pA$, $m^7GpppC_{2'OMe}pC$, $m^7GpppC_{2'OMe}pG$, $m^7GpppC_{2'OMe}pU$, $m^7GpppG_{2'OMe}pA$, $m^7GpppG_{2'OMe}pC$, $m^7GpppG_{2'OMe}pG$, $m^7GpppG_{2'OMe}pU$, $m^7GpppU_{2'OMe}pA$, $m^7GpppU_{2'OMe}pC$, $m^7GpppU_{2'OMe}pG$, and $m^7GpppU_{2'OMe}pU$.

In some embodiments, a trinucleotide cap comprises $m^7GpppA_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7GpppA_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises $m^7GpppA_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7GpppA_{2'OMe}pU$. In some embodiments, a trinucleotide cap comprises $m^7GpppC_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7GpppC_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises $m^7GpppC_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7GpppC_{2'OMe}pU$. In some embodiments, a trinucleotide cap comprises $m^7GpppG_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7GpppG_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises $m^7GpppG_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7GpppG_{2'OMe}pU$. In some embodiments, a trinucleotide cap comprises $m^7GpppU_{2'OMe}pA$. In some embodiments, a trinucleotide cap comprises $m^7GpppU_{2'OMe}pC$. In some embodiments, a trinucleotide cap comprises $m^7GpppU_{2'OMe}pG$. In some embodiments, a trinucleotide cap comprises $m^7GpppU_{2'OMe}pU$.

In some embodiments, a trinucleotide cap comprises GAG. In some embodiments, a trinucleotide cap comprises GCG. In some embodiments, a trinucleotide cap comprises GUG. In some embodiments, a trinucleotide cap comprises GGG.

In Vitro Transcription Methods

Some aspects of the present disclosure provide methods of producing (synthesizing) a RNA transcript (e.g., mRNA transcript) comprising contacting a DNA template with a RNA

polymerase (e.g., a T7 RNA polymerase such as a T7 RNA polymerase variant) under conditions that result in the production of RNA transcript.

In some embodiments, the methods comprise contacting a DNA template with a T7 RNA polymerase (e.g., SEQ ID NO:1, 99, or 100) variant having a S43A substitution (e.g., a S43A T7 RNAP variant or a S43A* T7 RNAP variant) under conditions that result in the production of RNA transcript. In other embodiments, the methods comprise contacting a DNA template with a T7 RNA polymerase (e.g., SEQ ID NO:1, 99, or 100) variant having a G47A substitution (e.g., a G47A T7 RNAP variant or a G47A* T7 RNAP variant) under conditions that result in the production of RNA transcript. In yet other embodiments, the methods comprise contacting a DNA template with a T7 RNA polymerase (e.g., SEQ ID NO:1, 99, or 100) variant having a R257A substitution (e.g., a R257A T7 RNAP variant or a R257A* T7 RNAP variant) under conditions that result in the production of RNA transcript. In still other embodiments, the methods comprise contacting a DNA template with a T7 RNA polymerase (e.g., SEQ ID NO:1, 99, or 100) variant having a G259A substitution (e.g., a G259A T7 RNAP variant or a G259A* T7 RNAP variant) under conditions that result in the production of RNA transcript.

In some embodiments, the methods comprise contacting a DNA template with a T7 RNA polymerase variant that comprises an (at least one) additional C terminal amino acid (e.g., Gly, Ala, GlyGly, AlaAla, GlyAla, or AlaGly).

In some embodiments, the methods comprise contacting a DNA template with a T7 RNA polymerase variant that comprises a E42R, S43A, S43E, S43L, S43R, E45R, E45L, M46A, G47A, G47E, G47L, G47R, N165W, E167M, E167N, E168I, E168T, E168V, A181F, A181W, G184M, E187F, A255Q, A255K, A255I, A255Y, R257A, R257E, R257L, R257W, G259A, G259E, G259L, G259R, A260W, or A260R substitution, or any combination of two or more of the foregoing substitutions and optionally an additional C terminal amino acid (e.g., Gly, Ala, GlyGly, AlaAla, GlyAla, or AlaGly).

In some aspects, the present disclosure provides methods of performing an IVT reaction, comprising contacting a DNA template with the RNA polymerase (e.g., a T7 RNA polymerase, such as a T7 RNA polymerase variant) in the presence of nucleoside triphosphates and buffer under conditions that result in the production of RNA transcripts.

Other aspects of the present disclosure provide co-transcriptional capping methods that comprise reacting a polynucleotide template with a T7 RNA polymerase variant, nucleoside triphosphates, and a cap analog under in vitro transcription reaction conditions to produce RNA transcript.

In some embodiments, a co-transcriptional capping method for RNA synthesis comprises reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a trinucleotide cap comprising sequence GpppA_{2'}OmepG, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.

IVT conditions typically require a purified linear DNA template containing a promoter, nucleoside triphosphates, a buffer system that includes dithiothreitol (DTT) and magnesium ions, and a RNA polymerase. The exact conditions used in the transcription reaction depend on the amount of RNA needed for a specific application. Typical IVT reactions are performed by incubating a DNA template with a RNA polymerase and nucleoside triphosphates, including GTP, ATP, CTP, and UTP (or nucleotide analogs) in a transcription buffer. A RNA transcript having a 5' terminal guanosine triphosphate is produced from this reaction.

A deoxyribonucleic acid (DNA) is simply a nucleic acid template for RNA polymerase. A DNA template may include a polynucleotide encoding a polypeptide of interest (e.g., an antigenic polypeptide). A DNA template, in some embodiments, includes a RNA polymerase promoter (e.g., a T7 RNA polymerase promoter) located 5' from and operably linked to polynucleotide encoding a polypeptide of interest. A DNA template may also include a nucleotide sequence encoding a polyadenylation (polyA) tail located at the 3' end of the gene of interest.

Polypeptides of interest include, but are not limited to, biologics, antibodies, antigens (vaccines), and therapeutic proteins. The term "protein" encompasses peptides.

A RNA transcript, in some embodiments, is the product of an IVT reaction. A RNA transcript, in some embodiments, is a messenger RNA (mRNA) that includes a nucleotide sequence encoding a polypeptide of interest linked to a polyA tail. In some embodiments, the mRNA is modified mRNA (mmRNA), which includes at least one modified nucleotide.

A nucleotide includes a nitrogenous base, a five-carbon sugar (ribose or deoxyribose), and at least one phosphate group. Nucleotides include nucleoside monophosphates, nucleoside diphosphates, and nucleoside triphosphates. A nucleoside monophosphate (NMP)

includes a nucleobase linked to a ribose and a single phosphate; a nucleoside diphosphate (NDP) includes a nucleobase linked to a ribose and two phosphates; and a nucleoside triphosphate (NTP) includes a nucleobase linked to a ribose and three phosphates.

Nucleotide analogs are compounds that have the general structure of a nucleotide or are structurally similar to a nucleotide. Nucleotide analogs, for example, include an analog of the nucleobase, an analog of the sugar and/or an analog of the phosphate group(s) of a nucleotide.

A nucleoside includes a nitrogenous base and a 5-carbon sugar. Thus, a nucleoside plus a phosphate group yields a nucleotide. Nucleoside analogs are compounds that have the general structure of a nucleoside or are structurally similar to a nucleoside. Nucleoside analogs, for example, include an analog of the nucleobase and/or an analog of the sugar of a nucleoside.

It should be understood that the term “nucleotide” includes naturally-occurring nucleotides, synthetic nucleotides and modified nucleotides, unless indicated otherwise. Examples of naturally-occurring nucleotides used for the production of RNA, e.g., in an IVT reaction, as provided herein include adenosine triphosphate (ATP), guanosine triphosphate (GTP), cytidine triphosphate (CTP), uridine triphosphate (UTP), and 5-methyluridine triphosphate (m^5 UTP). In some embodiments, adenosine diphosphate (ADP), guanosine diphosphate (GDP), cytidine diphosphate (CDP), and/or uridine diphosphate (UDP) are used.

Examples of nucleotide analogs include, but are not limited to, antiviral nucleotide analogs, phosphate analogs (soluble or immobilized, hydrolyzable or non-hydrolyzable), dinucleotide, trinucleotide, tetranucleotide, e.g., a cap analog, or a precursor/substrate for enzymatic capping (vaccinia or ligase), a nucleotide labeled with a functional group to facilitate ligation/conjugation of cap or 5' moiety (IRES), a nucleotide labeled with a 5' PO_4 to facilitate ligation of cap or 5' moiety, or a nucleotide labeled with a functional group/protecting group that can be chemically or enzymatically cleaved. Examples of antiviral nucleotide/nucleoside analogs include, but are not limited, to Ganciclovir, Entecavir, Telbivudine, Vidarabine and Cidofovir.

Modified nucleotides may include modified nucleobases. For example, a RNA transcript (e.g., mRNA transcript) of the present disclosure may include a modified nucleobase selected from pseudouridine (ψ), 1-methylpseudouridine ($m1\psi$), 1-ethylpseudouridine, 2-thiouridine, 4'-thiouridine, 2-thio-1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-pseudouridine, 2-thio-5-aza-uridine, 2-thio-dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-pseudouridine, 4-methoxy-2-thio-pseudouridine, 4-methoxy-

pseudouridine, 4-thio-1-methyl-pseudouridine, 4-thio-pseudouridine, 5-aza-uridine, dihydropseudouridine, 5-methyluridine, 5-methoxyuridine (mo5U) and 2'-O-methyl uridine. In some embodiments, a RNA transcript (e.g., mRNA transcript) includes a combination of at least two (e.g., 2, 3, 4 or more) of the foregoing modified nucleobases.

The nucleoside triphosphates (NTPs) as provided herein may comprise unmodified or modified ATP, modified or unmodified UTP, modified or unmodified GTP, and/or modified or unmodified CTP. In some embodiments, NTPs of an IVT reaction comprise unmodified ATP. In some embodiments, NTPs of an IVT reaction comprise modified ATP. In some embodiments, NTPs of an IVT reaction comprise unmodified UTP. In some embodiments, NTPs of an IVT reaction comprise modified UTP. In some embodiments, NTPs of an IVT reaction comprise unmodified GTP. In some embodiments, NTPs of an IVT reaction comprise modified GTP. In some embodiments, NTPs of an IVT reaction comprise unmodified CTP. In some embodiments, NTPs of an IVT reaction comprise modified CTP.

The concentration of nucleoside triphosphates and cap analog present in an IVT reaction may vary. In some embodiments, NTPs and cap analog are present in the reaction at equimolar concentrations. In some embodiments, the molar ratio of cap analog (e.g., trinucleotide cap) to nucleoside triphosphates in the reaction is greater than 1:1. For example, the molar ratio of cap analog to nucleoside triphosphates in the reaction may be 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 15:1, 20:1, 25:1, 50:1, or 100:1. In some embodiments, the molar ratio of cap analog (e.g., trinucleotide cap) to nucleoside triphosphates in the reaction is less than 1:1. For example, the molar ratio of cap analog (e.g., trinucleotide cap) to nucleoside triphosphates in the reaction may be 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:15, 1:20, 1:25, 1:50, or 1:100.

The composition of NTPs in an IVT reaction may also vary. For example, ATP may be used in excess of GTP, CTP and UTP. As a non-limiting example, an IVT reaction may include 7.5 millimolar GTP, 7.5 millimolar CTP, 7.5 millimolar UTP, and 3.75 millimolar ATP. The same IVT reaction may include 3.75 millimolar cap analog (e.g., trinucleotide cap). In some embodiments, the molar ratio of G:C:U:A:cap is 1:1:1:0.5:0.5. In some embodiments, the molar ratio of G:C:U:A:cap is 1:1:0.5:1:0.5. In some embodiments, the molar ratio of G:C:U:A:cap is 1:0.5:1:1:0.5. In some embodiments, the molar ratio of G:C:U:A:cap is 0.5:1:1:1:0.5.

In some embodiments, a RNA transcript (e.g., mRNA transcript) includes a modified nucleobase selected from pseudouridine (ψ), 1-methylpseudouridine ($m^1\psi$), 5-methoxyuridine (mo^5U), 5-methylcytidine (m^5C), α -thio-guanosine and α -thio-adenosine. In

some embodiments, a RNA transcript (e.g., mRNA transcript) includes a combination of at least two (e.g., 2, 3, 4 or more) of the foregoing modified nucleobases.

In some embodiments, a RNA transcript (e.g., mRNA transcript) includes pseudouridine (ψ). In some embodiments, a RNA transcript (e.g., mRNA transcript) includes 1-methylpseudouridine ($m^1\psi$). In some embodiments, a RNA transcript (e.g., mRNA transcript) includes 5-methoxyuridine (m^5U). In some embodiments, a RNA transcript (e.g., mRNA transcript) includes 5-methylcytidine (m^5C). In some embodiments, a RNA transcript (e.g., mRNA transcript) includes α -thio-guanosine. In some embodiments, a RNA transcript (e.g., mRNA transcript) includes α -thio-adenosine.

In some embodiments, the polynucleotide (e.g., RNA polynucleotide, such as mRNA polynucleotide) is uniformly modified (e.g., fully modified, modified throughout the entire sequence) for a particular modification. For example, a polynucleotide can be uniformly modified with 1-methylpseudouridine ($m^1\psi$), meaning that all uridine residues in the mRNA sequence are replaced with 1-methylpseudouridine ($m^1\psi$). Similarly, a polynucleotide can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as any of those set forth above. Alternatively, the polynucleotide (e.g., RNA polynucleotide, such as mRNA polynucleotide) may not be uniformly modified (e.g., partially modified, part of the sequence is modified). Each possibility represents a separate embodiment of the present invention.

In some embodiments, the buffer system contains tris. The concentration of tris used in an IVT reaction, for example, may be at least 10 mM, at least 20 mM, at least 30 mM, at least 40 mM, at least 50 mM, at least 60 mM, at least 70 mM, at least 80 mM, at least 90 mM, at least 100 mM or at least 110 mM phosphate. In some embodiments, the concentration of phosphate is 20-60 mM or 10-100 mM.

In some embodiments, the buffer system contains dithiothreitol (DTT). The concentration of DTT used in an IVT reaction, for example, may be at least 1 mM, at least 5 mM, or at least 50 mM. In some embodiments, the concentration of DTT used in an IVT reaction is 1-50 mM or 5-50 mM. In some embodiments, the concentration of DTT used in an IVT reaction is 5 mM.

In some embodiments, the buffer system contains magnesium. In some embodiments, the molar ratio of NTP to magnesium ions (Mg^{2+} ; e.g., $MgCl_2$) present in an IVT reaction is 1:1 to 1:5. For example, the molar ratio of NTP to magnesium ions may be 1:1, 1:2, 1:3, 1:4 or 1:5.

In some embodiments, the molar ratio of NTP plus cap analog (e.g., trinucleotide cap, such as GAG) to magnesium ions (Mg^{2+} ; e.g., $MgCl_2$) present in an IVT reaction is 1:1 to 1:5. For example, the molar ratio of NTP+trinucleotide cap (e.g., GAG) to magnesium ions may be 1:1, 1:2, 1:3, 1:4 or 1:5.

In some embodiments, the buffer system contains Tris-HCl, spermidine (e.g., at a concentration of 1-30 mM), TRITON[®] X-100 (polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether) and/or polyethylene glycol (PEG).

The addition of nucleoside triphosphates (NTPs) to the 3' end of a growing RNA strand is catalyzed by a polymerase, such as T7 RNA polymerase, for example, any one or more of the T7 RNA polymerase variants (e.g., S43A and/or G47A) of the present disclosure. In some embodiments, the RNA polymerase (e.g., T7 RNA polymerase variant) is present in a reaction (e.g., an IVT reaction) at a concentration of 0.01 mg/ml to 1 mg/ml. For example, the RNA polymerase may be present in a reaction at a concentration of 0.01 mg/mL, 0.05 mg/ml, 0.1 mg/ml, 0.5 mg/ml or 1.0 mg/ml.

Surprisingly, use of the combination of a T7 RNAP variant (e.g., E42, S43, Y44, E45, M46, G47, A255, R257, or G259, e.g., S43A or G47A) as provided herein with a cap analog (e.g., GpppA_{2'}Ome pG), in an in vitro transcription reaction, for example, results in the production of RNA transcript, wherein greater than 80% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 85% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 90% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 95% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 96% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 97% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 98% of the RNA transcript produced includes a functional cap. In some embodiments, greater than 99% of the RNA transcript produced includes a functional cap.

Also surprising was the finding that use of a polynucleotide template that includes a 2'-deoxythymidine residue or 2'-deoxycytidine residue at template position +1 results in the production of RNA transcript, wherein greater than 80% (e.g., greater than 85%, greater than 90%, or greater than 95%) of the RNA transcript produced includes a functional cap. Thus, in some embodiments, a polynucleotide (e.g., DNA) template used, for example, in an IVT reaction, includes a 2'-deoxythymidine residue at template position +1. In other

embodiments, a polynucleotide (e.g., DNA) template used, for example, in an IVT reaction, includes a 2'-deoxycytidine residue at template position +1.

Applications

The RNA transcripts produced according to the present disclosure include mRNA (including modified mRNA and/or unmodified RNA), lncRNA, self-replicating RNA, circular RNA, CRISPR guide RNA, and the like. In embodiments, the RNA is RNA (e.g., mRNA or self-replicating RNA) that encodes a polypeptide (e.g., a therapeutic polypeptide). Thus, the RNA transcripts produced using RNA polymerase variants of the present disclosure may be used in a myriad of applications.

For example, the RNA transcripts may be used to produce polypeptides of interest, e.g., therapeutic proteins, vaccine antigen, and the like. In some embodiments, the RNA transcripts are therapeutic RNAs. A therapeutic mRNA is an mRNA that encodes a therapeutic protein (the term 'protein' encompasses peptides). Therapeutic proteins mediate a variety of effects in a host cell or in a subject to treat a disease or ameliorate the signs and symptoms of a disease. For example, a therapeutic protein can replace a protein that is deficient or abnormal, augment the function of an endogenous protein, provide a novel function to a cell (e.g., inhibit or activate an endogenous cellular activity, or act as a delivery agent for another therapeutic compound (e.g., an antibody-drug conjugate). Therapeutic mRNA may be useful for the treatment of the following diseases and conditions: bacterial infections, viral infections, parasitic infections, cell proliferation disorders, genetic disorders, and autoimmune disorders. Other diseases and conditions are encompassed herein.

A protein of interest encoded by an mRNA as provided herein can be essentially any protein. In some embodiments, the therapeutic protein is a cytokine, a growth factor, an antibody or a fusion protein. Non-limiting examples of therapeutic proteins include blood factors (such as Factor VIII and Factor VII), complement factors, Low Density Lipoprotein Receptor (LDLR) and MUT1. Non-limiting examples of cytokines include interleukins, interferons, chemokines, lymphokines and the like. Non-limiting examples of growth factors include erythropoietin, EGFs, PDGFs, FGFs, TGFs, IGFs, TNFs, CSFs, MCSFs, GMCSFs and the like. Non-limiting examples of antibodies include adalimumab, infliximab, rituximab, ipilimumab, tocilizumab, canakinumab, itolizumab, talokinumab. Non-limiting examples of fusion proteins include, for example, etanercept, abatacept and belatacept.

In some embodiments, the protein of interest is human erythropoietin, LDLR (for use in inhibiting cholesterol), or MUT1 (for use in the treatment of methylmalonic acidemia (MMA)). In other embodiments, the protein of interest encoded by the mRNA is a therapeutic antibody, including but not limited to the antibodies listed above.

A RNA transcript produced using a RNA polymerase variant as disclosed herein may encode one or more biologics. A biologic is a polypeptide-based molecule that may be used to treat, cure, mitigate, prevent, or diagnose a serious or life-threatening disease or medical condition. Biologics include, but are not limited to, allergenic extracts (e.g. for allergy shots and tests), blood components, gene therapy products, human tissue or cellular products used in transplantation, vaccines, monoclonal antibodies, cytokines, growth factors, enzymes, thrombolytics, and immunomodulators, among others.

One or more biologics currently being marketed or in development may be encoded by the RNA of the present invention. While not wishing to be bound by theory, it is believed that incorporation of the encoding polynucleotides of a known biologic into the RNA of the present disclosure will result in improved therapeutic efficacy due at least in part to the specificity, purity and/or selectivity of the construct designs.

A RNA transcript produced using a RNA polymerase variant as disclosed herein may encode one or more antibodies. The term “antibody” includes monoclonal antibodies (including full length antibodies which have an immunoglobulin Fc region), antibody compositions with polypeptidic specificity, multispecific antibodies (e.g., bispecific antibodies, diabodies, and single-chain molecules), as well as antibody fragments. The term “immunoglobulin” (Ig) is used interchangeably with “antibody” herein. A monoclonal antibody is an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations and/or post-translation modifications (e.g., isomerizations, amidations) that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site.

Monoclonal antibodies specifically include chimeric antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is(are) identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity. Chimeric antibodies include, but are not

limited to, “primatized” antibodies comprising variable domain antigen-binding sequences derived from a non-human primate (e.g., Old World Monkey, Ape etc.) and human constant region sequences.

Antibodies encoded in the RNA of the present disclosure may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, blood, cardiovascular, CNS, poisoning (including antivenoms), dermatology, endocrinology, gastrointestinal, medical imaging, musculoskeletal, oncology, immunology, respiratory, sensory and anti-infective.

A RNA transcript produced using a RNA polymerase variant as disclosed herein may encode one or more vaccine antigens. A vaccine antigen is a biological preparation that improves immunity to a particular disease or infectious agent. One or more vaccine antigens currently being marketed or in development may be encoded by the RNA of the present disclosure. Vaccine antigens encoded in the RNA may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, cancer, allergy and infectious disease. In some embodiments, a cancer vaccine may be a personalized cancer vaccine in the form of a concatemer or individual RNAs encoding peptide epitopes or a combination thereof.

A RNA transcript produced using a RNA polymerase variant as disclosed herein may be designed to encode one or more antimicrobial peptides (AMP) or antiviral peptides (AVP). AMPs and AVPs have been isolated and described from a wide range of animals such as, but not limited to, microorganisms, invertebrates, plants, amphibians, birds, fish, and mammals. The anti-microbial polypeptides may block cell fusion and/or viral entry by one or more enveloped viruses (e.g., HIV, HCV). For example, the anti-microbial polypeptide can comprise or consist of a synthetic peptide corresponding to a region, e.g., a consecutive sequence of at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, or 60 amino acids of the transmembrane subunit of a viral envelope protein, e.g., HIV-1 gp120 or gp41. The amino acid and nucleotide sequences of HIV-1 gp120 or gp41 are described in, e.g., Kuiken et al., (2008). "HIV Sequence Compendium," Los Alamos National Laboratory.

In some embodiments, RNA transcripts are used as radiolabeled RNA probes. In some embodiments, RNA transcripts are used for non-isotopic RNA labeling. In some embodiments, RNA transcripts are used as guide RNA (gRNA) for gene targeting. In some embodiments, RNA transcripts (e.g., mRNA) are used for in vitro translation and micro injection. In some embodiments, RNA transcripts are used for RNA structure, processing and catalysis studies. In some embodiments, RNA transcripts are used for RNA

amplification. In some embodiments, RNA transcripts are used as anti-sense RNA for gene expression experiment. Other applications are encompassed by the present disclosure.

Compositions

The T7 RNAP variants of the present disclosure, in some embodiments when used in combination with a cap analog, such as a trinucleotide cap, in an IVT reaction produces RNA that does not induce a detectable cytokine response, even in the absence of post-IVT purification. Thus, provided herein, in some embodiments, are composition comprising IVT RNA and a pharmaceutically acceptable excipient, wherein the composition is substantially free (e.g., does not comprise, or comprises less than 10%, less than 1%, less than 0.1%, or less than 0.01%) of cytokine-inducing RNA contaminant in the absence of post-IVT purification.

As described elsewhere herein, the T7 RNAP variants and methods of the present disclosure produce RNA transcript, wherein at least 80% (e.g., 80%-90%, 80-95%, 90-95%, 80-99%, 90-99%, or 90-100%) of the transcript includes a functional cap. Thus, also provided herein are composition comprising an IVT RNA and a pharmaceutically acceptable excipient, wherein the composition comprises less than 20% (e.g., 5-15%, 5-10%, 1-15%, or 1-10%) uncapped RNA species. In some embodiments, the composition comprises less than 15% (e.g., 5-10%, 1-10%, or 1-5%) uncapped RNA species. In some embodiments, the composition comprises less than 10% uncapped RNA species. In some embodiments, the composition comprises less than 5% uncapped RNA species.

In some embodiments, greater than 80% of the IVT RNA includes a functional cap. In some embodiments, greater than 85% of the IVT RNA includes a functional cap. In some embodiments, greater than 90% of the IVT RNA includes a functional cap. In some embodiments, greater than 95% of the IVT RNA includes a functional cap.

The IVT RNA, in some embodiments, is not chemically modified, while in other embodiments, the IVT RNA is chemically modified.

In some embodiments, unexpectedly, greater than 80% of the IVT RNA comprises single-stranded full-length transcripts. For example, greater than 85% of the IVT RNA may comprise single-stranded full-length transcripts. In some embodiments, greater than 90% of the IVT RNA comprises single-stranded full-length transcripts. greater than 95% of the IVT RNA comprises single-stranded full-length transcripts.

In some embodiments, the RNA is produced by a process comprising reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino

acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a trinucleotide cap comprising sequence GpppA_{2'}Ome pG, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.

Kits

Also provided herein are kits, such as in vitro transcription kits comprising a RNA polymerase variant of the present disclosure. The kits may comprise any one or more (at least one) IVT component described herein and any one or more (at least one) RNA polymerase variant. For example, a kit may include a buffer system, NTPs and a T7 RNA polymerase variant having an amino acid sequence of SEQ ID NO:1, 99, or 100 having at least one (or at least two, or at least three) amino acid substitution at position E42, S43, Y44, E45, M46, G47, A255, R257, A258, G259, A260, L261 and/or A262 and optionally at least one amino acid addition on the C-terminal end.

Additional Embodiments

Additional embodiments of the present disclosure are encompassed by the following numbered paragraphs:

1. A ribonucleic acid (RNA) polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex.
2. The RNA polymerase variant of paragraph 1, wherein the at least one amino acid substitution has a higher helix propensity, relative to wild-type amino acid.
3. The RNA polymerase variant of paragraph 1 or 2, wherein the RNA polymerase is a T7 RNA polymerase.
4. The RNA polymerase variant of any one of paragraphs 1-3, wherein the at least one loop structure is in a C-helix structure.
5. The RNA polymerase variant of any one of paragraphs 1-4, wherein the at least one loop structure is in a C-linker structure.

6. The RNA polymerase variant of any one of paragraphs 1-4, wherein the at least one amino acid substitution is at least one high-helix propensity amino acid substitution.
7. The RNA polymerase variant of paragraph 6, wherein the at least one high-helix propensity acid substitution is selected from alanine, isoleucine, leucine, methionine, lysine, glutamine, and/or glutamate.
8. The RNA polymerase variant of paragraph 7, wherein the at least one high-helix propensity amino acid substitution is alanine.
9. The RNA polymerase variant of any one of paragraphs 4-8, wherein the T7 RNA polymerase comprising an amino acid sequence identified by SEQ ID NO:1, SEQ ID NO:99, or SEQ ID NO:100 modified to include at least one amino acid substitution of a high-helix propensity amino acid at a position selected from E42, S43, Y44, E45, M46, G47, R257, and G259.
10. The RNA polymerase variant of paragraph 9, wherein that at least one amino acid substitution comprises S43A.
11. The RNA polymerase variant of paragraph 9, wherein that at least one amino acid substitution comprises G47A.
12. The RNA polymerase variant of any one of paragraphs 5-8, wherein the T7 RNA polymerase comprising an amino acid sequence identified by SEQ ID NO:1, SEQ ID NO:99, or SEQ ID NO:100 modified to include at least one amino acid substitution of a high-helix propensity amino acid at a position selected from R257, A258, G259, A260, L261 and A262.
13. The RNA polymerase variant of paragraph 12, wherein the at least one amino acid substitution comprises R257A.
14. The RNA polymerase variant of paragraph 12, wherein the at least one amino acid substitution comprises G259A.
15. A T7 ribonucleic acid (RNA) polymerase comprising an amino acid sequence identified by SEQ ID NO:1, SEQ ID NO:99, or SEQ ID NO:100 modified to include an amino acid substitution of a high-helix propensity amino acid at position G47, S43, R257, or G259.
16. The T7 RNA polymerase of paragraph 15, wherein the high-helix propensity amino acid is selected from alanine, isoleucine, leucine, methionine, lysine, glutamine, and/or glutamate.
17. The T7 RNA polymerase of paragraph 16, wherein the high-helix propensity amino acid is alanine.

18. A T7 RNA polymerase comprising an amino acid sequence identified by SEQ ID NO:2, SEQ ID NO:107, or SEQ ID NO:108.
19. A T7 RNA polymerase comprising an amino acid sequence identified by SEQ ID NO:3, SEQ ID NO:109, or SEQ ID NO:110.
20. A T7 RNA polymerase comprising an amino acid sequence identified by SEQ ID NO:4, SEQ ID NO:111, or SEQ ID NO:112.
21. A T7 RNA polymerase comprising an amino acid sequence identified by SEQ ID NO:5, SEQ ID NO:113, or SEQ ID NO:114.
22. A method of producing a ribonucleic acid (RNA) comprising contacting a DNA template with the RNA polymerase of any one of paragraphs 1-21 under conditions that result in the production of RNA transcript.
23. A method of performing an in vitro transcription (IVT) reaction, comprising contacting a DNA template with the RNA polymerase of any one of paragraphs 1-21 in the presence of nucleoside triphosphates and buffer under conditions that result in the production of RNA transcripts.
24. The method of paragraph 23, wherein the RNA transcript produced, when delivered to cells, optionally in unpurified form, stimulates cytokine response that is at least 50% lower relative to RNA produced using wild-type RNA polymerase.
25. The method of any one of paragraphs 23-24, wherein the concentration of double-stranded RNA transcript produced by IVT is at least 50% lower relative to dsRNA transcript produced using wild-type polymerase.
26. The method of any one of paragraphs 23-25, wherein less than 20% of the RNA transcripts produced exhibit 3' heterogeneity.
27. The method of any one of paragraphs 23-26, wherein less than 50% of the RNA transcript produced is truncated RNA transcript.
28. The method of any one of paragraphs 23-28, wherein less than 50% of the RNA transcript produced is run-on RNA transcript.
29. The method of any one of paragraphs 23-28, wherein the amount of full-length RNA transcript produced is at least 15 times greater than the amount of the DNA template
30. The method of any one of paragraphs 23-29, wherein the ratio of truncated RNA transcript:full-length RNA transcript produced is less than 1:1.
31. The method of any one of paragraphs 23-30, wherein the RNA transcript produced has less than 1 mutation per 100 nucleotides relative to the DNA template.
32. A nucleic acid encoding the RNA polymerase of any one of paragraphs 1-21.

33. A vector comprising the nucleic acid of paragraph 33.
34. A host cell comprising the nucleic acid of paragraph 33 or the vector of paragraph 34.
35. A kit comprising the RNA polymerase of any one of paragraphs 1-21.
36. A composition comprising the RNA polymerase of any one of paragraphs 1-21.
37. A ribonucleic acid (RNA) produced by the method of any one of paragraphs 22-31.
38. The RNA of paragraph 37 formulated in a lipid nanoparticle.
39. The RNA of paragraph 38, wherein the lipid nanoparticle comprises a molar ratio of 20-60% ionizable amino lipid, 5-25% non-cationic lipid, 25-55% sterol, and 0.5-15% PEG-modified lipid.
40. A co-transcriptional capping method for ribonucleic acid (RNA) synthesis, the method comprising reacting a polynucleotide template with a T7 RNA polymerase variant, nucleoside triphosphates, and a cap analog under in vitro transcription reaction conditions to produce RNA transcript.
41. The method of paragraph 40, wherein greater than 80%, greater than 85%, or greater than 90% of the RNA transcript produced includes a functional cap.
42. The method of paragraph 41, wherein greater than 95% of the RNA transcript produced includes a functional cap.
43. The method of any one of paragraphs 40-42, wherein the nucleoside triphosphates comprise unmodified or modified ATP, modified or unmodified UTP, modified or unmodified GTP, and/or modified or unmodified CTP.
44. The method of any one of paragraphs 40-43, wherein the T7 RNA polymerase variant is a T7 polymerase variant of any one of paragraph 1-21.
45. The method of paragraph 44, wherein the T7 polymerase variant comprises an amino acid sequence identified by SEQ ID NO:1, SEQ ID NO:99, or SEQ ID NO:100 modified to include at least one amino acid substitution of a high-propensity amino acid at a position selected from E42, S43, Y44, E45, M46, G47, R257, and G259.
46. The method of any one of paragraphs 40-45, the T7 polymerase variant comprises an amino acid sequence identified by SEQ ID NO:1, SEQ ID NO:99, or SEQ ID NO:100 modified to include amino acid substitution of G47A.
47. The method of any one of paragraphs 40-45, the T7 polymerase variant comprises an amino acid sequence identified by SEQ ID NO:1, SEQ ID NO:99, or SEQ ID NO:100 modified to include amino acid substitution of S43A.
48. The method of any one of paragraphs 40-47, wherein the nucleoside triphosphates and cap analog are present in the reaction at equimolar concentrations.

49. The method of any one of paragraphs 40-47, wherein a molar ratio of cap analog to nucleoside triphosphates in the reaction is greater than 1:1.
50. The method of any one of paragraphs 40-47, wherein a molar ratio of cap analog to nucleoside triphosphates in the reaction is less than 1:1.
51. The method of any one of paragraphs 40-50, wherein the cap analog is a dinucleotide cap, a trinucleotide cap, or a tetranucleotide cap.
52. The method of any one of paragraphs 40-50, wherein the cap analog is a trinucleotide cap.
53. The method of paragraph 51, wherein the trinucleotide cap comprises a sequence selected from the following sequences: GAA, GAC, GAG, GAU, GCA, GCC, GCG, GCU, GGA, GGC, GGG, GGU, GUA, GUC, GUG, and GUU.
54. The method of paragraph 53, wherein the trinucleotide cap comprises a sequence selected from the following sequences: m⁷GpppApA, m⁷GpppApC, m⁷GpppApG, m⁷GpppApU, m⁷GpppCpA, m⁷GpppCpC, m⁷GpppCpG, m⁷GpppCpU, m⁷GpppGpA, m⁷GpppGpC, m⁷GpppGpG, m⁷GpppGpU, m⁷GpppUpA, m⁷GpppUpC, m⁷GpppUpG, and m⁷GpppUpU.
55. The method of paragraph 53, wherein the trinucleotide cap comprises a sequence selected from the following sequences: m⁷G_{3'OMe}pppApA, m⁷G_{3'OMe}pppApC, m⁷G_{3'OMe}pppApG, m⁷G_{3'OMe}pppApU, m⁷G_{3'OMe}pppCpA, m⁷G_{3'OMe}pppCpC, m⁷G_{3'OMe}pppCpG, m⁷G_{3'OMe}pppCpU, m⁷G_{3'OMe}pppGpA, m⁷G_{3'OMe}pppGpC, m⁷G_{3'OMe}pppGpG, m⁷G_{3'OMe}pppGpU, m⁷G_{3'OMe}pppUpA, m⁷G_{3'OMe}pppUpC, m⁷G_{3'OMe}pppUpG, and m⁷G_{3'OMe}pppUpU.
56. The method of paragraph 53, wherein the trinucleotide cap comprises a sequence selected from the following sequences: m⁷G_{3'OMe}pppA_{2'OMe}pA, m⁷G_{3'OMe}pppA_{2'OMe}pC, m⁷G_{3'OMe}pppA_{2'OMe}pG, m⁷G_{3'OMe}pppA_{2'OMe}pU, m⁷G_{3'OMe}pppC_{2'OMe}pA, m⁷G_{3'OMe}pppC_{2'OMe}pC, m⁷G_{3'OMe}pppC_{2'OMe}pG, m⁷G_{3'OMe}pppC_{2'OMe}pU, m⁷G_{3'OMe}pppG_{2'OMe}pA, m⁷G_{3'OMe}pppG_{2'OMe}pC, m⁷G_{3'OMe}pppG_{2'OMe}pG, m⁷G_{3'OMe}pppG_{2'OMe}pU, m⁷G_{3'OMe}pppU_{2'OMe}pA, m⁷G_{3'OMe}pppU_{2'OMe}pC, m⁷G_{3'OMe}pppU_{2'OMe}pG, and m⁷G_{3'OMe}pppU_{2'OMe}pU.
57. The method of paragraph 53, wherein the trinucleotide cap comprises a sequence selected from the following sequences: m⁷GpppA_{2'OMe}pA, m⁷GpppA_{2'OMe}pC, m⁷GpppA_{2'OMe}pG, m⁷GpppA_{2'OMe}pU, m⁷GpppC_{2'OMe}pA, m⁷GpppC_{2'OMe}pC, m⁷GpppC_{2'OMe}pG, m⁷GpppC_{2'OMe}pU, m⁷GpppG_{2'OMe}pA, m⁷GpppG_{2'OMe}pC,

m⁷GpppG_{2'OMe}pG, m⁷GpppG_{2'OMe}pU, m⁷GpppU_{2'OMe}pA, m⁷GpppU_{2'OMe}pC, m⁷GpppU_{2'OMe}pG, and m⁷GpppU_{2'OMe}pU.

58. The method of any one of paragraphs 53-57, wherein the trinucleotide cap comprises a sequence selected from the following sequences: GAG, GCG, GUG, and GGG.
59. The method of paragraph 58, wherein the trinucleotide cap comprises sequence GAG.
60. The method of paragraph 59, wherein the trinucleotide cap comprises GpppA_{2'OMe}pG.
61. The method of any one of paragraphs 40-60, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.
62. The method of any one of paragraphs 40-60, wherein the polynucleotide template includes a 2'-deoxycytidine residue at template position +1.
63. A co-transcriptional capping method for RNA synthesis, the method comprising reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a trinucleotide cap comprising sequence GpppA_{2'OMe}pG, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.
64. The method of any one of paragraph 40-63, wherein the RNA transcript produced, when delivered to cells, optionally in unpurified form, does not stimulate a detectable cytokine response.
65. A composition comprising an in vitro-transcribed (IVT) RNA and a pharmaceutically acceptable excipient, wherein the composition is substantially free of cytokine-inducing RNA contaminant in the absence of post-IVT purification.
66. A composition comprising an in vitro-transcribed (IVT) RNA and a pharmaceutically acceptable excipient, wherein the composition has less than 5% uncapped RNA species.
67. The composition of paragraph 65 or 66, wherein greater than 80%, 85%, or 90% of the IVT RNA include a functional cap.
68. The composition of paragraph 67, wherein greater than 95% of the IVT RNA include a functional cap.
69. The composition of any one of paragraphs 65-68, wherein the IVT RNA is not chemically modified.

70. The composition of any one of paragraphs 65-68, wherein the IVT RNA is chemically modified.
71. The composition of any one of paragraphs 65-70, wherein greater than 95% of the IVT RNA comprises single-stranded full-length transcripts.
72. The composition of any one of paragraphs 65-71, wherein the RNA is produced by a process comprising:
- reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a trinucleotide cap comprising sequence GpppA_{2'}Ome_{3'}pG, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.
73. A T7 RNA polymerase comprising at least one additional C-terminal amino acid relative to wild-type T7 RNA polymerase.
74. The T7 RNA polymerase of paragraph 73 comprising at least two additional C-terminal amino acids.
75. The T7 RNA polymerase of paragraph 74, wherein the at least two additional C-terminal amino acids comprise the same type of amino acid or at least two different types of amino acids.
76. The T7 RNA polymerase of any one of paragraphs 73-75 comprising 1 to 10 additional C-terminal amino acids.
77. The T7 RNA polymerase of paragraph 76 comprising 1 to 5 additional C-terminal amino acids.
78. The T7 RNA polymerase of any one of paragraphs 73-77, wherein the T7 RNA polymerase comprises a C terminus that comprises a FAFAX_n motif, wherein X is any amino acid and n is any integer greater than zero.
79. The T7 RNA polymerase of paragraph 78, wherein X is G or A, and optionally wherein X_n is GG or AA.
80. The T7 RNA polymerase of paragraph 78 or 79, wherein n is 1, 2, 3, 4, or 5.
81. A T7 RNA polymerase comprising a C terminus that comprises a FAFAG motif.

82. A T7 RNA polymerase comprising a C terminus that comprises a XAFAX_n motif, a FXFAX_n motif, FAXAX_n motif, or a FAFXX_n motif, wherein each X is any amino acid and n is any integer greater than zero.
83. The T7 RNA polymerase of any one of paragraphs 73-82, wherein the T7 RNA polymerase comprises at least one substitution at a position corresponding to position S43, G47, R257, or G259 of a wild-type T7 RNA polymerase, and optionally at least one additional amino acid substitution.
84. The T7 RNA polymerase of paragraph 83, wherein the wild-type T7 RNA polymerase comprises an amino acid sequence identified by SEQ ID NO:1.
85. A T7 RNA polymerase comprising an amino acid sequence of SEQ ID NO:99 modified to include at least one substitution selected from G47, S43, R257, and G259, and optionally at least one additional amino acid substitution.
86. A T7 RNA polymerase comprising an amino acid sequence of any one of SEQ ID NOs:100, 294, 296, or 296 modified to include at least one substitution selected from G47, S43, R257, and G259, and optionally at least one additional amino acid substitution.
87. The T7 RNA polymerase of any one of paragraphs 83-86, comprising at least one substitution selected from S43A, G47A, R257A, and G259A.
88. A T7 RNA polymerase comprising an amino acid sequence of any one of SEQ ID NOs: 294-313, wherein x is any amino acid and n is any integer, e.g., between 1 and 5 (e.g., 1, 2, 3, 4, or 5), and optionally wherein the T7 RNA polymerase further comprises at least one additional amino acid substitution.
89. A method of performing an in vitro transcription (IVT) reaction, comprising contacting a DNA template with the RNA polymerase of any one of paragraphs 73-88 in the presence of nucleoside triphosphates and buffer under conditions that result in the production of RNA transcripts.
90. The method of paragraph 89, wherein the RNA produced, when delivered to cells stimulates a cytokine response that is at least 50% lower relative to the dsRNA transcript produced using a WT T7 RNAP.
91. The method of paragraph 90, wherein the RNA is delivered to the cells in unpurified form.
92. The method of any one of paragraphs 89-92, wherein less than 30% of the RNA transcripts produced exhibit 3' heterogeneity.
93. A nucleic acid encoding the RNA polymerase of any one of paragraphs 73-92.
94. A vector comprising the nucleic acid of paragraph 93.

95. A host cell comprising the nucleic acid of paragraph 93 or the vector of paragraph 94.
96. A kit comprising the RNA polymerase of any one of paragraphs 73-92.
97. A composition comprising the RNA polymerase of any one of paragraphs 73-92.
98. A ribonucleic acid (RNA) produced by the method of any one of paragraphs 89-92.
99. The RNA of paragraph 98 formulated in a lipid nanoparticle.
100. The RNA of paragraph 98, wherein the lipid nanoparticle comprises a molar ratio of 20-60% ionizable amino lipid, 5-25% non-cationic lipid, 25-55% sterol, and 0.5-15% PEG-modified lipid.
101. An RNA polymerase comprising at least one additional C-terminal amino acid relative to a corresponding wild-type RNA polymerase, optionally wherein the at least one additional C-terminal amino acid comprises glycine (G) and/or alanine (A).
102. The RNA polymerase of paragraph 101, wherein the RNA polymerase is selected from T7 RNA polymerases, T3 RNA polymerases, and SP6 RNA polymerases.
103. The RNA polymerase of paragraph 101 or 102, wherein the RNA polymerase further comprises at least one additional amino acid substitution, optionally further comprising an amino acid substitution corresponding to an amino acid substitution of SEQ ID NO:1 selected from S43A, G47A, R257A, and G259A.
104. The RNA polymerase of any one of paragraphs 101-104, wherein the RNA polymerase is a T7 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:1 modified to include an amino acid substitution at position 43, 47, 257, and/or 259, optionally wherein the amino acid substitution is alanine (A).
105. The RNA polymerase of any one of paragraphs 101-104, wherein the RNA polymerase is a T3 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:6 modified to include an amino acid substitution at position 44, 48, 258, and/or 260, optionally wherein the amino acid substitution is alanine (A).
106. The RNA polymerase of any one of paragraphs 101-104, wherein the RNA polymerase is a SP6 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:7 modified to include an amino acid substitution at position 15, 19, 230, and/or 232, optionally wherein the amino acid substitution is alanine (A).
107. A T7 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:1 modified to include an amino acid substitution at

position 43, 47, 257, and/or 259, optionally wherein the amino acid substitution is alanine (A).

108. A T3 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, or at least 98% identical to SEQ ID NO:6 modified to include an amino acid substitution at position 44, 48, 258, and/or 260, optionally wherein the amino acid substitution is alanine (A).

109. A SP6 RNA polymerase comprising an amino acid sequence at least 90%, at least 95%, or at least 98% % identical to SEQ ID NO:7 modified to include an amino acid substitution at position 15, 19, 230, and/or 232, optionally wherein the amino acid substitution is alanine (A).

Table 1. RNA Polymerase Sequences

RNA Polymerase	Amino Acid Sequence	SEQ ID NO
T7 RNA Polymerase	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRET VDITYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	1
T3 RNA Polymerase	MNIIENIEKNDFSEIELAAIPFNTLADHYGSALAKEQLALEHESYELGERRF LKMLERQAKAGEIADNAAKPLLATLLPKLTTRIVEWLEEYASKKGRKPSAY APLQLLKPEASAFITLKVILASLTSTNMTTI QAAAGMLGKAIEDEARFGRIR DLEAKHFKKHVEEQLNKRHGQVYKKAQFMQVVEADMIGRGLLGGEAWSSWDKE TTMHVGIRLIEMLIESTGLVELQRHNAGNAGSDHEALQLAQEYVDVLAKRAG ALAGISPMFQPCVVPPPKPWVAITGGGYWANGRRPLALVRTHSKKGLMRYEDV YMPEVYKAVNLAQNTAWKINKKVLAVVNEIVNWKNCPVADIPSLERQELPPK PDDIDTNEAALKEWKKAAAGIYRLDKARVSRRI SLEFMLEQANKFASKKAIW FPYNMDWRGRVYAVPMFNPQGNDMTKGLLTAKGKPIGEEGFYWLKIHGANC AGVDKVPFPERIAFIEKHVDDILACAKDPINNTWWAEQDSPFCFLAFCFEYA GVTHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY GIVAQKVNEILKQDAINGTPNEMITVTDKDTGEISEKLKLGTSSTLAQQWLAY GVTRSVTKRSVMTLAYGSKEFGFRQQVLD D TIQPAIDSGKGLMFTQPNQAAG YMAKLIWDVSVTVVAAVEAMNWLKSAKLLAAEVKDKKTKEILRHRCVHVW TTPDGFPVWQ EYRKPLQKRLDMI FLGQFRLQPTINTLKDSGIDAHKQESGIA PNFVHSQDGSHLRMT VVYAHEKYGIESFALIHDSFGTIPADAGKLFKAVRET MVITYENNDVLADFYQS FADQLHETQLDKMPPLPKKGNLNLQDILKSDFAF	6

SP6 RNA Polymerase	MQDLHAIQLQLEEEMFNGGIRRFEDQQQRQIAAGSESDTAWNRRLLSELIAP MAEGIQAYKEEYEGKKGRAPRALAFLQCVENEVAAYITMKVMDMLNTDATL QAIAMSVAERIEDQVRFSKLEGHAAKYFEKVKKSLKASRTKSYRHAHNVAVV AEKSVAEKDADFDRWEAWPKETQLQIGTTLLEILEGVSFYNGEPVFMAMRT YGGKTIYYLQTSSESVQWISAFKEHVAQLSPAYAPCVIPPRPWRTPFNGGFH TEKVASRIRLVKGNREHVRKLTQKQMPKVYKAINALQNTQWQINKDVLAVIE EVIRLDLGYGVPSFKPLIDKENKPANPVPVEFQHLRGRELKEMLSPEQWQQF INWKGECARLYTAETKRGSKSAAVVRMVGQARKYSAFESIYFVYAMDSRSRV YVQSSTLSPQSNDLGKALLRFTEGRPVNGVEALKWFCINGANLWGWDDKTFD VRVSNVLDEEFQDMCRDIAADPLTFTQWAKADAPYEFLOWCFEYAYQLDLVD EGRADEFRTHLFPVHQDGSCSGIQHYSAMLRDEVGAKAVNLKPSDAPQDIYGA VAQVVIKKNALYMDADATTFTSGSVTLSGTEL RAMASAWDSIGITRSLTKK PVMTLPGYSTRLTCRESVIDYIVDLEEKEAQKAVAEGRGTANKVHPFEDDRQD YLTPGAAYNYMTALIWPSISEVVKAPIVAMKMIRQLARFAAKRNEGLMYTLP TGFILEQKIMATEMLRVRTCLMGDIKMSLQVETDIVDEAAMMGAAAPNFVHG HDASHLILTVCELVDKGVTSLIAVIHDSFGTHADNTLTLRVALKGQMVAMYID GNALQKLLLEEHEVRWMDTGIEVPEQGEFDLNEIMDSEYVFA	7
K11 RNA Polymerase	MNALNIGRNDSEIELAAIPYNI LSEHYGDQAAAREQLALEHEAYELGRQRFL KMLERQVKAGEFADNAAKPLVLT LHPQLTKRIDDWKEEQANARGKKPRAYY PIKHGVASELAVSMGAEVLKEKRGVSSEAIALLTIKVVLGNAHRPLKGHNPA VSSQLGKALEDEARFGRIREQEAAAYFKKNVADQLDKRVGHVYKKAQMVEA DMISKGMLGGDNWASWKTDEQMHVGTKLLELLIEGTGLVEMTKNKMADGSDD VTSMQMVQLAPAFVELLSKRAGALAGISPMHQPCVVPKPKWVETVGGGYWSV GRRPLALVRTHSKKALRRYADVHMPEVYKAVNLAQNTPWKVNKKVLAVVNEI VNWKHCPVGDVPAIEREELPPRPDDIDTNEVARKAWRKEAAAVYRKDKARQS RRCRCEFMVAQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNMTKGS LT LAKGKPIGLDGFYWLKIHGANCAVVDKVPFPERIKFIEENEGNILASAADPL NNTWWTQQDSPFCFLAFCFEYAGVKHHGLNYNCSLPLAFDGS CSGIQHFSAM LRDSIGGRAVNLLP SDTVQDIYKIVADKVNEVLHQHAVNGSQTVVEQIADKE TGEFHEKVTLGESVLAAQWLQYGVTRKVTKRSVMTLAYGSKESLVRQQVLED TIQPAIDNGEGLMFTHPNQAAGYMAKLIWDAVTVTVVAAVEAMNWLKSAAKL LAAEVKDKKTKVLRKRCALHVVTPDGFVPWQEQYRKQONQARLKLVLGQANV KMTYNTGKDSEIDA HKQESGIAPNFVHSQDGSHLRMTVVHANEVYGIDSFAL IHDSSGTIPADAGNLFKAVRETMVKTYEDNDVIADFYDQFADQLHESQLDKM PAVPAKGDLNLRDILESDFAFAX	8

Table 2. RNA Polymerase C-Terminal Variant Sequences

RNA Polymerase	Amino Acid Sequence	SEQ ID NO
T7 RNA Polymerase C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQ LNKRVGHVYKKAQMVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQNA GVVGDSETIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTKGIELLRKCAVHV TPDGFVPWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGDLNLRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	99
T7 RNA	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR	100

RNA Polymerase	Amino Acid Sequence	SEQ ID NO
Polymerase C-Terminal Variant (additional G)	KMFERQLKAGEVADNAAKPLITTTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLT SADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHF KKNVVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TI ELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEAL TAWKRAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKI HGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIA NPFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAA NLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG	
T3 RNA Polymerase C-Terminal Variant	MNIIENIEKND FSEIELAAIPFNTLADHYGSALAKEQLALEHESYELGERRF LKMLERQAKAGEIADNAAKPLLATLLPKLTTRIVEWLEEYASKKGRKPSAY APLQLLKPEASAFITLKVILASLTSTNMTTIQAAAGMLGKAIEDEARFGRIR DLEAKHF KKHVVEEQ LNKRHGQVYKKA FMQVVEADMI GRGLLGGEAWSSWDKE TTMHVGIRLIEMLIESTGLVELQRHNAGNAGSDHEALQLAQEYVDVLAKRAG ALAGISPMFQPCVVPPKPWVAITGGGYWANGRRPLALVRTHSKKGLMRYEDV YMPEVYKAVNLAQNTAWKINKKVLAVVNEIVNWKNCPVADIPSLERQELPPK PDDIDTNEAALKEWK KAAAGIYRLDKARVSRRI SLEFMLEQANKFASKKAIW FPYNMDWRGRVYAVPMFNPQGNDMTKGLLTLAKGKPIGEEGFYWLKI HGANCA AGVDKVPFFPERIAFIEKHVDDILACAKDPINNTWWAEQDSPFCFLAFCFEYA GVTHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY GIVAQKVNEILKQDAINGTPNEMITVTDKDTGEISEKLKLGTS TLAQQWLAY GVTRSVTKRSVMTLAYGSKEFGFRQQVLD D TIQPAIDSGKGLMFTQPNQAAG YMAKLIWD AVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTKEILRHRC AVHW TTPDGFVPVWQEYRKPLQKRLDMI FLGQFRLQPTINTLKDSGIDA HKQESGIA PNFVHSQDGSHLRMT VVYAHEKYGIESFALIHDSFGTI PADAGKLFKAVRET MVITYENNDVLADFY SQFADQLHETQLDKMPPLPKKGNLNLQDILKSDFAFA X_n , where X is any amino acid and n is any integer, e.g., between 1 and 5	101
T3 RNA Polymerase C-Terminal Variant	MNIIENIEKND FSEIELAAIPFNTLADHYGSALAKEQLALEHESYELGERRF LKMLERQAKAGEIADNAAKPLLATLLPKLTTRIVEWLEEYASKKGRKPSAY APLQLLKPEASAFITLKVILASLTSTNMTTIQAAAGMLGKAIEDEARFGRIR DLEAKHF KKHVVEEQ LNKRHGQVYKKA FMQVVEADMI GRGLLGGEAWSSWDKE TTMHVGIRLIEMLIESTGLVELQRHNAGNAGSDHEALQLAQEYVDVLAKRAG ALAGISPMFQPCVVPPKPWVAITGGGYWANGRRPLALVRTHSKKGLMRYEDV YMPEVYKAVNLAQNTAWKINKKVLAVVNEIVNWKNCPVADIPSLERQELPPK PDDIDTNEAALKEWK KAAAGIYRLDKARVSRRI SLEFMLEQANKFASKKAIW FPYNMDWRGRVYAVPMFNPQGNDMTKGLLTLAKGKPIGEEGFYWLKI HGANCA AGVDKVPFFPERIAFIEKHVDDILACAKDPINNTWWAEQDSPFCFLAFCFEYA GVTHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY GIVAQKVNEILKQDAINGTPNEMITVTDKDTGEISEKLKLGTS TLAQQWLAY GVTRSVTKRSVMTLAYGSKEFGFRQQVLD D TIQPAIDSGKGLMFTQPNQAAG YMAKLIWD AVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTKEILRHRC AVHW TTPDGFVPVWQEYRKPLQKRLDMI FLGQFRLQPTINTLKDSGIDA HKQESGIA PNFVHSQDGSHLRMT VVYAHEKYGIESFALIHDSFGTI PADAGKLFKAVRET MVITYENNDVLADFY SQFADQLHETQLDKMPPLPKKGNLNLQDILKSDFAFA G	102

SP6 RNA Polymerase C-Terminal Variant	MQDLHAIQLQLEEEMFNGGIRRFEDQQQRQIAAGSESDTAWNRRLLSELIAP MAEGIQAYKEEYEGKKGRAPRALAFLQCVENEVAAYITMKVMDMLNTDATL QAIAMSVAERIEDQVRFSKLEGHAAKYFEKVKKSLKASRTKSYRHAHNVAVV AEKSVAEKDADFDRWEAWPKETQLQIGTTLLEILEGVSFYNGEPVFMRAMRT YGGKTIYYLQTSSESVGWISAFKEHVAQLSPAYAPCVIPPRPWRTPFNGGFH TEKVASRIRLVKGNREHVRKLTQKQMPKVYKAINALQNTQWQINKDVLAVIE EVIRLDLGYGVPSFKPLIDKENKPANPVPVEFQHLRGRELKEMLSPEQWQQF INWKGEARLYTAETKRGSKSAAVVRMVGQARKYSAFESIYFVYAMDSRSRV YVQSSTLSPOQNDLGKALLRFTEGRPVNGVEALKWFCINGANLWGWDDKTFD VRVSNVLDEEFQDMCRDIAADPLTFTQWAKADAPYEFWCFEYAYQYLDLVD EGRADEFRTHLVPVHQDGS CSGIQHYSAMLRDEVGAKAVNLKPSDAPQDIYGA VAQVVIKKNALYMDADDATTTFTSGSVTLSGTEL RAMASAWDSIGITRSLTKK PVMITLPYGSTRLTCRESVIDYIVDLEEKEAQKAVAEGRGTANKVHPFEDDRQD YLTPGAAYNYMTALIWPSISEVVKAPIVAMKMIRQLARFAAKRNEGLMYTLP TGFILEQKIMATEMLRVRTCLMGDIKMSLQVETDIVDEAAMMGAAAPNFVHG HDASHLILTVCELVDKGVTIAVIHDSFGTHADNTLTLRVALKGQMVAMYID GNALQKLEEEHEVRWMDTGIEVPEQGEFDLNEIMDSEYVFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	103
SP6 RNA Polymerase C-Terminal Variant	MQDLHAIQLQLEEEMFNGGIRRFEDQQQRQIAAGSESDTAWNRRLLSELIAP MAEGIQAYKEEYEGKKGRAPRALAFLQCVENEVAAYITMKVMDMLNTDATL QAIAMSVAERIEDQVRFSKLEGHAAKYFEKVKKSLKASRTKSYRHAHNVAVV AEKSVAEKDADFDRWEAWPKETQLQIGTTLLEILEGVSFYNGEPVFMRAMRT YGGKTIYYLQTSSESVGWISAFKEHVAQLSPAYAPCVIPPRPWRTPFNGGFH TEKVASRIRLVKGNREHVRKLTQKQMPKVYKAINALQNTQWQINKDVLAVIE EVIRLDLGYGVPSFKPLIDKENKPANPVPVEFQHLRGRELKEMLSPEQWQQF INWKGEARLYTAETKRGSKSAAVVRMVGQARKYSAFESIYFVYAMDSRSRV YVQSSTLSPOQNDLGKALLRFTEGRPVNGVEALKWFCINGANLWGWDDKTFD VRVSNVLDEEFQDMCRDIAADPLTFTQWAKADAPYEFWCFEYAYQYLDLVD EGRADEFRTHLVPVHQDGS CSGIQHYSAMLRDEVGAKAVNLKPSDAPQDIYGA VAQVVIKKNALYMDADDATTTFTSGSVTLSGTEL RAMASAWDSIGITRSLTKK PVMITLPYGSTRLTCRESVIDYIVDLEEKEAQKAVAEGRGTANKVHPFEDDRQD YLTPGAAYNYMTALIWPSISEVVKAPIVAMKMIRQLARFAAKRNEGLMYTLP TGFILEQKIMATEMLRVRTCLMGDIKMSLQVETDIVDEAAMMGAAAPNFVHG HDASHLILTVCELVDKGVTIAVIHDSFGTHADNTLTLRVALKGQMVAMYID GNALQKLEEEHEVRWMDTGIEVPEQGEFDLNEIMDSEYVFAG	104
K11 RNA Polymerase C-Terminal Variant	MNALNIGRDNDFSEIELAAIPYNILSEHYGDQAAREQLALEHEAYELGRQRFL KMLERQVKAGEFADNAAKPLVLTLPQLTKRIDDWKEEQANARGKKPRAYY PIKHGVASELAVSMGAEVLKEKRGVSSEAIALLTIKVV LGNAHRPLKGNHPA VSSQLGKALEDEARFGRIREQEAAAYFKKNVADQLDKRVGHVYKKAQVVEA DMI SKGMLGGDNWASWKTDEQMHVGT KLLELLIEGTGLVEMTKNKMADGSDD VTSMQMVQLAPAFVELLSKRAGALAGISPMHQPCVVPKPPWVETVGGGYWSV GRRPLALVRTHSKKALRRYADVHMPEVYKAVNLAQNTPWKNKKVLAVVNEI VNWKHCPVGDVPAIEREELPPRPDDIDTNEVARKAWRKEAAAVYRKDKARQS RRCRCFEMVAQANKFANHKAIFWFPYNMDWRGRVYAVSMFNPQNDMTKGS LT LAKGKPIGLDGFYWLKIHGANCAGVDKVPFPERIKFIEENEGNILASAADPL NNTWWTQQDSPFCFLAFCFEYAGVKHHGLNYNCSLPLAFDGS CSGIQHFSAM LRDSIGGRAVNLLPSDTVQDIYKIVADKVNEVLHQHAVNGSQTVVEQIADKE TGEFHEKVTLGESVLAAQWLQYGVTRKVTKRSMVTLAYGSKESLVRQQVLED TIQPAIDNNEGGLMFTHPNQAAGYMAKLIWDAVTVTVAAVEAMNWLKSAAKL LAAEVKDKKTKEVLRKRCIAIHWVTPDGFVPWQEYRKQONQARLKLVLGQANV KMTYNTGKDSEIDAHKQESGIAPNFVHSQDGS HLRMTVVHANEVYGIDSFAL IHDSSGTIPADAGNLFKAVRETVMKTYEDNDVIADFYDQFADQLHESQLDKM PAVPAKGDNLNRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	105
K11 RNA Polymerase C-Terminal Variant	MNALNIGRDNDFSEIELAAIPYNILSEHYGDQAAREQLALEHEAYELGRQRFL KMLERQVKAGEFADNAAKPLVLTLPQLTKRIDDWKEEQANARGKKPRAYY PIKHGVASELAVSMGAEVLKEKRGVSSEAIALLTIKVV LGNAHRPLKGNHPA VSSQLGKALEDEARFGRIREQEAAAYFKKNVADQLDKRVGHVYKKAQVVEA DMI SKGMLGGDNWASWKTDEQMHVGT KLLELLIEGTGLVEMTKNKMADGSDD VTSMQMVQLAPAFVELLSKRAGALAGISPMHQPCVVPKPPWVETVGGGYWSV	106

RNA Polymerase	Amino Acid Sequence	SEQ ID NO
	GRRPLALVRTHSKKALRRYADVHMPVEVYKAVNLAQNTPWKVNKKVLAVVNEI VNWKHCVPVGDVPAIEREELPPRPDDIDTNEVARKAWRKEAAAVYRKDKARQS RRCRCFEMVAQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQNDMTKGS LT LAKGKPIGLDGFYWLKIHGANCAGVDKVPFPERIKFIEENEGNILASAADPL NNTWWTQQDSPFCFLAFCFEYAGVKHHGLNYNCSLPLAFDGS CSGIQHFSAM LRDSIGGRAVNLLPSETVQDIYKIVADKVNELVHQA VNGSQT VVEQIADKE TGEFHEKVTLGESVLAAQWLQYGVTRKVTKR SVMTLAYGS KESLVRQQVLED TIQPAIDNGEGLMFTHPNQAGYMAKLIWDAVTVTVA AVEAMNWLKSAAKL LAAEVKDKKTKEVLRKRC AIHWVTPDGF PVWQ EYRKQNQARLKL VFLGQANV KMTYNTGKDSEIDAHKQESGIAPNFVHSQDGS HLRMTVVHANEVY GIDSFAL IHDS SGTIPADAGNLFKAVRETMVKTYEDNDVIADFYDQFADQLHESQLDKM PAVPAKGDNLNRDILESDFAFAG	

EXAMPLES

Example 1. Production of variant T7 RNA polymerases.

C-helix and C-linker T7 RNA polymerase variants were generated with the substitutions shown in Tables 3 and 4.

Table 3. T7 RNA Polymerase Variants

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
S43A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEAAYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIW F PYNMDWRGRVYAVSMFNPQNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKS PLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGF PVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNRDILESDFAFAG	2
G47A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMAEARFR KMFERQLKAGEVADNAAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIW F PYNMDWRGRVYAVSMFNPQNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKS PLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV	3

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
R257A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLT SADNTTVQ AVASAI GRAI ED EARFGR IRD LEAKHF KKNV EEQLNKR VGHVYK KAFMQVVEADMLS SKGL LGGEAWSSWHKED SIHVG VRCI EMLIESTGMVSLHRQ NAGVVG QDSE TIELA PEYAE AIAT AAGA LAGIS PMFQ PCV VPPK PWTG ITGGGY WANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREEL PMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFM LEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGL LLTLAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQD SPFCFLA FCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQH FSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAID SGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMN NWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	4
G259A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLT SADNTTVQ AVASAI GRAI ED EARFGR IRD LEAKHF KKNV EEQLNKR VGHVYK KAFMQVVEADMLS SKGL LGGEAWSSWHKED SIHVG VRCI EMLIESTGMVSLHRQ NAGVVG QDSE TIELA PEYAE AIAT RAAA LAGIS PMFQ PCV VPPK PWTG ITGGGY WANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREEL PMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFM LEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGL LLTLAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQD SPFCFLA FCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQH FSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAID SGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMN NWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	5
E42A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHASYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLT SADNTTVQ AVASAI GRAI ED EARFGR IRD LEAKHF KKNV EEQLNKR VGHVYK KAFMQVVEADMLS SKGL LGGEAWSSWHKED SIHVG VRCI EMLIESTGMVSLHRQ NAGVVG QDSE TIELA PEYAE AIAT TRAGA LAGIS PMFQ PCV VPPK PWTG ITGGGY WANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREEL PMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFM LEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGL LLTLAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQD SPFCFLA FCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQH FSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAID SGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMN NWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	9
E42I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHISYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLT SADNTTVQ AVASAI GRAI ED EARFGR IRD LEAKHF KKNV EEQLNKR VGHVYK KAFMQVVEADMLS SKGL LGGEAWSSWHKED	10

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
E42L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHLSYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	11
E42M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHMSYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	12
E42K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHKS YEMGEARFR KMFERQLKAGEVADNAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG	13

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSTFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
E42Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHQS YEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSTFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	14
S43I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEIYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSTFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	15
S43L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHELYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSTFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	16
S43M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEMYEMGEARFR	17

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	
S43K	MNTINIAKNDFS DIELAAIPFNTLADHYGERLAREQLALEHEKYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	18
S43Q	MNTINIAKNDFS DIELAAIPFNTLADHYGERLAREQLALEHEQYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	19
S43E	MNTINIAKNDFS DIELAAIPFNTLADHYGERLAREQLALEHEEYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF	20

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	
Y44A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESAEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	21
Y44I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESIEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	22
Y44L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESLEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP	23

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
Y44M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESMEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVQDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	24
Y44K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESKEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVQDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	25
Y44Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESQEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVQDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	26
Y44E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESEEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA	27

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
E45A	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYAMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	28
E45I	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYIMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	29
E45L	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYLMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG	30

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
E45M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYMMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	31
E45K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYKMGGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	32
E45Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYQMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	33
M46A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEAGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ	34

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	
M46I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEIGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	35
M46L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYELGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	36
M46K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEKG EARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA	38

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
M46Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEQGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	39
M46E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEEGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	40
G47I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMIEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM	41

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	
G47L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMLEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	42
G47M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMMEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	43
G47K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMKEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	44
G47Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMQEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY	45

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF A	
G47E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMEEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF A	46
A255I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIIT IAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF A	315
A255K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIKTIAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY	316

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
A255Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIQTIAGA LAGISPMFQPCVVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	317
A255Y	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIYTIAGA LAGISPMFQPCVVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	318
R257I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATIAGA LAGISPMFQPCVVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	47
R257L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD	48

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATLAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
R257M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATMAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	49
R257K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATKAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	50
R257Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATQAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG	51

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	
R257E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATEAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	52
A258I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRIGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	53
A258L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRIGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	54

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
A258M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATR MGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRRCAVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF A	55
A258K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATR KGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRRCAVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF A	56
A258Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRQGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRRCAVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF A	57
A258E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATR EGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP	58

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLNTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
G259I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQLNKR VGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAIA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLNTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	59
G259L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQLNKR VGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRALA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLNTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	60
G259M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQLNKR VGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAMA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLNTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW	61

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
G259K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAKA LAGISPMFQPCVVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	62
G259Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAQA LAGISPMFQPCVVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	63
G259E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAEA LAGISPMFQPCVVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	64
A260I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED	65

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGI LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
A260L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGL LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	66
A260M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGM LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	67
A260K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGK LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG	68

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	
A260Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGQ LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	69
A260E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGE LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	70
L261A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA AAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAF	71
L261I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR	72

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA IAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	
L261M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA MAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	73
L261K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA KAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAF A	74
L261Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA QAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF	75

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
L261E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA EAGISPMFQPCVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	76
A262I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LIGISPMFQPCVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	77
A262L	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LLGISPMFQPCVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP	78

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
A262M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LMGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	79
A262K	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LKGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	80
A262Q	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LQGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	81
A262E	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA	82

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	LEGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
S43R	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESEYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	83
E45R	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	84
G47R	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG	85

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
R257W	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAVVGQDSEITIELAPEYAEAIATWAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	86
G259R	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAVVGQDSEITIELAPEYAEAIATRAA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	87
A260R	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAVVGQDSEITIELAPEYAEAIATRAGR LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	88
N165W	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ	89

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	FLQEIKP EAVAYITIKTTLACLT SADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAF A	
E167M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKP EAVAYITIKTTLACLT SADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVMEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAF A	90
E167N	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKP EAVAYITIKTTLACLT SADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVNEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAF A	91
E168I	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKP EAVAYITIKTTLACLT SADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVNEIQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA	92

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	
E168T	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVE<u>T</u>QLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	93
E168V	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVE<u>V</u>QLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFafa	94
A181F	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKK<u>F</u>FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM	95

T7 RNA Polymerase Variants	Amino Acid Sequence	SEQ ID NO
	VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	
A181W	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKWFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	96
Q184M	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	97
E187F	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVFADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAF	98

Table 4. T7 RNA Polymerase C-Terminal Variants

T7 RNA Polymerase C-Terminal Variant	Amino Acid Sequence	SEQ ID NO
S43A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHE<u>A</u>YEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	107
S43A C-Terminal Variant (S43A C-terminal G Variant or S43A*)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHE<u>A</u>YEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	108
G47A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEM<u>A</u>EARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX _n, where X is any amino acid and n is any integer,	109

	e.g., between 1 and 5	
G47A C-Terminal Variant (G47A C-terminal G Variant or G47A*)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMAEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRCAVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	110
R257A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATAAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRCAVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	111
R257A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATAAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRCAVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	112
G259A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP	113

	EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
G259A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAA A LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	114
E42A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHAS YEMGEARFR KMFERQLKAGEVADNAAAKPLITTTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	115
E42A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHAS YEMGEARFR KMFERQLKAGEVADNAAAKPLITTTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY	116

	MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
E42I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHISYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	117
E42I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHISYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	118
E42L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHLSYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	119
E42L C-Terminal	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHLSYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ	120

Variant	FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	
E42M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHMSYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	121
E42M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHMSYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	122
E42K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHKS YEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG	123

	VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
E42K C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHKS YEMGEARFR KMFERQLKAGEVADNAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKR VGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKI HGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	124
E42Q C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHQS YEMGEARFR KMFERQLKAGEVADNAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKR VGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKI HGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	125
E42Q C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHQS YEMGEARFR KMFERQLKAGEVADNAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKR VGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKI HGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM	126

	VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	
S43I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEIYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	127
S43I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEIYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	128
S43L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHELYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	129
S43L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHELYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSE TIELAPEYAEAIATRAGA	130

	LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
S43M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEMYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	131
S43M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEMYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	132
S43K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEKYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLTPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY	133

	MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
S43K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEKYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAG	134
S43Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEQYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	135
S43Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEQYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAG	136
S43E C-Terminal	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEEYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ	137

Variant	FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLI ESTGMVSLHRQ NAGVVGQDSE TI ELAPEYAEAIATRAGA LAGIS PMFQPCVVP PKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAA VYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSK EFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGF PVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQD GSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
S43E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEEYEMGEARFR KMFERQLKAGEVADNAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLI ESTGMVSLHRQ NAGVVGQDSE TI ELAPEYAEAIATRAGA LAGIS PMFQPCVVP PKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAA VYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSK EFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGF PVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQD GSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAFAG	138
Y44A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESAEMGEARFR KMFERQLKAGEVADNAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLI ESTGMVSLHRQ NAGVVGQDSE TI ELAPEYAEAIATRAGA LAGIS PMFQPCVVP PKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAA VYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSK EFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGF PVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQD GSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	139
Y44A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESAEMGEARFR KMFERQLKAGEVADNAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVR CIEMLI ESTGMVSLHRQ NAGVVGQDSE TI ELAPEYAEAIATRAGA LAGIS PMFQPCVVP PKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAA VYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSK EFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGF PVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQD GSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNLRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	140

	PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAG	
Y44I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESIEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAG_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	141
Y44I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESIEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAG	142
Y44L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESIEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM	143

	VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
Y44L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESLEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANC A GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNL LPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	144
Y44M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESMEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANC A GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNL LPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	145
Y44M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESMEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANC A GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNL LPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	146
Y44K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESKEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA	147

	LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
Y44K C-Terminal Variant	MNTINIAKND FSDIELAAIPFNTLADHYGERLAREQLALEHESKEMGEARFR KMFERQLKAGEVADNAAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	148
Y44Q C-Terminal Variant	MNTINIAKND FSDIELAAIPFNTLADHYGERLAREQLALEHESQEMGEARFR KMFERQLKAGEVADNAAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	149
Y44Q C-Terminal Variant	MNTINIAKND FSDIELAAIPFNTLADHYGERLAREQLALEHESQEMGEARFR KMFERQLKAGEVADNAAAKPLITTL LPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG	150

	IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	
Y44E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESEEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAA VYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	151
Y44E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESEEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAA VYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	152
E45A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYAMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAA VYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	153

E45A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYAMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	154
E45I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYIMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	155
E45I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYIMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	156
E45L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYLMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF	157

	PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
E45L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYLMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEAREFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	158
E45M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYMMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEAREFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	159
E45M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYMMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEAREFGRI RD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV	160

	TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
E45K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYKMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	161
E45K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYKMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	162
E45Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYQMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	163
E45Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYQMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD	164

	LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
M46A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEAGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	165
M46A C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEAGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	166
M46I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEAGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG	167

	IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
M46I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEIGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	168
M46L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYELGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	169
M46L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYELGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	170

M46K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEKGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNL RDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	173
M46K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEKGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNL RDILESDFAFAG	174
M46Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEQGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNL RDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	175
M46Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEQGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY	176

	MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
M46E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEEGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG X n, where X is any amino acid and n is any integer, e.g., between 1 and 5	177
M46E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEEGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	178
G47I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMI EARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW	179

	TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	
G47I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMIEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	180
G47L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMLEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	181
G47L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMLEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDI PAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDQDGSRLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	182
G47M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMMEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD	183

	LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
G47M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMMEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	184
G47K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMKEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	185
G47K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMKEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRVNNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	186

	GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	
G47Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMQEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	187
G47Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMQEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	188
G47E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMEEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	189

	n , where X is any amino acid and n is any integer, e.g., between 1 and 5	
G47E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMEEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG	190
R257I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATIAGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG $\mathbf{X}$ n , where X is any amino acid and n is any integer, e.g., between 1 and 5	191
R257I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATIAGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG	192
R257L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATLAGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY	193

	MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACA KSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
R257L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATLAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACA KSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	194
R257M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATMAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACA KSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	195
R257M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATMAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACA KSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	196

	VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSTGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
R257K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATKAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSTGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	197
R257K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATKAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSTGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	198
R257Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATQAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSTGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	199
R257Q C-	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR	200

Terminal Variant	KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATQAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
R257E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATEAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	201
R257E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATEAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	202
A258I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRIGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA	203

	GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
A258I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAFFMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRLGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANC GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	204
A258L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAFFMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRLGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANC GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	205
A258L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAFFMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRLGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANC GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRRRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP	206

	NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
A258M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATR MGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGC SGIQHFSAMLRDE VGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	207
A258M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATR MGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGC SGIQHFSAMLRDE VGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	208
A258K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATR MGA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGC SGIQHFSAMLRDE VGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	209
A258K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED	210

	SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRKA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
A258Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVEEQNLKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRQGA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	211
A258Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVEEQNLKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRQGA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	212
A258E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVEEQNLKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATREGA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG	213

	VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
A258E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAVVGQDSEITIELAPEYAEAIATREGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	214
G259I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAVVGQDSEITIELAPEYAEAIATRAIA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	215
G259I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAVVGQDSEITIELAPEYAEAIATRAIA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	216
G259L C-	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR	217

Terminal Variant	KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRALA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
G259L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRALA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	218
G259M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAMA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	219
G259M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAMA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	220

	EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQW LAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG	
G259K C- Terminal Variant	MNTINIAKNDFS DIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLIT TLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQNLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAKA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQW LAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	221
G259K C- Terminal Variant	MNTINIAKNDFS DIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLIT TLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQNLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAKA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQW LAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGS HLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG	222
G259Q C- Terminal Variant	MNTINIAKNDFS DIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLIT TLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNV EEQNLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAQA LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQW LAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP	223

	NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	
G259Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSDANTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAQA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	224
G259E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSDANTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAEA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	225
G259E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSDANTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAEA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	226
A260I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSDANTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED	227

	SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGI LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
A260I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGI LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	228
A260L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGL LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	229
A260L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGL LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	230

	VQHHGLSYNCSLPLAFDGGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	
A260M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGM LAGISPMFQPCVVP PKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	231
A260M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGM LAGISPMFQPCVVP PKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	232
A260K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGK LAGISPMFQPCVVP PKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG n, where X is any amino acid and n is any integer,	233

	e.g., between 1 and 5	
A260K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATRAGK LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRRCAVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	234
A260Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATRAGQ LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRRCAVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	235
A260Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATRAGQ LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRRCAVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	236
A260E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAGVVGDSETIELAPEYAEAIATRAGE LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP	237

	EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
A260E C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKR VGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGE LAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	238
L261A C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKR VGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA AAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
L261A C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKNVEEQLNKR VGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQDSE TIELAPEYAEAIATRAGA AAGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKI HGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY G IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY	239

	MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
L261I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA IAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	240
L261I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA IAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	241
L261M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA MAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	242
L261M C-Terminal	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ	243

Variant	FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA MAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	
L261K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA KAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	244
L261K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA KAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	245
L261Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA QAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG	246

	VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
L261Q C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA QAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	247
L261E C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA EAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	248
L261E C- Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA EAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHD SFGTIPADAANLFKAVRETM	249

	VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	
A262I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVQDSE TIELAPEYAEAIATRAGA LIGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	250
A262I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVQDSE TIELAPEYAEAIATRAGA LIGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	251
A262L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVQDSE TIELAPEYAEAIATRAGA LLGISPMFQPCVVPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	252
A262L C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVQDSE TIELAPEYAEAIATRAGA	253

	LLGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFBVHSQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
A262M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LMGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFBVHSQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	254
A262M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LMGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFBVHSQDGSRLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	255
A262K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVVEEQLNKRVGHVYKKAQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LKGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY	256

	MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
A262K C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSETIELAPEYAEAIATRAGA LKGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAG	257
A262Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSETIELAPEYAEAIATRAGA LQGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	258
A262Q C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSETIELAPEYAEAIATRAGA LQGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFAG	259
A262E C-Terminal	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ	260

Variant	FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LEGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
A262E C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LEGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	261
S43R C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHERYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	262
S43R C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHERYEMGEARFR KMFERQLKAGEVADNAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF	263

	PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
E45R C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYRMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	264
E45R C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYRMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	265
G47R C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMREARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM	266

	VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
G47R C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMREARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNL LPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLRDILESDFAFAG	267
R257W C-terminal variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATWAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNL LPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLRDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	268
R257W C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATWAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRVNL LPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEI SEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLRDILESDFAFAG	269
G259R C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRI RD LEAKHFKKVVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATRARA	270

	LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
G259R C-Terminal Variant	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVEEQNLNKRGVHVKKAFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRA <u>RA</u> LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	271
A260R C-Terminal Variant	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVEEQNLNKRGVHVKKAFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRA <u>GR</u> LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	272
A260R C-Terminal Variant	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKNVEEQNLNKRGVHVKKAFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRA <u>GR</u> LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPPVWQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	273

	IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
N165W C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKWVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYVRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIYIG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	274
N165W C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKWVVEEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAAL AGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYM PEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKPE DIDMNPEALTAWKRAAAAAYVRKDKARKSRRI SLEFMLEQANKFANHKAIFW YNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWYWLKIHGANCA VDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGV QHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIYIGI VAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYGV TRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGYM AKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWV PDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAPN FVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETMV DTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	275
E167M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKEPAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVMEQLNKRVGHVYKKAQFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQONAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAAYVRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTLAKGKPIGKEGYWYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGS CSGIQHFSAMLRDEVGGRVNLPLSETVQDIYIG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVWV TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	276

E167M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHF KKNVMEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	277
E167N C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHF KKNVNEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX _n, where X is any amino acid and n is any integer, e.g., between 1 and 5	278
E167N C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHF KKNVNEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRKRC AVHWV TPDGFVPVWQ EYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	279
E168I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAI EDEARFGRIRD LEAKHF KKNVIEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF	280

	PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
E168I C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEAREFGRI RD LEAKHFKKNVEI¹QLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	281
E168T C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEAREFGRI RD LEAKHFKKNVEI²QLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAX_n, where X is any amino acid and n is any integer, e.g., between 1 and 5	282
E168T C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEAREFGRI RD LEAKHFKKNVEI³QLNKRVGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHWV	283

	TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
E168V C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEV QLNKR VGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQD SETIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVR THSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLTLAKGKPIGKEGYWLKIH GANCA GVDKVPFPERIKFIEENHENIMACA KSPLENTWWAEQDSPFCFLAFCFEYAG VQHHLGSLYNCSLPLAFD GSCSGIQHFSAMLRDEVG GRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVD TENTGEISEKVKLG TKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQ QVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	284
E168V C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEV QLNKR VGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQD SETIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVR THSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLTLAKGKPIGKEGYWLKIH GANCA GVDKVPFPERIKFIEENHENIMACA KSPLENTWWAEQDSPFCFLAFCFEYAG VQHHLGSLYNCSLPLAFD GSCSGIQHFSAMLRDEVG GRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVD TENTGEISEKVKLG TKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQ QVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	285
A181F C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHF KKNVEE QLNKR VGHVYKKA FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQ NAGVVGQD SETIELAPEYAEAIATRAGA LAGISPMFQPCVVPK PWTGITGGGYWANGRRPLALVR THSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGN DMTKGLLTLAKGKPIGKEGYWLKIH GANCA GVDKVPFPERIKFIEENHENIMACA KSPLENTWWAEQDSPFCFLAFCFEYAG VQHHLGSLYNCSLPLAFD GSCSGIQHFSAMLRDEVG GRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVD TENTGEISEKVKLG TKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQ QVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFPVWQ EYKKPIQ TRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG n, where X is any amino acid and n is any integer, e.g., between 1 and 5	286
A181F C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIK TTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD	287

	LEAKHFKKNVEEQLNKRVGHVYKK F FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	
A181W C-Terminal Variant	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKK W FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG X _n , where X is any amino acid and n is any integer, e.g., between 1 and 5	288
A181W C-Terminal Variant	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKK W FMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDCKTGEILRKRCVHVW TPDGFVPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFDYDQFADQLHESQLDKMPALPAKGNLNLRLDILESDFAFAG	289
Q184M C-Terminal Variant	MNTINIAKNDFSDI ELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTL PKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVEEQLNKRVGHVYKK AFM VEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVPPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNDMTKGLLLTLAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG	290

	IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	
Q184M C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFM ^M VVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	291
E187F C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVV ^F ADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAX n, where X is any amino acid and n is any integer, e.g., between 1 and 5	292
E187F C-Terminal Variant	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVV ^F ADMLSKGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQNAAGVVGQDSEITIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIY IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFPVWQEQYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG	293

T7 RNA Polymerase C-Terminal Variant (C-terminal GG)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLT SADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGDSE TI ELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAG G	294
T7 RNA Polymerase C-Terminal (C-terminal A)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLT SADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGDSE TI ELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAA A	295
T7 RNA Polymerase C-Terminal Variant (C-terminal AA)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLT SADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGDSE TI ELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNDMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKFEGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKD KKTGEILRKRC AVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTI PADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKG NNLNRDILESDFAFAA A	296
G47A C-Terminal Variant (C-terminal GG)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFR KMFERQLKAGEVADNAAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLT SADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHF KKNVEEQ LNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCEIMLIESTGMVSLHRQ NAGVVGDSE TI ELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNPEALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKA IWF	297

	PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAG G	
G47A C-Terminal Variant (C-terminal A)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMAEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAAA	298
G47A C-Terminal Variant (C-terminal AA)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMAEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLNRDILESDFAFAAA	299
S43A/G47A	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEAYEMAEARFR KMFERQLKAGEVADNAAAKPLITTLPKMIARINDWFEEVKAKRGKRPTAFQ FLQEI KPEAVAYITIKTTLACLTSADNTTVQAVASAI GRAIEDEARFGRIRD LEAKHFKKNVVEEQLNKRVGHVYKKAFFMQVVEADMLSKGLLGGEAWSSWHKED SIHVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRAGA LAGISPMFQPCVVPKPKWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKP EDIDMNPALTAWKRAAAAVYRKDKARKSRRI SLEFMLEQANKFANHKAIFW PYNMDWRGRVYAVSMFNPQGNMTKGLLTAKGKPIGKEGYWLKIHGANCA GVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFDGCSCGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCVHVW TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAP NFVHSDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM	300

	VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLRDILESDFAFAG	
S43A/G47A C-Terminal Variant (C-terminal G)	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHEAAYEMA E EARFR KMFERQLKAGEVADNAAKPLITTLPLKMIARINDWFEEVKAKRGKRPTAFQ FLQEIKPEAVAYITIKTTLACLT SADNTTVQAVASAI GRAI EDEARFGRI RD LEAKHFKKNVEEQLNKRVGHVYKKA FMQVVEADMLS KGLLGGEAWSSWHKED SIHVGVRCIEMLIESTGMVSLHRQNA GVVGDSE TIELAPEYAEAIATRAGA LAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPA IEREELPMKP EDIDMNP EALTAWKRAAAAVYRKDKARKSRRLSLEFMLEQANKFANHKA IWF PYNMDWRGRVYAVSMFNPQGNMTKGLLT LAKGKPIGKEGYWLKIHGANCA GVDKVPFFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAG VQHHGLSYNCSLPLAFD GSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYG IVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGT KALAGQWLAYG VTRSVTKRSVMTLAYGSKEFGFRQQVLED TIQPAIDSGKGLMFTQPNQAAGY MAKLIWESVSVTVVA AVEAMNWLKSAAKLLAAEVKDKKTGEILRRCAVHWV TPDGFVPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDA HKQESGIAP NFVHSQDGSHLRKT VVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETM VDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNNLRDILESDFAFAG	301

Example 2. The purity of hEPO mRNA produced using the T7 RNA polymerase variants is comparable to the purity of hEPO mRNA produced using a wild-type T7 RNA polymerase.

In vitro transcription reactions were performed using hEPO DNA template and (1) Wild-type (WT) T7 RNA polymerase (WT #1 and WT #2), (2) a G47A C-terminal G T7 RNA polymerase variant (“G47A*”) (SEQ ID NO: 110), and (3) a S43A C-terminal G T7 RNA polymerase variant (“S43A*”) (SEQ ID NO: 108). DBAA (dibutylammonium acetate) HPLC analyses showed that the purity of all hEPO transcripts was approximately the same (FIG. 1).

Example 3. hEPO mRNA produced using the T7 RNA polymerase variants does not elicit a cellular cytokine response.

BJ fibroblasts were transfected with hEPO mRNA produced by the IVT reactions described in Example 2. The mRNA was either unpurified or purified by reverse phase (RP) chromatography. As shown in FIG. 2 (left graph), unpurified mRNA produced using the S43A* T7 RNA polymerase variant or the G47A* T7 RNA polymerase variant was as ‘cold’ (did not elicit a cytokine response) as purified mRNA produced using WT T7 polymerase. Approximately equivalent mRNA expression levels are shown in FIG. 2 (right graph). The mRNA used in these experiments did not contain any nucleotide modifications.

A similar cell transfection assay was repeated using hEPO mRNA containing either m¹ψ nucleotide modifications (FIG. 3, top graph) or mo⁵U nucleotide modifications (FIG. 3, bottom graph). BJ fibroblast cells transfected with unpurified m¹ψ hEPO mRNA produced using the G47A* and S43A* T7 RNA polymerase variants exhibited lower IFNβ responses

than cells transfected with unpurified m¹ψ hEPO mRNA produced using WT T7 RNA polymerase. Approximately equivalent mRNA expression levels are shown in FIG. 4.

Monocyte-derived macrophages (MDMs) were then used to examine cytokine response further. An IP10 to hypoxanthine-guanine phosphoribosyltransferase (HPRT) ratio was measured in MDMs transfected with hEPO mRNA generated under the conditions described in Example 2. MDMs are very sensitive to the unmodified chemistry, so drastic differences between samples were not seen using unmodified hEPO mRNA. However, MDMs transfected with unpurified m¹ψ hEPO mRNA produced using the G47A* and S43A* T7 RNA polymerase variants exhibited lower IP10 responses than MDMs transfected with purified m¹ψ hEPO mRNA produced using WT T7 RNA polymerase (FIG. 5, top graph).

Example 4. The T7 RNA polymerase variants produce mRNA associated with less contaminating dsRNA than the WT T7 RNA polymerase.

A standard dsRNA ELISA was used to assess dsRNA contaminants (e.g., longer than 40 nucleotide base pairs) in IVT reactions used to produce unmodified hEPO mRNA (FIG. 7) or hEPO mRNA containing either m¹ψ nucleotide modifications (FIG. 8, top graph) or mo⁵U nucleotide modifications (FIG. 8, bottom graph). The dsRNA assay demonstrated that the S43A* T7 RNA polymerase variant and the G47A* T7 RNA polymerase variant produced less dsRNA than WT T7 RNA polymerase (FIGS. 7 and 8).

Example 5. The T7 RNA polymerase variants reduce 3' heterogeneity of mRNA.

The 3' heterogeneity of transcripts can be measured using a RNase T1 digest. RNase T1 cleaves mRNA specifically after a G nucleotide. Endonucleolytic cleavage results in a 5' hydroxide (OH) and 3' monophosphate (mP) 'scar'. Thus, a RNase T1 digest can be used to differentiate between transcripts that do and do not have non-templated additions at the 3' end. In this Example, the hEPO mRNA produced in Example 2 terminate with a polyA tail and a XbaI restriction site (e.g. A_nUCUAG). hEPO mRNA produced using WT T7 RNA polymerase, the S43A* T7 RNA polymerase variant, or the G47A* T7 RNA polymerase variant were digested with RNase T1 and analyzed by LCMS to generate oligo fingerprints. A clean 3' end (5' OH/3' OH) was found to peak at 32780 Da, while a 'scar' (5' OH/3' mP) peaks at 32860 Da (FIG. 9A). A 'scar' indicates that the transcript had non-templated additions at the carboxy terminus. hEPO mRNA produced using the T7 RNA polymerase variants, S43A* or G47A*, had a higher 3' end population distribution, indicating that they have cleaner 3' ends and improved 3' end heterogeneity (FIG. 9B).

Example 6. T7 RNA Polymerase Variants Produce ‘Cleaner’ 3' Ends

An assay was used to probe the extent of non-templated additions produced by T7 RNAP variants, G47A* and S43A*. In this assay, ligation leftmers were made using an enzymatic polymerase (see, e.g., FIG. 11A). The leftmers were then annealed to a fully-complimentary DNA splint immediately adjacent to a 5'-monophosphorylated rightmer RNA. A DNA splint is typically 40 nt long and the rightmer is typically decorated with a fluorophore at the 5'-end. DNA ligase I was then added under catalytic conditions and ligation was allowed to proceed for a desired reaction time before the reaction was quenched with EDTA. The mixture was then denatured in 4M urea at 95°C for 5 minutes before being loaded onto a denaturing polyacrylamide gel (PAGE-D) (FIG. 11B). The percentage of acrylamide in the gel depends on the length of the expected construct but is typically 6% acrylamide for leftmers > 50 nt in length and 20% acrylamide for leftmers approaching the length of typical mRNA. 6% acrylamide gels ran for 25 minutes at a constant 180V and 20% acrylamide gels ran for 120 minutes at a constant 180V. The gels were then imaged on a Typhoon or similar scanner using excitation and emission wavelengths appropriate to the fluorophore present on the rightmer and the newly ligated product. Rightmer and ligated products migrate differently on the gel based upon size, and the intensity of the two bands will exactly correspond to the efficiency of the ligation reaction. The greater the extent of untemplated nucleotide addition during the enzymatic synthesis of the leftmer, the lower the expected ligation yield.

Ligation yield was calculated as a ratio of bands corresponding to unligated rightmer and ligated full-length product. The reaction was terminated early to exaggerate differences in reaction yields. Based on these results T7 RNAP variant G47A* incorporated the fewest untemplated nucleotides at the 3'-end of the leftmer RNA.

Example 7. Production of capped mRNAs in a Single In Vitro Transcription Reaction

In a standard in vitro transcription (IVT) assay, T7 RNA polymerase prefers to initiate transcription with a 5'GTP. A single stranded RNA (ssRNA) molecule produced by such standard IVT assays requires a separate enzymatic capping step (FIG. 13). For example, a cap1 RNA (7mGpppN2'-Om-RNA) can be produced in a capping assay using a Vaccinia capping enzyme and a 2'-O-methyl transferase (2'OMTase). Cap1 is typically used for

efficient protein translation and mRNA stability in cells. However, some mRNA sequences are difficult to cap, and capping/methylation enzymes are very expensive.

The present disclosure provides, in some embodiments, methods of capping a ssRNA (e.g., mRNA) with trinucleotides co-transcriptionally in an in vitro transcription assay using a T7 RNA polymerase variant described herein (e.g., T7 RNA polymerase variant G47A*). Efficient co-transcriptional capping typically includes a double-stranded DNA (dsDNA) template that initiates transcription with 5'ATP and equimolar concentrations of NTPs and trinucleotides. Normally, T7 RNA polymerase has severely reduced initiation activity with 5'ATP. However, in the presence of XAG (wherein X is any nucleotide, e.g., 5^mGppp) trinucleotides, the limited initiation activity of T7 RNA polymerase with 5'ATP drives initiation with the trinucleotide rather than 5'ATP, generating capped mRNA products co-transcriptionally.

Exemplary commercially available di- and trinucleotide caps are shown in FIG. 14A. The Vaccinia cap1 is a typical di-nucleotide cap and cannot be added co-transcriptionally. Another di-nucleotide cap that can be added co-transcriptionally is the Anti-Reverse Cap Analog (ARCA), also commercially available (e.g., from Thermo Fisher, Catalog number: AM8045). ARCA is a modified cap analog in which the 3' OH group (closer to m⁷G) is replaced with –OCH₃. ARCA is used as a control in some of the co-transcriptional capping assays described in this study.

In a co-transcriptional capping assay, human EPO (hEPO) DNA template was used. Also present in the assay mixture were 5 mM NTP with or without 5 mM trinucleotides (m⁷GpppAG, m⁷GmpppAG, and ARCA). The unmodified mRNA products were analyzed by RNase H cap assay or LCMS. As shown in Table 5, both wild type (WT) T7 RNA polymerase and the T7 RNA polymerase variant G47A* were able to produce fully capped mRNA with equally high efficiency. RNA yield was comparable among different IVT reactions (FIG. 14B) and mRNA products with high degree of integrity were produced (FIG. 14C).

Table 5. Capping Efficiency of WT T7 RNA polymerase or T7 RNA polymerase variant G47A*

T7 RNA polymerase	Trinucleotide	Capping efficiency (%)
WT	-	0.0
T7 RNA polymerase variant G47A*	-	0.0
WT	m ⁷ GpppAG	93.9

T7 RNA polymerase	Trinucleotide	Capping efficiency (%)
T7 RNA polymerase variant G47A*	m ⁷ GpppAG	100.0
WT	m ⁷ GpppAG	100.0
T7 RNA polymerase variant G47A*	m ⁷ GpppAG	100.0

The immune-stimulatory activity of the mRNA products was assessed by evaluating their ability to induce cytokine (IFN β) production in BJ fibroblasts. Interestingly, WT T7 RNA polymerase produced unmodified mRNA that induced a high cytokine response, while T7 RNA polymerase variant G47A* produced unmodified mRNA that did not stimulate cytokine production in BJ fibroblasts (FIG. 14D), minimizing the need to further purify the RNA after the IVT/capping assay. hEPO expression from the mRNA products were also evaluated. No hEPO expression was observed from mRNA produced by WT T7 RNA polymerase, while mRNA produced by T7 RNA polymerase variant G47A* led to comparable hEPO expression as mRNA modified with 1-methyl-pseudourine capped with Vaccinia cap1 (FIG. 14E). LCMS experiments also shows that T7 RNA polymerase variant G47A* produced cleaner RNA than WT T7 RNA polymerase, though the trinucleotide incorporation efficiency appeared to be equivalent for the two enzymes (FIG. 15).

In a different experiment, the capping efficiency of T7 RNA polymerase variant G47A* was compared with the efficiency achieved in a standard capping assay using the Vaccinia cap1. In this experiment, the reaction mixture contained a hEPO dsDNA template with a deoxythymidine at position +1 on the template strand (also termed “Astart” for the first templated nucleotide), the T7 RNA polymerase variant G47A*, and equimolar (7.5mM) NTPs and GAG trinucleotides. The resulting mRNA products were purified using oligo dT and analyzed for capping efficiency, in vitro cytokine response, and in vitro expression. The results show that T7 RNA polymerase variant G47A* produced capped mRNA with comparably high efficiency as the standard capping process (Table 6), that the mRNA did not induce cytokine response in BJ fibroblasts (FIG. 16A), and that the mRNA lead to high hEPO expression in BJ fibroblasts (FIG. 16B).

Table 6. Capping Efficiency of Co-transcriptional Capping and Standard Process

Cap identity	% functional cap
Uncap	0
GAG (one-pot)	96.0
Vaccinia cap1 (+RP)	99.8

In summary, it was demonstrated herein that the T7 RNA polymerase variant G47A* can efficiently initiate transcription at the 3'dTTP of the template strand in the presence of XAG trinucleotides (e.g., GAG or GmAG). Provided at equimolar concentrations as the NTPs in the assays, the trinucleotides can be added co-transcriptionally to produce fully capped mRNAs that does not stimulate an immune response and results in high protein expression (FIG. 17).

Example 8. Comparison of transcription initiation with 5'GTP (Gstart) or 5'ATP (Astart)

Using the co-transcriptional capping assays described in Example 7, the capping efficiency on model constructs where transcription initiation requires 5'-GTP (Gstart) or a 5'-ATP (Astart) were compared. A model construct encoding a 5'UTR and a 47-mer RNA oligonucleotide was used as template for IVT. The total RNA products were assessed by mass spectrometry and the results show that Astart constructs more efficiently incorporate the GAG or GmAG trinucleotides, compared to Gstart constructs (FIG.18A). Though the data is not shown here, it was also observed that Cstart and Ustart constructs produced low amounts of capped products.

Further, the RNA products were subject to RNase H cleavage and the cleaved 5' ends of the RNA products were analyzed using mass spectrometry to determine the presence of the cap. Similar results were obtained as that of the total RNA products (FIG. 18B).

A trinucleotide titration experiment was also performed using the model construct used in FIGs. 18A and 18B. In the co-transcriptional capping assay, 5 mM of NTPs (A, modified U, G, C at equimolar ratio) were used while the concentration of the trinucleotide (GAG or GmAG) varied to evaluate capping efficiency at different molar ratios of the NTP and the trinucleotide. The crude RNA products were then analyzed by LC-MS. The result shows that capping efficiency is the highest when the NTP and the trinucleotide are at equimolar ratio for both GAG and GmAG (Table 7).

Table 7. Trinucleotide Titration

Trinucleotide concentration (mM)	% Cap (GAG)	% Cap (GmAG)
5 (EQ)	96.0	88.3
4	92.8	85.6
3	90.3	82.8
2	81.7	74.5
1	66.1	57.0
0.5	42.5	48.6

Trinucleotide concentration (mM)	% Cap (GAG)	% Cap (GmAG)
0	0	0

Example 9. Capping of mRNAs chemically modified with 1-methyl-pseudouridine

Co-transcriptional capping of mRNAs chemically modified with 1-methyl-pseudouridine were also assessed. Three model constructs were used: hEPO, Luc (encoding luciferase), and eGFP. All IVT templates were Astart templates. The templates may be PCR fragments or plasmids. The capping assay reaction mixture contained the template, T7 RNA polymerase variant G47A*, 7.5 mM of NTPs, and 7.5 mM of trinucleotides (GAG or GmAG). The UTPs in the NTPs were replaced by 1-methylpseudouridine to produce chemically-modified mRNA. Vaccinia cap1 was used as a production control in a standard capping assay. The mRNA products were analyzed using RNase H cap assay, T1 fingerprinting, and a fragment analyzer to determine their integrity. The abilities of the mRNA products in inducing cytokine response and expressing the encoded protein in BJ fibroblasts were also tested.

As shown in Table 8, the mRNA chemically modified with 1-methyl-pseudouridine for all three model constructs produced from PCR fragment templates were efficiently capped, the capping efficiency was comparable to the vaccinia cap1 control. Chemistry A denotes use of unmodified uridine triphosphate nucleotide and Chemistry B denotes use of 1-methyl-pseudouridine triphosphate nucleotide. Chemistry C denotes 5-methyl-cytidine triphosphate and 1-methyl-pseudouridine triphosphate nucleotides. Similar results were obtained for mRNA chemically modified with 1-methyl-pseudouridine produced from plasmid templates (Table 9). Further, RNase T1 fingerprinting analysis showed that the mRNAs produced from the model constructs using PCR fragment templates were of correct sequence (FIG. 19A). The mRNA produced from either PCR fragment templates (FIG. 19B) or plasmid templates (FIG. 20C) were also shown to have high degree of integrity. The capped mRNAs produced from PCR fragment templates as well as plasmid templates did not induce any cytokine expression in BJ fibroblasts (FIG. 19C, FIG. 20D). Further, mRNAs produced from plasmid templates resulted in expression of the encoded proteins in BJ fibroblasts (FIGs. 20E-20G). The expression levels for GAG-capped mRNAs were slightly higher, comparable or slightly lower than the expression levels from the Vaccinia cap1 control mRNAs.

Table 8. Capping Efficiency of mRNA chemically modified with 1-methyl-pseudouridine - PCR Fragment Templates

Construct	Cap	% Cap (% reported)
hEPO	Uncap	0
hEPO	Cap1 (vaccinia)	99.8 (n/a)
hEPO	GAG (trinucleotide)	96.1
hEPO	GmAG (trinucleotide)	94.9
Luc	Uncap	0
Luc	Cap1 (vaccinia)	85.7 (100%)
Luc	GAG (trinucleotide)	94.4
Luc	GmAG (trinucleotide)	93.3
eGFP	Uncap	0
eGFP	Cap1 (vaccinia)	99.4 (88%)
eGFP	GAG (trinucleotide)	93.8
eGFP	GmAG (trinucleotide)	94.9

Table 9. Capping Efficiency of mRNA chemically modified with 1-methyl-pseudouridine - Plasmid Templates

Construct	Cap	% Cap (% reported)
hEPO	Uncap	0
hEPO	Cap1 (vaccinia)	99.8 (n/a)
hEPO	GAG (trinucleotide)	96.4
hEPO	GmAG (trinucleotide)	96.3
Luc	Uncap	0
Luc	Cap1 (vaccinia)	85.7 (100%)
Luc	GAG (trinucleotide)	95.7
Luc	GmAG (trinucleotide)	99.6
eGFP _{egd}	Uncap	0
eGFP _{deg}	Cap1 (vaccinia)	92.1 (n/a)
eGFP _{deg}	GAG (trinucleotide)	94.3
eGFP _{deg}	GmAG (trinucleotide)	92.1

The NTP consumption in the co-transcriptional capping assay were also assessed for reactions using plasmid templates for model constructs hEPO (FIG. 20A) and eGFP (FIG. 20B). The results show that the trinucleotides GAG and GmAG were hardly consumed throughout the duration of the assay (2 hours). This is consistent with the fact that the concentration of the trinucleotides in the assay was in extreme excess. It may be possible to recover trinucleotides after the assay is completed.

Example 10. Substitutions at the T7 RNA Polymerase C-terminus impact RNA yields

G47A* polymerase contains an extra glycine at the C-terminus (“foot glycine”) of the translated protein. The presence of this foot glycine was confirmed by trypsin digest and mass spectrometry analysis of purified G47A* protein (data not shown). The role of the C-

terminal foot region of G47A* was examined in greater detail because of its proximity to the active site of G47A*.

In vitro transcription (IVT) reactions were performed using hEPO DNA template and (1) a wild-type (WT) T7 RNA polymerase (WT), (2) a WT T7 RNA polymerase with a foot glycine (WT + G), (3) a WT T7 RNA polymerase with two foot glycines (WT + GG), (4) a wild-type RNA polymerase with two foot alanines (WT + AA), (5) a G47A T7 RNA polymerase variant, (6) a G47A* T7 RNA polymerase variant with a foot glycine (G47A + G), (7) a G47A T7 RNA polymerase variant with two foot glycines (G47A + GG), (8) a G47A T7 RNA polymerase variant with two foot alanines (G47A + AA), (9) a S43A/G47A T7 RNA polymerase variant (S43A/G47A), and (10) a S43A/G47A T7 RNA polymerase variant with a foot glycine (S43A/G47A + G). IVT reaction mixtures contained hEPO DNA template, one of the T7 RNA polymerase variants listed above, and NTPs. IVT reactions with unmodified chemistry (FIG. 22) were performed with standard NTPs (ATP, CTP, GTP, UTP), while IVT reactions with pseudouridine-modified (m1ψ) chemistry (FIG. 23) contained 1-methylpseudouridine instead of UTP to produce chemically-modified mRNA. RNA yield was measured by UV absorption.

hEPO mRNA yield decreased with the presence of a foot glycine in either WT or G47A* variant T7 RNA polymerase (FIGs. 22-23). mRNA yield continued to decrease with either two foot glycines or two foot alanine residues, although there was less of a decrease with G47A* than with WT T7 RNA polymerase.

Example 11. Impact of T7 RNA polymerase foot glycine on 3'-mRNA heterogeneity

hEPO mRNA produced using WT or G47A* T7 RNAP with a foot glycine was digested with RNase T1 and analyzed by LCMS to generate oligo fingerprints as in Example 5. WT T7 RNAP IVT reactions were conducted with either equimolar concentrations of NTPs (WT EQ) or with molar excess of GTP and ATP (WT alpha (A)). LCMS analysis revealed a clean 3' end (5' OH/3' OH) peak at 32780 Da, and a 'scar' (5' OH/3' mP) peak at 32860 Da (FIG. 24A). A 'scar' indicates that the transcript had non-templated additions at the 3' end. hEPO mRNA produced using WT T7 RNAP and equimolar concentrations of NTPs (WT EQ) had decreased 3' end population distribution, indicating decreased 3' homogeneity and more run-on products with non-templated additions (FIG. 24B). However, hEPO mRNA produced using G47A* T7 RNAP with a foot glycine had a higher 3' end population distribution, indicating that they have cleaner 3' ends and improved 3' end heterogeneity (FIG. 24B).

To directly examine the effect of the T7 RNAP foot region on hEPO 3' mRNA homogeneity, IVT reactions were performed using hEPO DNA template and (1) a wild-type T7 RNAP (WT), (2) a WT T7 RNAP with a foot glycine (WT + G), (3) a WT T7 RNAP with two foot glycines (WT + GG), (4) a wild-type T7 RNAP with two foot alanines (WT + AA), (5) a G47A T7 RNAP variant, (6) a G47A* T7 RNAP variant with a foot glycine (G47A + G), (7) a G47A T7 RNAP variant with two foot glycines (G47A + GG), (8) a G47A T7 RNAP variant with two foot alanines (G47A + AA), (9) a S43A/G47A T7 RNAP variant (S43A/G47A), and (10) a S43A/G47A T7 RNAP variant with a foot glycine (S43A/G47A + G). IVT reaction mixtures contained hEPO DNA template, one of the T7 RNAP variants listed above, and equimolar NTPs. IVT reactions were conducted which produced either unmodified mRNA (FIG. 24C) or pseudouridine modified mRNA (FIG. 24D). The hEPO mRNA was digested with RNase T1 as in Example 5 and above. The presence of a foot glycine increased hEPO mRNA 3' end homogeneity by 60% (FIGs. 24C-24D). Two foot glycines or two foot alanines in T7 RNAP also improved 3' end homogeneity by 60-70% over WT or G47A T7 RNAPs (FIGs. 24C-24D). These results indicate the presence of a single foot glycine in T7 RNAP can increase the 3' end homogeneity observed in FIG. 24B.

Example 12. Impact of T7 RNAP foot glycine on immune stimulation

The role of T7 RNAP foot glycine in stimulating an immune response to hEPO mRNA products was assessed by evaluating their ability to induce cytokine (IFN β) production in BJ fibroblasts as in Example 7. Interestingly, while WT T7 RNAP-produced unmodified and pseudouridine-modified hEPO mRNA induced a high cytokine response, WT T7 RNAP with a foot glycine-produced unmodified and pseudouridine-modified hEPO mRNA stimulated less cytokine production in BJ fibroblasts (FIG. 25). hEPO mRNA produced by G47A* variant T7 RNAP with a foot glycine failed to stimulate cytokine production in BJ fibroblasts, indicating an additive effect between the foot glycine and G47A T7 RNA polymerase variant in failing to induce an immune response (FIG. 25).

Example 13. Impact of T7 RNA polymerase foot glycine on mRNA production

An assay was developed to probe the effect of a foot glycine residue on hEPO mRNA species produced by WT and G47A variant T7 RNAPs. The T7 RNAP enzymes tested were: (1) WT + G (foot glycine), (2) G47A variant, or (3) G47A variant + G (G47A* variant). IVT reactions contained one of the T7 RNAPs listed, shORFan template, NTPs (either unmodified or modified UTP) and ^{32}P -CTP. Reactions were terminated with EDTA

after the desired amount of time. The reaction mixture was then denatured in 4M urea at 95°C for 5 minutes before being loaded onto a denaturing polyacrylamide gel (20% acrylamide) (FIG. 26A). The gel ran for 30 min at 20 watts and then an additional 2 h at 40 watts. The image was transferred onto a phosphorimager screen (exposure time of 1 h) prior to imaging using the Typhoon scanner. Use of ³²P-CTP labels full-length product (arrow denoting shORFan mRNA), run-on transcription products (box above) (FIG. 26A), and reverse complement mRNA products (FIG. 26A, lower box). The results are also summarized in **Table 10** below. These data demonstrate that T7 RNAP with a foot glycine is associated with less contaminating dsRNA than either WT or G47A T7 RNAP (FIG. 26A (lower box) and FIG. 26B). Furthermore, a foot glycine is associated with decreased run-on transcription (FIG. 26A (upper box), FIG. 26B) compared with WT or G47A variant T7 RNAP.

Table 10. Ratio reverse complement (RC): full length (FL) mRNAs

Mutant	RC:FL	Full length (compared to G47A +G)
WT + G (Unmodified mRNA)	-	0.3
G47A (Unmodified mRNA)	0.16	1
G47A + G (G47A*) (Unmodified mRNA)	-	1
WT + G (Pseudouridine mRNA)	-	0.2
G47A (Pseudouridine mRNA)	0.18	1.2
G47A + G (Pseudouridine mRNA)	-	1

Example 14. Evaluation of Repeat Dosing of trinuc-capped G47A* mRNA Encoding Firefly Luciferase

In vitro transcribed mRNA is typically purified, for example, by reverse phase chromatography (RP) to remove cytokine-inducing impurities (dsRNA) and process impurities (e.g., protein/DNA), and to improve total RNA purity. This purification process, however, often results in a significant loss of mRNA product. Eliminating RP reduces turn-around-time, eliminates operational and engineering complexity at scale, and is cost saving. Further, because the high temperatures used during RP can hydrolyze a fraction of mRNA, eliminating RP increases mRNA expression levels. Thus, we have developed a mRNA production and purification strategy to eliminate the need for RP. The mRNAs produced by the methods provided in this Example are referred to as “trinuc-capped G47A* mRNAs,” which are mRNAs capped with the trinucleotide GpppA_{2'OME}pG co-transcriptionally in an *in vitro* transcription assay using the G47A* T7 RNAP variant (having a G47A substitution and an additional C-terminal G).

In this Example, a single 0.5 mpk dose of mRNA encoding firefly luciferase (ffLuc) was formulated in MC3 lipid nanoparticles and administered weekly to C57Bl/6 mice (n=5) for 6 weeks (on Day 1, Day 8, Day 15, Day 22, Day 29, and Day 36). The following mRNAs were administered: (1) unmodified mRNA produced using wild-type T7 polymerase in the presence of equimolar concentrations of NTPs and purified by oligo dT purification (Chemistry A Process 1 dT); (2) pseudouridine-modified mRNA produced using wild-type T7 polymerase in the presence of an excess concentration of GTP/ATP, and purified by reverse phase chromatography (Chemistry B Process 2 RP); (3) pseudouridine-modified mRNA produced using wild-type T7 polymerase in the presence of an excess concentration of GTP/ATP and purified by oligo dT purification (Chemistry B Process 2 dT); (4) pseudouridine-modified mRNA produced using wild-type T7 polymerase in the presence of equimolar concentrations of NTPs and purified by oligo dT purification (Chemistry B process 1 dT); (5) pseudouridine-modified trinuc-capped G47A* mRNA (trinucleotide capped and produced using the G47A* T7 RNAP variant) purified by reverse phase chromatography (Chemistry B Process 3 RP); (6) pseudouridine-modified trinuc-capped G47A* mRNA purified by oligo dT purification (Chemistry B Process 3 dT); and (7) uncapped pseudouridine-modified mRNA produced using the G47A* T7 RNAP variant and purified by oligo dT purification (Chemistry B Process 4 dT)). Six hours following each dose, the following were assessed: serum cytokine levels, mRNA expression, and anti-PEG IgM levels. B cell activation was assessed six hours following the Day 36 dose. Chemistry A denotes use of unmodified uridine triphosphate nucleotide and Chemistry B denotes use of 1-methyl-pseudouridine triphosphate nucleotide. Process 1 refers to equimolar NTPs in the IVT and vaccinia cap1, Process 2 refers to an IVT containing 4:2:1:1 GTP:ATP:CTP:UTP (see WO 2018/053209 A1, published 22 March 2018, incorporated herein by reference in its entirety) and vaccinia cap1, Process 3 refers to equimolar NTPs and GAG in the IVT, and Process 4 refers to equimolar NTPs in the IVT and no cap.

Results from the serum cytokine assessment (IP-10) in mice at Day 1 and Day 29 are presented in FIG. 27A and FIG. 27B, showing that trinuc-capped G47A* mRNAs induce serum cytokine levels in vivo, similar to alpha mRNA controls. The results were similar for in vitro experiments in which the mRNAs were delivered to BJ human fibroblasts (FIG. 28A) and monocyte-derived macrophages (MDMs) (FIG. 28B).

The mRNA expression studies (FIG. 29) showed that trinuc-capped G47A* mRNAs maintained high expression in vivo following 6 weekly doses.

The anti-PEG IgM assessment (FIG. 30) showed that N1U trinuc-capped G47A* mRNAs produce low anti-PEG IgM levels following 6 weekly doses.

The B cell activation analysis (FIG. 31) showed that activation of splenic B cells transfected with trinuc-capped G47A* mRNAs was low, similar to the alpha mRNA controls.

Collectively, these results indicate that trinuc-capped G47A* mRNAs encoding ffLuc, which are not subjected to RP purification, do not induce a cytokine response above baseline, maintain high expression levels in vivo, maintain low anti-PEG IgM levels in vivo, and exhibit low B cell activation.

Example 15. Evaluation of Repeat Dosing of Trinuc-capped G47A* mRNA Encoding Human Erythropoietin

In this Example, a single 0.5 mpk dose of mRNA encoding human erythropoietin (hEPO) was formulated in MC3 lipid nanoparticles and administered weekly to C57Bl/6 mice (n=5) for 6 weeks (on Day 1, Day 8, Day 15, Day 22, Day 29, and Day 36). The following mRNAs were administered: (1) unmodified mRNA produced using wild-type T7 polymerase in the presence of equimolar concentrations of NTPs and purified by oligo dT purification (Chemistry A Process 1 dT); (2) pseudouridine-modified mRNA produced using wild-type T7 polymerase in the presence of an excess concentration of GTP/ATP and purified by reverse phase chromatography (Chemistry B Process 2 RP); (3) pseudouridine-modified mRNA produced using wild-type T7 polymerase in the presence of an excess concentration of GTP/ATP and purified by oligo dT purification (Chemistry B Process 2 dT); (4) pseudouridine-modified mRNA produced using wild-type T7 polymerase in the presence of equimolar concentrations of NTPs and purified by oligo dT purification (Chemistry B Process 1 dT); (5) pseudouridine-modified trinuc-capped G47A* mRNA (trinucleotide capped and produced using the G47A* T7 RNAP variant) purified by reverse phase chromatography (Chemistry B Process 3 RP); (6) pseudouridine-modified trinuc-capped G47A* mRNA purified by oligo dT purification (Chemistry B Process 3 dT); and (7) uncapped pseudouridine-modified mRNA produced using the G47A* T7 RNAP variant and purified by oligo dT purification (Chemistry B Process 4 dT)). Six hours following each dose, the following were assessed: serum cytokine levels, mRNA expression, and anti-PEG IgM levels. B cell activation was assessed six hours following the Day 36 dose. Chemistry A denotes use of unmodified uridine triphosphate nucleotide and Chemistry B denotes use of 1-methyl-pseudouridine triphosphate nucleotide. Process 1 refers to equimolar NTPs in the IVT and vaccinia cap1, Process 2 refers to an IVT containing 4:2:1:1 GTP:ATP:CTP:UTP and

vaccinia cap1, Process 3 refers to equimolar NTPs and GAG in the IVT, and Process 4 refers to equimolar NTPs in the IVT and no cap.

Results from the serum cytokine assessment (IP-10) in mice at Day 1 and Day 22 are presented in FIG. 32A and FIG. 32B, showing that trinuc-capped G47A* mRNAs induce serum cytokine levels in vivo, similar to alpha mRNA controls. The results were similar for in vitro experiments in which the mRNAs were delivered to BJ human fibroblasts (FIG. 33A) and monocyte-derived macrophages (MDMs) (FIG. 33B).

The mRNA expression studies (FIG. 34) showed that trinuc-capped G47A* mRNAs maintained high expression in vivo following 6 weekly doses.

The anti-PEG IgM assessment (FIG. 35) showed that trinuc-capped G47A* mRNAs produce low anti-PEG IgM levels following 6 weekly doses.

The B cell activation analysis (FIG. 36) showed that activation of splenic B cells transfected with trinuc-capped G47A* mRNAs was low, similar to the alpha mRNA controls.

Collectively, these results indicate that trinuc-capped G47A* mRNAs encoding hEPO, which are not subjected to RP purification, do not induce a cytokine response above baseline, maintain high expression levels in vivo, maintain low anti-PEG IgM levels in vivo, and exhibit low B cell activation.

Example 16. Evaluation of Indel Frequency and Point Mutation Frequency

mRNAs were prepared with either WT or G47A* enzymes using Process 1 (equimolar NTPs) or Process 2 (4:2:1:1 GTP:ATP:CTP:UTP). For next generation sequencing (NGS), the mRNAs were converted to cDNA by reverse transcriptase and adaptors were ligated to prepare the sequencing library. The NGS data was compared to the parent sequence and any indels observed were tabulated. The indel frequency was comparable for WT and G47A* (FIG. 37A). Likewise, the NGS data was analyzed for any point mutations. The point mutation frequency was comparable for WT and G47A* (FIG. 37B).

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.”

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

In the claims, as well as in the specification above, all transitional phrases such as “comprising,” “including,” “carrying,” “having,” “containing,” “involving,” “holding,” “composed of,” and the like are to be understood to be open-ended, i.e., to mean including but not limited to. Only the transitional phrases “consisting of” and “consisting essentially of” shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.

CLAIMS

What is claimed is:

1. A ribonucleic acid (RNA) polymerase variant comprising an amino acid substitution, relative to wild-type RNA polymerase, that causes a loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex.
2. The RNA polymerase variant of claim 1, wherein the amino acid substitution has a higher helix propensity, relative to wild-type amino acid.
3. The RNA polymerase variant of claim 1 or 2, wherein the RNA polymerase is a T7 RNA polymerase.
4. The RNA polymerase variant of any one of claims 1-3, wherein the loop structure is in a C-helix structure.
5. The RNA polymerase variant of any one of claims 1-4, wherein the loop structure is in a C-linker structure.
6. The RNA polymerase variant of any one of claims 1-4, wherein the amino acid substitution is a high-helix propensity amino acid substitution.
7. The RNA polymerase variant of claim 6, wherein the high-helix propensity acid substitution is selected from alanine, isoleucine, leucine, methionine, lysine, glutamine, and glutamate.
8. The RNA polymerase variant of claim 7, wherein the high-helix propensity amino acid substitution is alanine.
9. The RNA polymerase variant of any one of claims 1-8, wherein the T7 RNA polymerase variant comprises an amino acid sequence having at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to SEQ ID NO:1.

10. The RNA polymerase variant of claim 9, wherein the T7 RNA polymerase comprises an amino acid sequence having at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity to SEQ ID NO:1 modified to include an amino acid substitution of a high-helix propensity amino acid at a position selected from E42, S43, Y44, E45, M46, G47, A255, R257, A258, G259, A260, L261, and A262.
11. The RNA polymerase variant of claim 10, wherein the amino acid substitution comprises S43A.
12. The RNA polymerase variant of claim 10, wherein the amino acid substitution comprises G47A.
13. The RNA polymerase variant of any one of claims 1-12 further comprising an additional C-terminal amino acid.
14. The RNA polymerase variant of claim 13, wherein the additional C-terminal amino acid comprises glycine (G).
15. The RNA polymerase variant of any one of claims 11, 13, or 14, wherein the RNA polymerase variant comprises the amino acid sequence of SEQ ID NO:108.
16. The RNA polymerase variant of any one of claims 12-14, wherein the RNA polymerase variant comprises the amino acid sequence of SEQ ID NO:110.
17. A ribonucleic acid (RNA) polymerase variant comprising an additional C-terminal amino acid relative to wild-type RNA polymerase.
18. The RNA polymerase variant of claim 17, wherein the RNA polymerase variant is a T7 RNA polymerase variant.
19. The RNA polymerase variant of claim 18, wherein the RNA polymerase variant has at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identity to wild-type T7 RNA polymerase.

20. The RNA polymerase variant of claim 17 or 18, wherein the RNA polymerase variant comprises an amino acid substitution relative to wild-type T7 RNA polymerase.
21. The RNA polymerase variant of any one of claims 17-20 comprising two additional C-terminal amino acids.
22. The RNA polymerase variant of claim 21, wherein the two additional C-terminal amino acids comprise the same type of amino acid or two different types of amino acids.
23. The RNA polymerase variant of any one of claims 17-22 comprising 1-5 or 1-10 additional C-terminal amino acids.
24. A RNA polymerase variant comprising a C terminus that comprises a FAFAX_n motif, wherein X is any amino acid and n is any integer greater than zero.
25. The RNA polymerase variant of claim 24, wherein X is G or A.
26. The RNA polymerase variant of claim 24 or 25, wherein n is any integer greater than zero, optionally wherein n is 1, 2, 3, 4, or 5.
27. A RNA polymerase variant, optionally a T7 RNA polymerase variant, comprising a C terminus that comprises a FAFAG motif.
28. A RNA polymerase variant, optionally a T7 RNA polymerase variant, comprising a C terminus that comprises a XAFAX_n motif, a FXFAX_n motif, a FAXAX_n motif, or a FAFXX_n motif, wherein each X is any amino acid, and n is any integer greater than zero.
29. The RNA polymerase variant of any one of claims 17-28 further comprising an amino acid substitution.
30. The RNA polymerase variant of claim 29, wherein the amino acid substitution is a high-helix propensity amino acid at a position selected from E42, S43, Y44, E45, M46, G47, R257, A258, G259, A260, L261, and A262 of a wild-type RNA polymerase, optionally

wherein the wild-type RNA polymerase comprises the amino acid sequence of SEQ ID NO:1 or an amino acid sequence having at least 90% identity with the amino acid sequence of SEQ ID NO:1.

31. The RNA polymerase variant of claim 30, wherein the high-helix propensity amino acid is alanine.

32. The RNA polymerase variant of claim 31, wherein the amino acid substitution is S43A or G47A.

33. A method of producing a ribonucleic acid (RNA) comprising contacting a DNA template with the RNA polymerase of any one of claims 1-32 under conditions that result in the production of RNA transcript.

34. A method of performing an in vitro transcription (IVT) reaction, comprising contacting a DNA template with the RNA polymerase of any one of claims 1-32 in the presence of nucleoside triphosphates and buffer under conditions that result in the production of RNA transcript.

35. The method of claim 33 or 34, wherein the RNA transcript produced, when delivered to cells, optionally in unpurified form, stimulates a cytokine response that is at least 50% lower relative to RNA produced using wild-type RNA polymerase.

36. The method of any one of claims 33-35, wherein the concentration of double-stranded RNA (dsRNA) transcript produced is at least 50% lower relative to dsRNA transcript produced using wild-type polymerase.

37. The method of any one of claims 33-36, wherein less than 20% of the RNA transcripts produced exhibit 3' heterogeneity.

38. The method of any one of claims 33-37, wherein less than 50% of the RNA transcript produced is double-stranded contaminant.

39. The method of any one of claims 33-38, wherein less than 50% of the RNA transcript produced is run-on RNA transcript.
40. The method of any one of claims 33-39, wherein the amount of full-length RNA transcript produced is at least 15 times greater than the amount of the DNA template.
41. The method of any one of claims 33-40, wherein the ratio of double-stranded contaminant:full-length RNA transcript produced is less than 1:1.
42. The method of any one of claims 33-41, wherein the RNA transcript produced has less than 1 mutation per 100 nucleotides relative to the DNA template.
43. A nucleic acid encoding the RNA polymerase variant of any one of claims 1-32.
44. A vector comprising the nucleic acid of claim 43.
45. A host cell comprising the nucleic acid of claim 43 or the vector of claim 44.
46. A composition or kit comprising the RNA polymerase variant of any one of claims 1-32 and optionally an in vitro transcription (IVT) reagent.
47. A ribonucleic acid (RNA), optionally a messenger RNA (mRNA), produced by the method of any one of claims 33-42.
48. The RNA of claim 47 formulated in a lipid nanoparticle, optionally wherein the lipid nanoparticle comprises a molar ratio of 20-60% ionizable amino lipid, 5-25% non-cationic lipid, 25-55% sterol, and 0.5-15% PEG-modified lipid.
49. A co-transcriptional capping method for ribonucleic acid (RNA) synthesis, the method comprising reacting a polynucleotide template with a RNA polymerase variant, optionally a T7 RNA polymerase variant, nucleoside triphosphates, and a cap analog under in vitro transcription reaction conditions to produce RNA transcript.

50. The method of claim 49, wherein greater than 80%, greater than 85%, greater than 90%, or greater than 95% of the RNA transcript produced includes a functional cap.
51. The method of claim 49 or 50, wherein the nucleoside triphosphates comprise unmodified or modified ATP, modified or unmodified UTP, modified or unmodified GTP, and/or modified or unmodified CTP.
52. The method of any one of claims 49-51, wherein the RNA polymerase variant is a RNA polymerase variant of any one of claims 1-32.
53. The method of any one of claims 49-52, wherein the nucleoside triphosphates and cap analog are present in the reaction at equimolar concentrations.
54. The method of any one of claims 49-53, wherein a molar ratio of cap analog to nucleoside triphosphates, optionally wherein the nucleoside triphosphates comprise GTP and/or ATP, in the reaction is greater than 1:1 or equal to 1:1.
55. The method of any one of claims 49-54, wherein the cap analog is a dinucleotide cap, a trinucleotide cap, or a tetranucleotide cap.
56. The method of claim 55, wherein the cap is a trinucleotide cap comprising a sequence selected from the following sequences: GAA, GAC, GAG, GAU, GCA, GCC, GCG, GCU, GGA, GGC, GGG, GGU, GUA, GUC, GUG, and GUU.
57. The method of claim 56, wherein the trinucleotide cap comprises a sequence selected from the following sequences:
- (a) $m^7GpppApA$, $m^7GpppApC$, $m^7GpppApG$, $m^7GpppApU$, $m^7GpppCpA$, $m^7GpppCpC$, $m^7GpppCpG$, $m^7GpppCpU$, $m^7GpppGpA$, $m^7GpppGpC$, $m^7GpppGpG$, $m^7GpppGpU$, $m^7GpppUpA$, $m^7GpppUpC$, $m^7GpppUpG$, and $m^7GpppUpU$;
 - (b) $m^7G_{3'OMe}pppApA$, $m^7G_{3'OMe}pppApC$, $m^7G_{3'OMe}pppApG$, $m^7G_{3'OMe}pppApU$, $m^7G_{3'OMe}pppCpA$, $m^7G_{3'OMe}pppCpC$, $m^7G_{3'OMe}pppCpG$, $m^7G_{3'OMe}pppCpU$, $m^7G_{3'OMe}pppGpA$, $m^7G_{3'OMe}pppGpC$, $m^7G_{3'OMe}pppGpG$, $m^7G_{3'OMe}pppGpU$, $m^7G_{3'OMe}pppUpA$, $m^7G_{3'OMe}pppUpC$, $m^7G_{3'OMe}pppUpG$, and $m^7G_{3'OMe}pppUpU$;

- (c) $m^7G_{3'OMe}pppA_{2'OMe}pA$, $m^7G_{3'OMe}pppA_{2'OMe}pC$, $m^7G_{3'OMe}pppA_{2'OMe}pG$,
 $m^7G_{3'OMe}pppA_{2'OMe}pU$, $m^7G_{3'OMe}pppC_{2'OMe}pA$, $m^7G_{3'OMe}pppC_{2'OMe}pC$, $m^7G_{3'OMe}pppC_{2'OMe}pG$,
 $m^7G_{3'OMe}pppC_{2'OMe}pU$, $m^7G_{3'OMe}pppG_{2'OMe}pA$, $m^7G_{3'OMe}pppG_{2'OMe}pC$,
 $m^7G_{3'OMe}pppG_{2'OMe}pG$, $m^7G_{3'OMe}pppG_{2'OMe}pU$, $m^7G_{3'OMe}pppU_{2'OMe}pA$,
 $m^7G_{3'OMe}pppU_{2'OMe}pC$, $m^7G_{3'OMe}pppU_{2'OMe}pG$, and $m^7G_{3'OMe}pppU_{2'OMe}pU$; or
- (d) $m^7GpppA_{2'OMe}pA$, $m^7GpppA_{2'OMe}pC$, $m^7GpppA_{2'OMe}pG$, $m^7GpppA_{2'OMe}pU$,
 $m^7GpppC_{2'OMe}pA$, $m^7GpppC_{2'OMe}pC$, $m^7GpppC_{2'OMe}pG$, $m^7GpppC_{2'OMe}pU$,
 $m^7GpppG_{2'OMe}pA$, $m^7GpppG_{2'OMe}pC$, $m^7GpppG_{2'OMe}pG$, $m^7GpppG_{2'OMe}pU$,
 $m^7GpppU_{2'OMe}pA$, $m^7GpppU_{2'OMe}pC$, $m^7GpppU_{2'OMe}pG$, and $m^7GpppU_{2'OMe}pU$.

58. The method of claim 56 or 57, wherein the trinucleotide cap comprises a sequence selected from the following sequences: GAG, GCG, GUG, and GGG.
59. The method of claim 58, wherein the trinucleotide cap comprises sequence GAG.
60. The method of claim 59, wherein the trinucleotide cap comprises $GpppA_{2'OMe}pG$.
61. The method of any one of claims 49-60, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.
62. The method of any one of claims 49-60, wherein the polynucleotide template includes a 2'-deoxycytidine residue at template position +1.
63. A co-transcriptional capping method for RNA synthesis, the method comprising reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a trinucleotide cap comprising sequence $GpppA_{2'OMe}pG$, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.

64. The method of any one of claim 49-63, wherein the RNA transcript produced, when delivered to cells, optionally in unpurified form, does not stimulate a detectable cytokine response or does not stimulate a cytokine response above a baseline control.
65. A composition comprising an in vitro-transcribed (IVT) RNA and a pharmaceutically acceptable excipient, wherein the composition is substantially free of cytokine-inducing RNA contaminant in the absence of post-IVT purification.
66. A composition comprising an in vitro-transcribed (IVT) RNA and a pharmaceutically acceptable excipient, wherein the composition has less than 5% uncapped RNA species.
67. The composition of claim 65 or 66, wherein greater than 80%, 85%, 90%, or 95% of the IVT RNA include a functional cap.
68. The composition of any one of claims 65-67, wherein the IVT RNA is not chemically modified.
69. The composition of any one of claims 65-67, wherein the IVT RNA is chemically modified.
70. The composition of any one of claims 65-69, wherein greater than 95% of the IVT RNA comprises single-stranded full-length transcripts.
71. The composition of any one of claims 65-70, wherein the RNA is produced by a process comprising:
reacting a polynucleotide template with (a) a T7 RNA polymerase variant comprising at least one amino acid substitution, relative to wild-type RNA polymerase, that causes at least one loop structure of the RNA polymerase variant to undergo a conformational change to a helix structure as the RNA polymerase variant transitions from an initiation complex to an elongation complex, (b) nucleoside triphosphates, and (c) a trinucleotide cap comprising sequence GpppA_{2'}OmepG, under in vitro transcription reaction conditions to produce RNA transcript, wherein the polynucleotide template includes a 2'-deoxythymidine residue at template position +1.

72. A T7 ribonucleic acid (RNA) polymerase comprising an amino acid sequence that has at least 90% identity with a wild-type T7 RNA polymerase and is modified to comprise (a) an alanine at an amino acid position corresponding to amino acid position 47 of SEQ ID NO:1 and (b) an additional C-terminal glycine.
73. The T7 RNA polymerase of claim 72 further comprising a trinucleotide cap.
74. The T7 RNA polymerase of claim 73, wherein the trinucleotide cap is GpppA_{2'}Ome pG.
75. A T7 ribonucleic acid (RNA) polymerase comprising an amino acid sequence that has at least 90% identity with a wild-type T7 RNA polymerase and is modified to comprise (a) an alanine at an amino acid position corresponding to amino acid position 43 of SEQ ID NO:1 and (b) an additional C-terminal glycine.
76. The T7 RNA polymerase of claim 75 further comprising a trinucleotide cap.
77. The T7 RNA polymerase of claim 76, wherein the trinucleotide cap is GpppA_{2'}Ome pG.

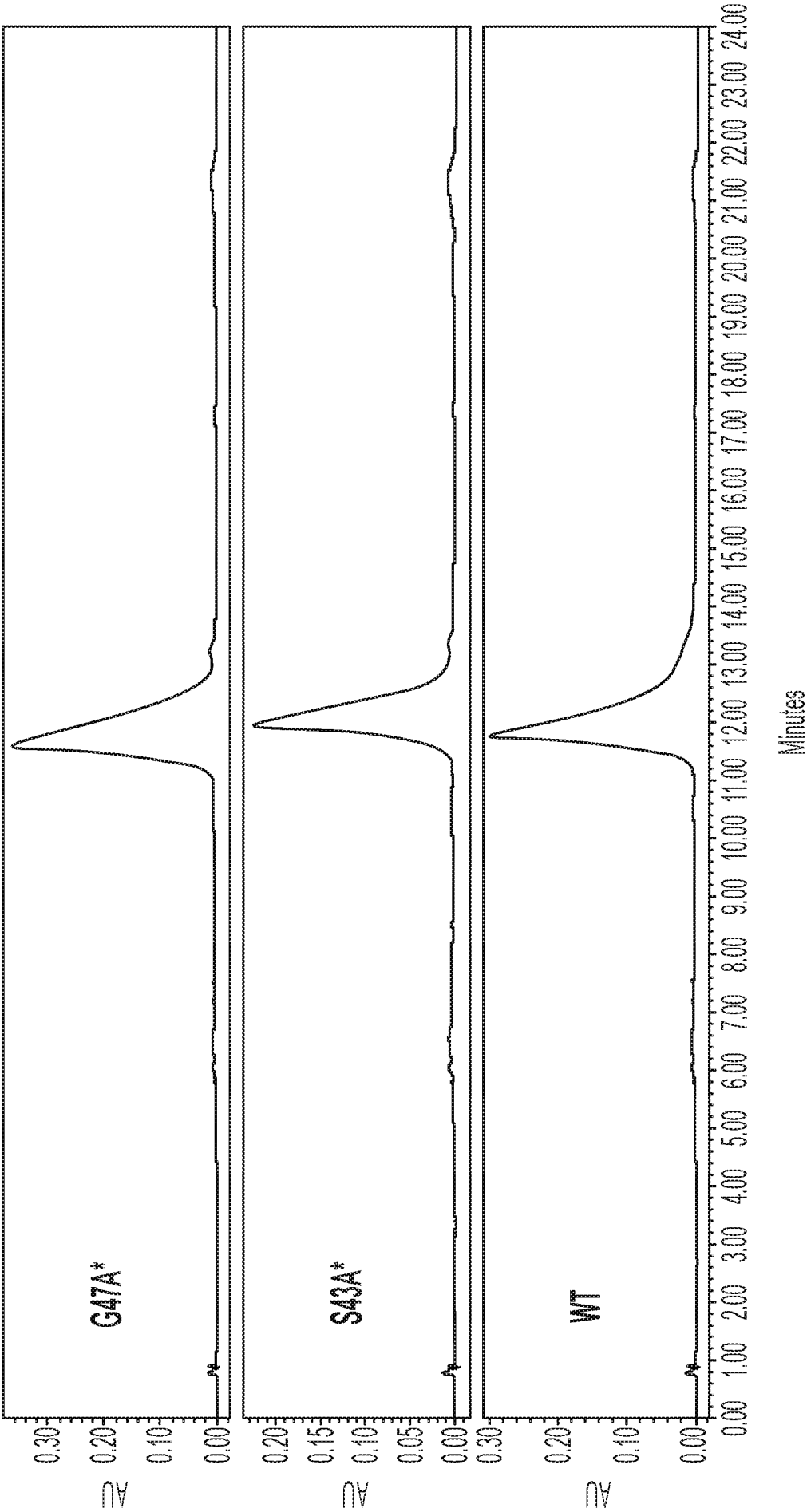


Figure 1

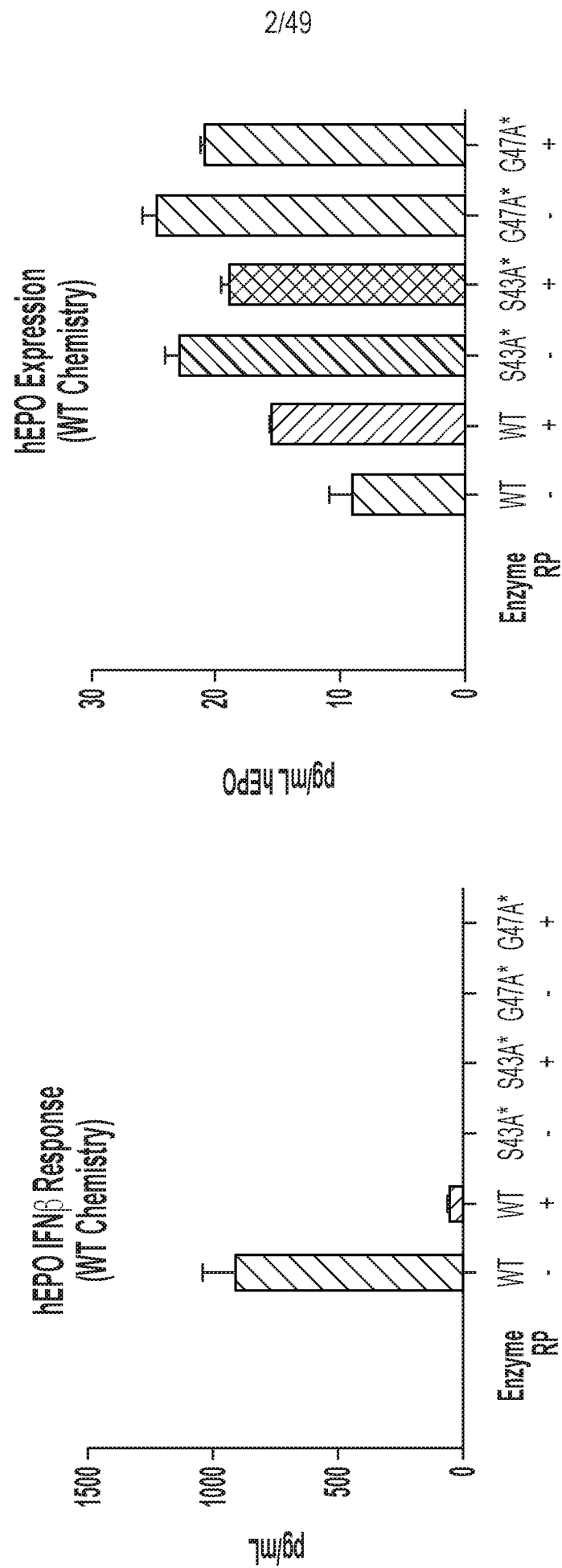


Figure 2

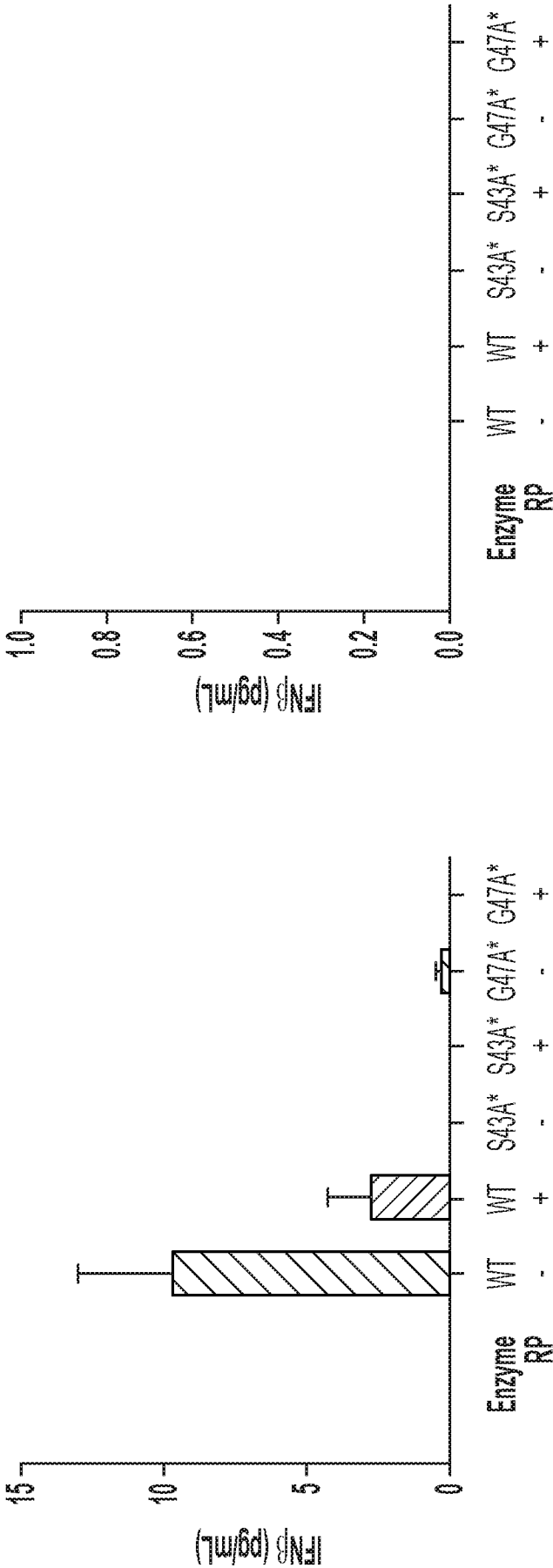


Figure 3

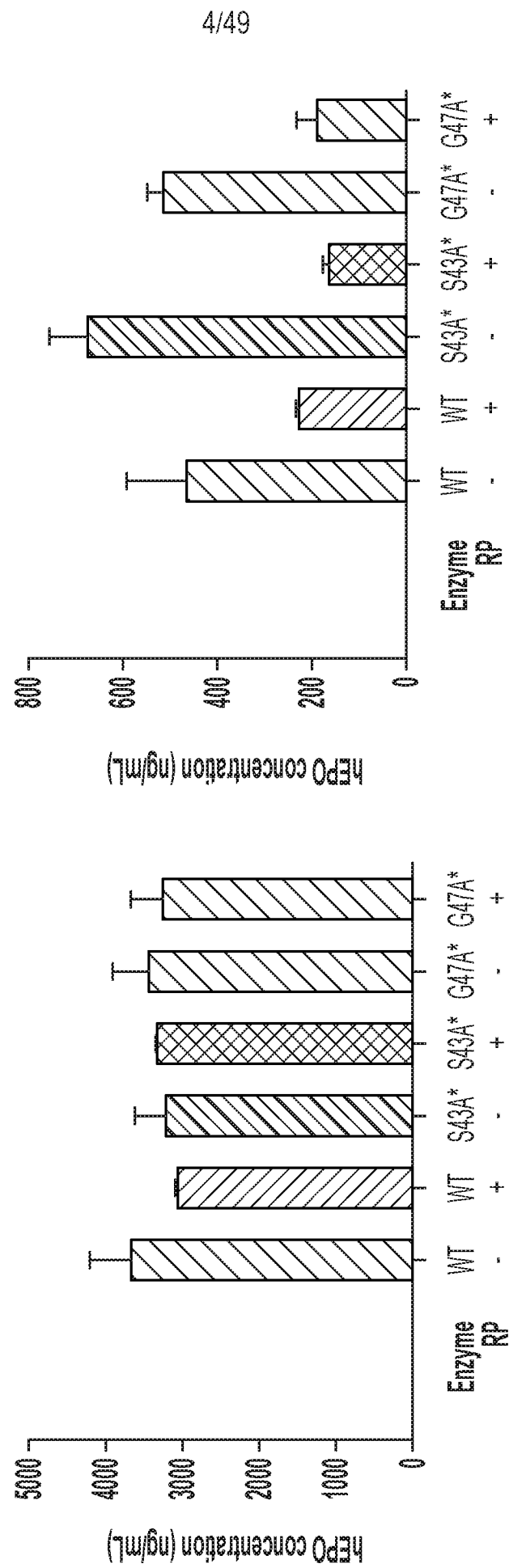


Figure 4

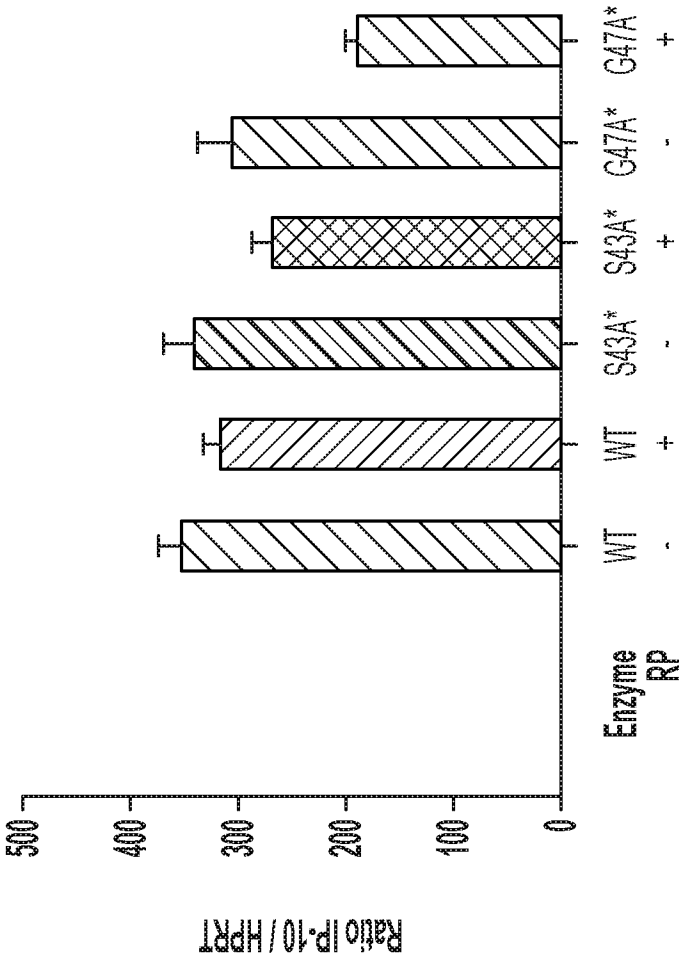


Figure 5

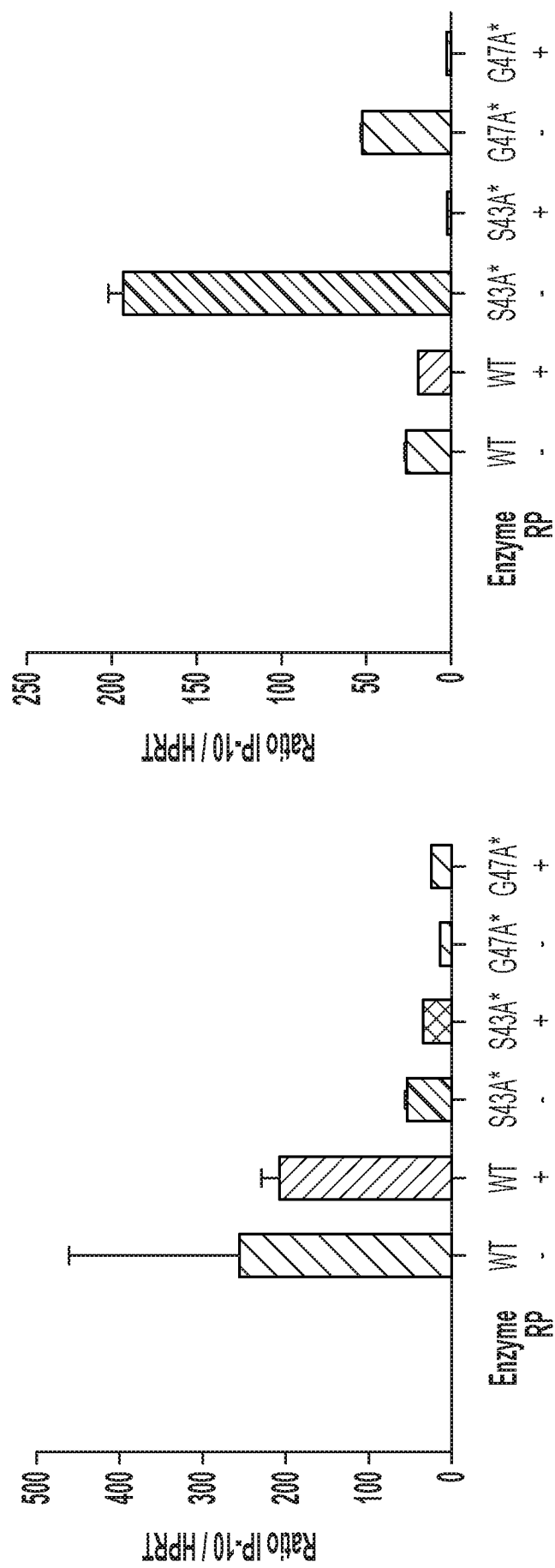


Figure 6

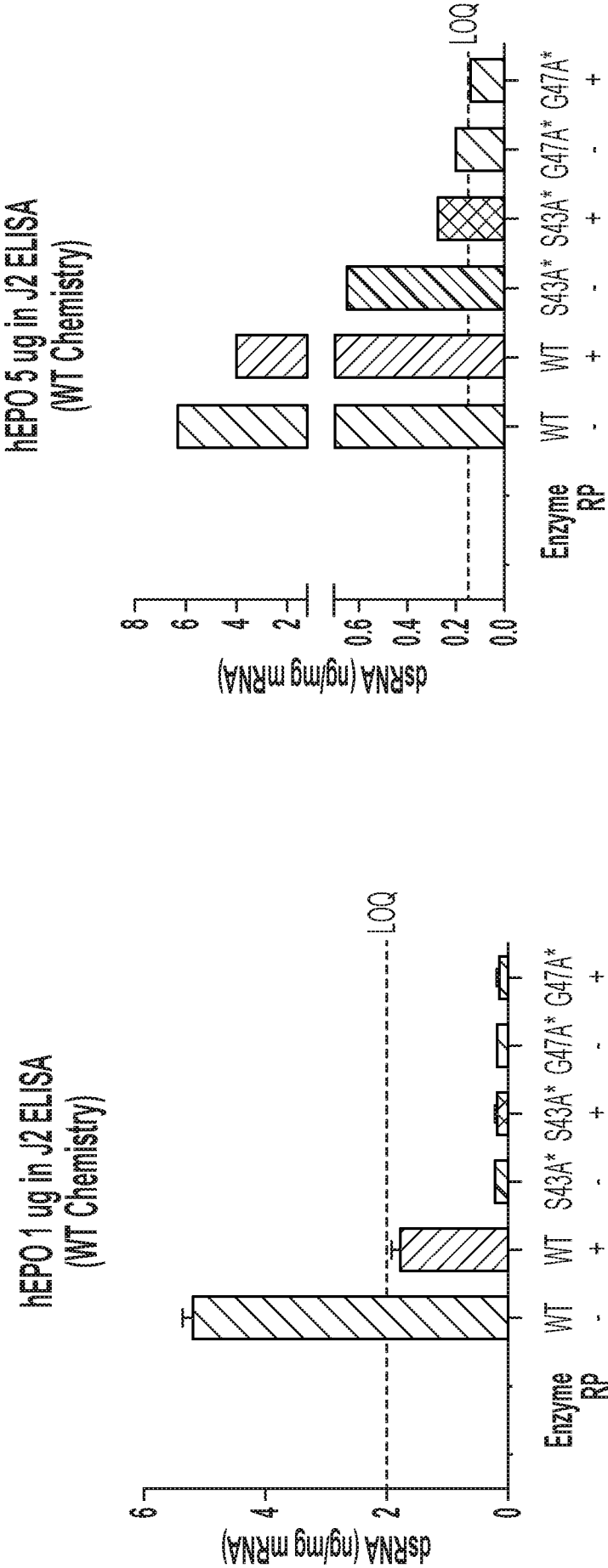


Figure 7

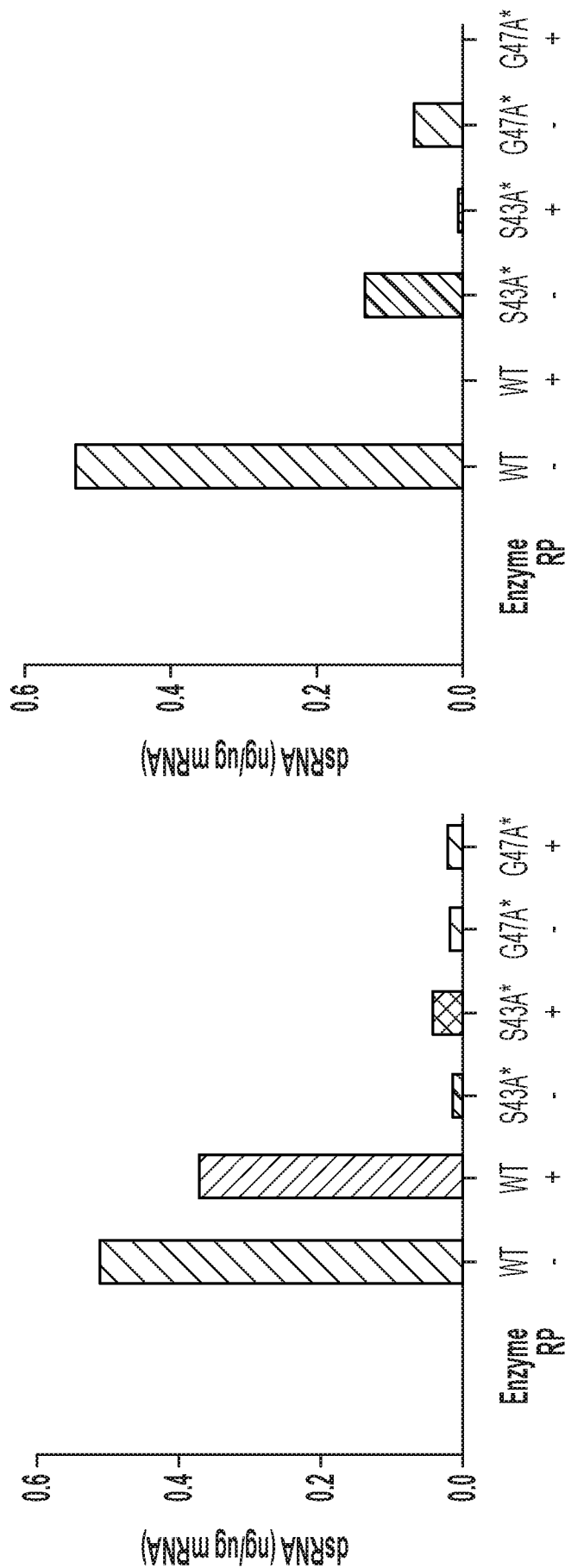


Figure 8

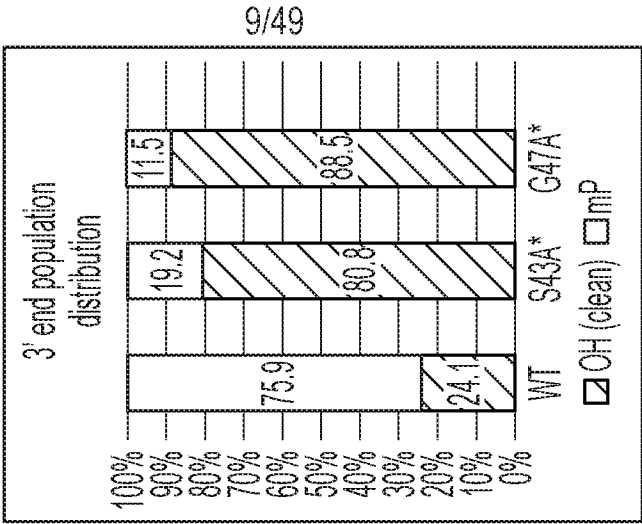
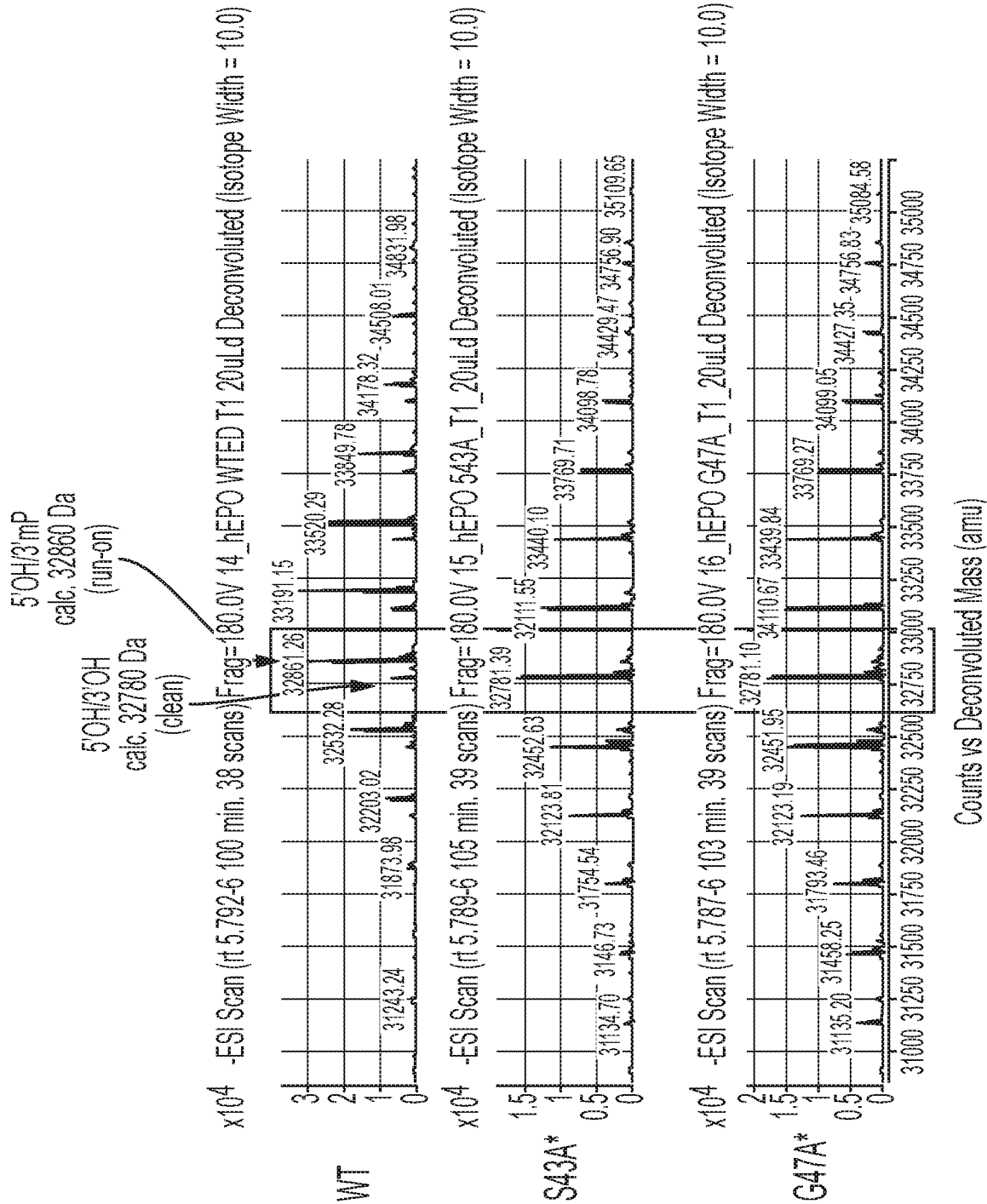


Figure 9B

Figure 9A

10/49

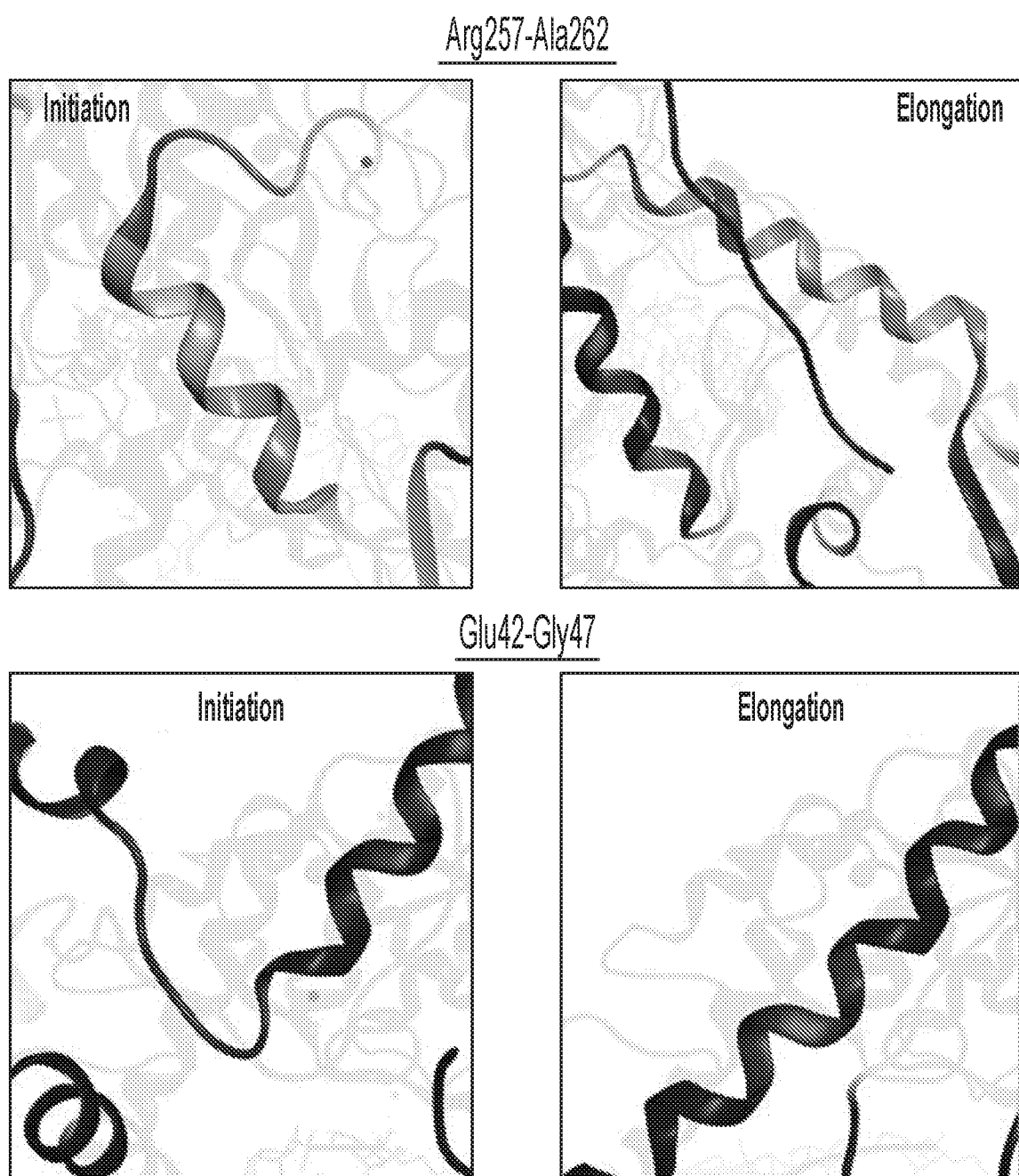


Figure 10

11/49

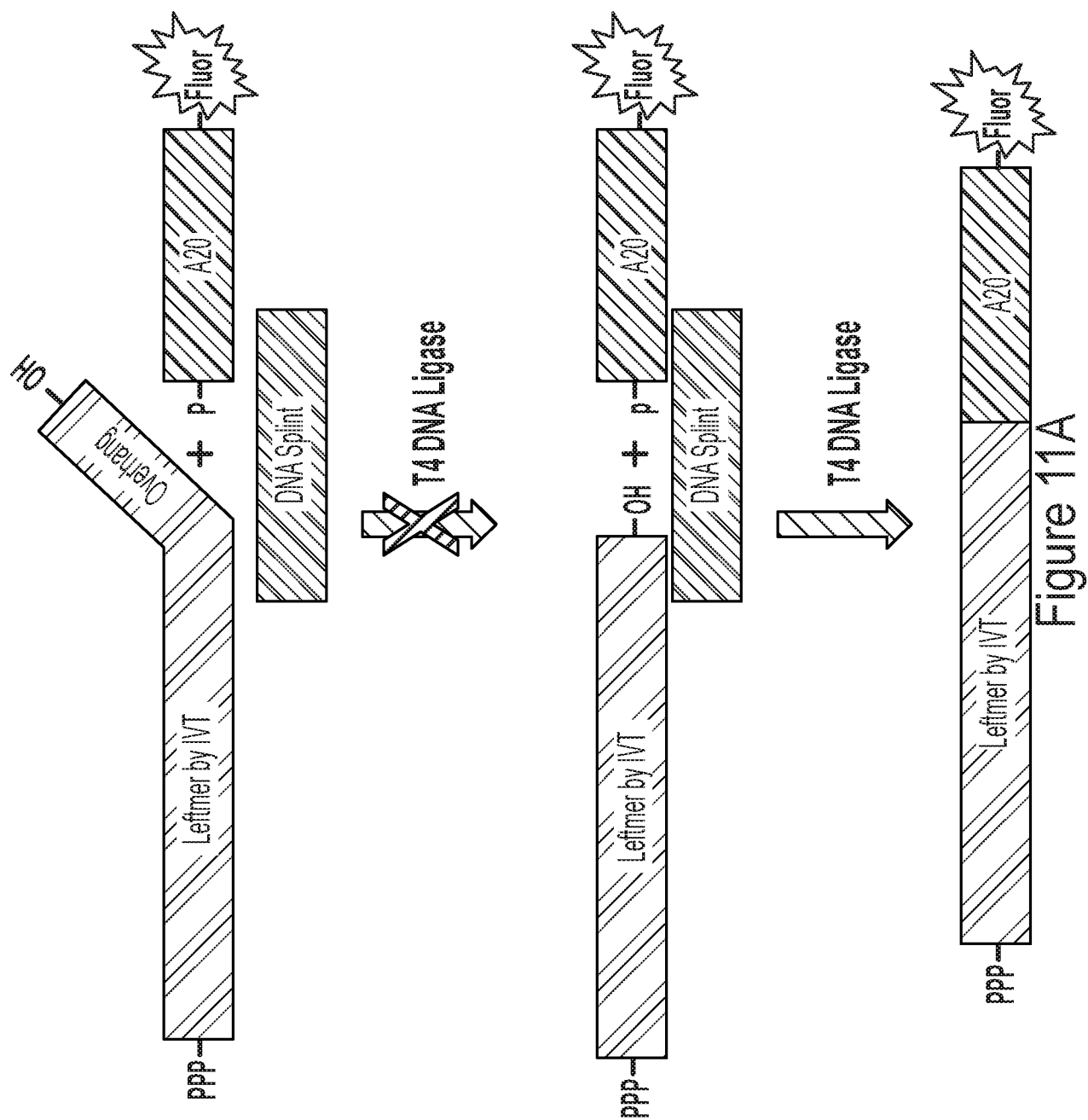


Figure 11A

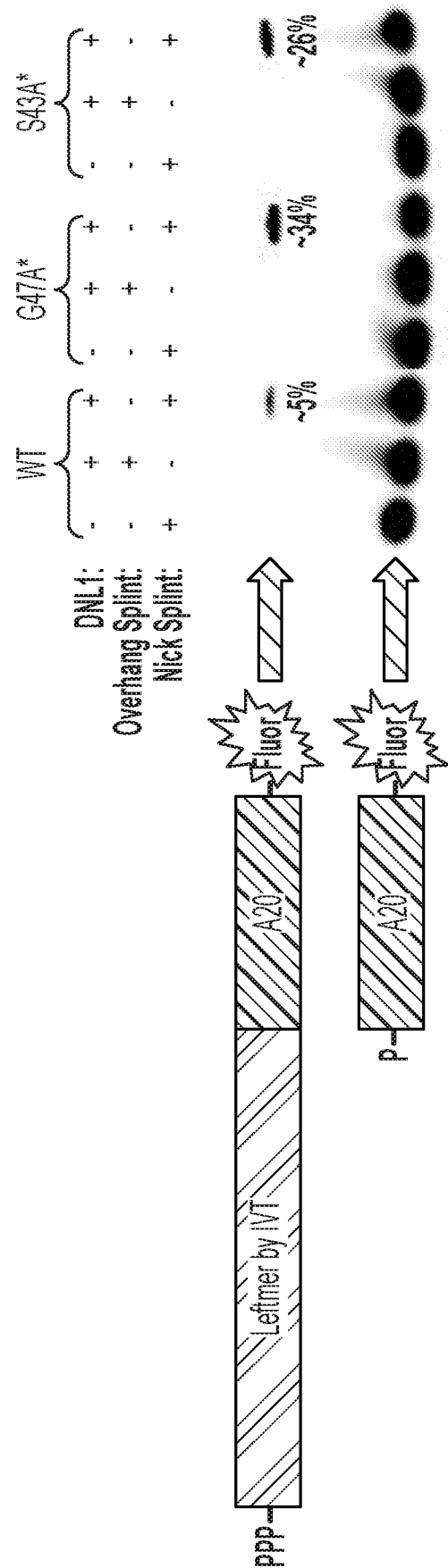


Figure 11B

13/49

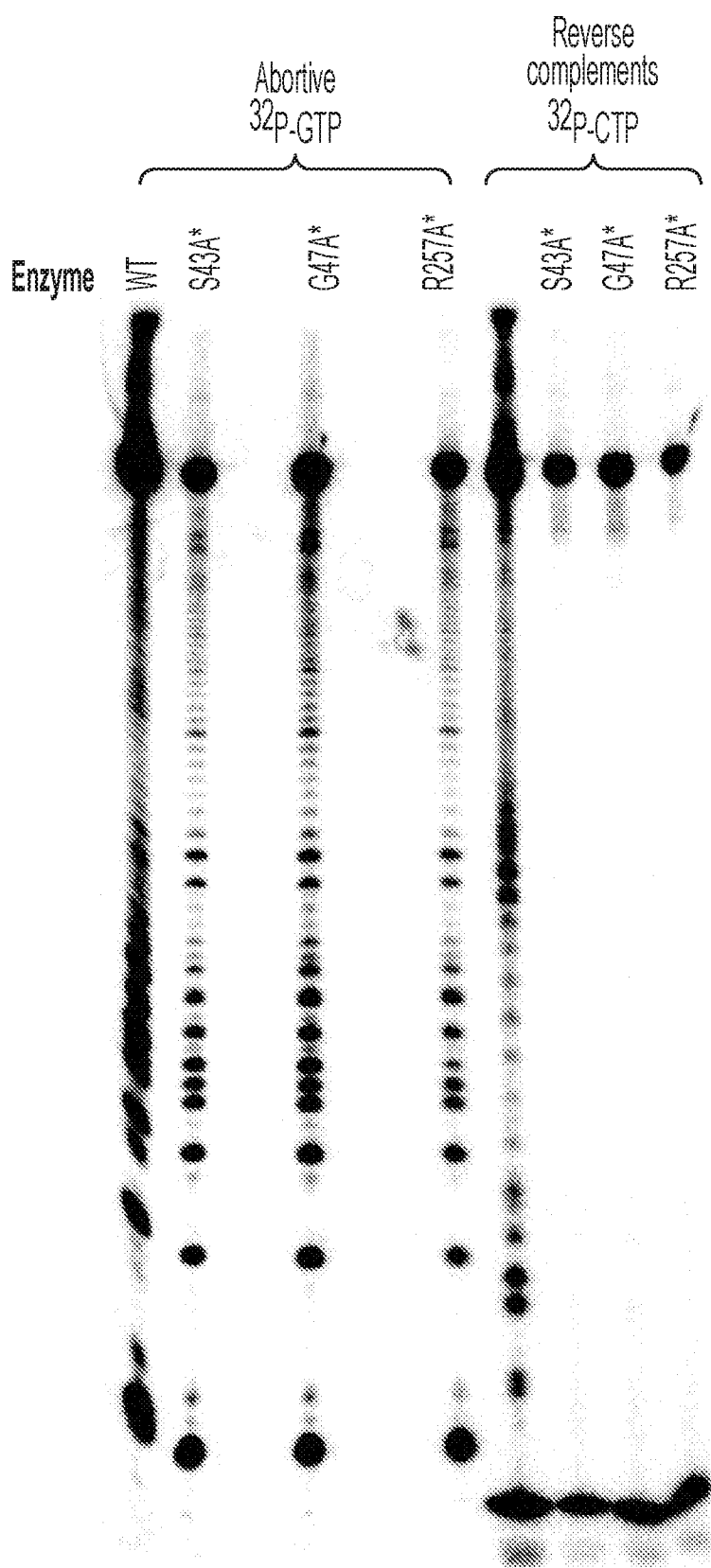


Figure 12

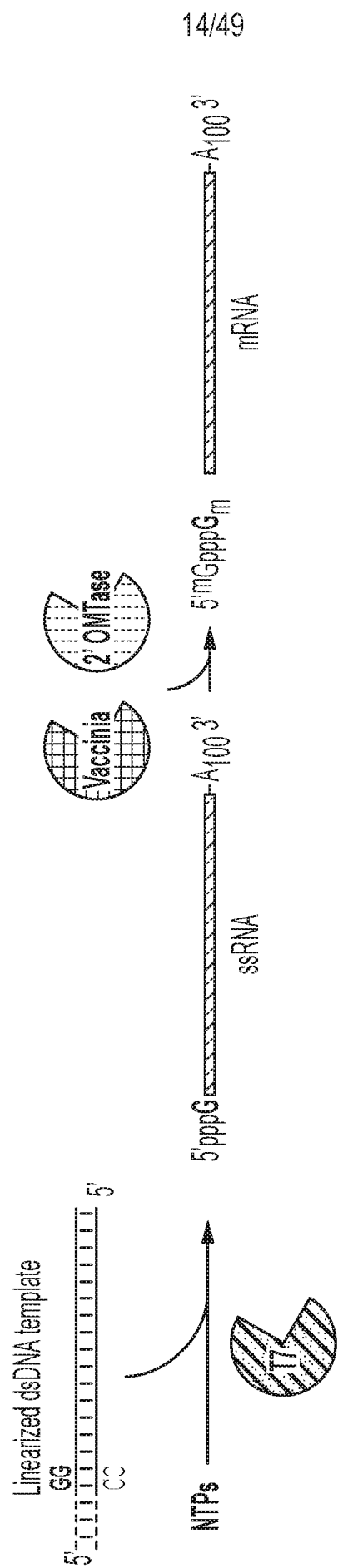
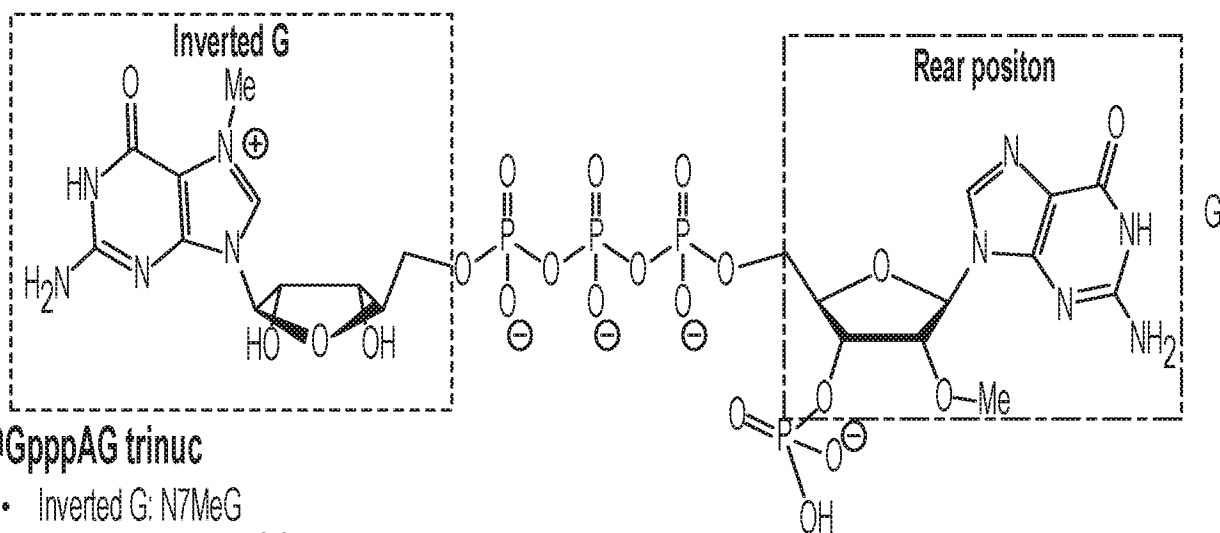


Figure 13

Vaccinia Cap1

15/49

- Inverted G: N7MeG
- Rear position: G with 2'OMe

**mGmpppAG trinuc**

- Inverted G: N7MeG with 3'OMe
- Rear position: A with 2'OMe

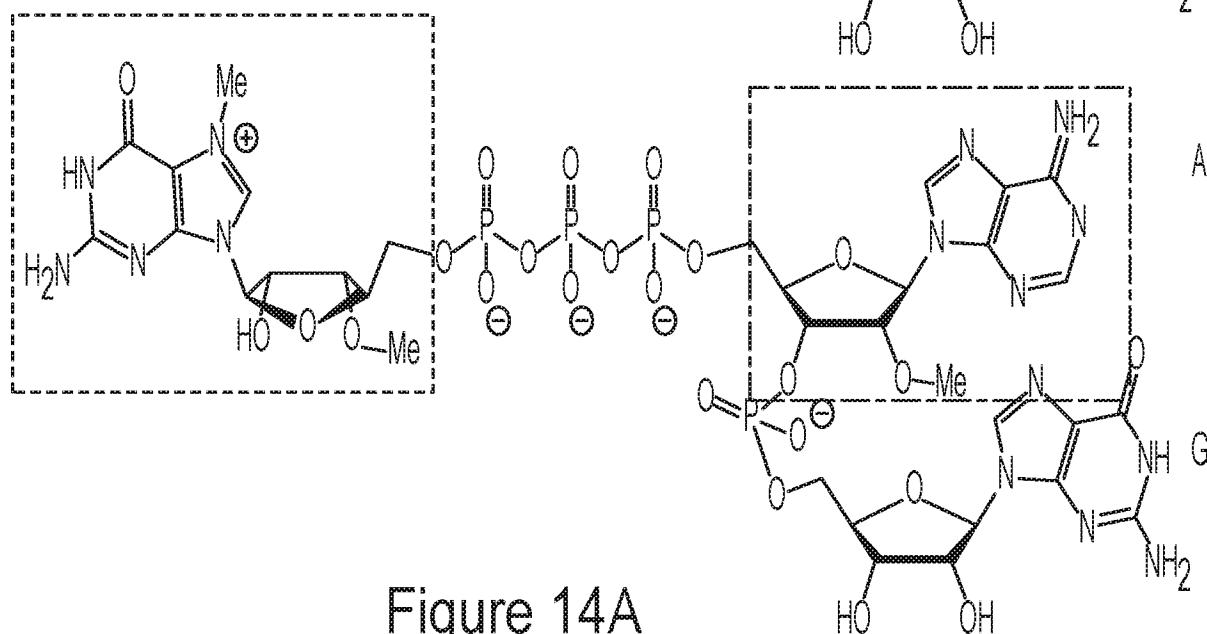


Figure 14A

16/49

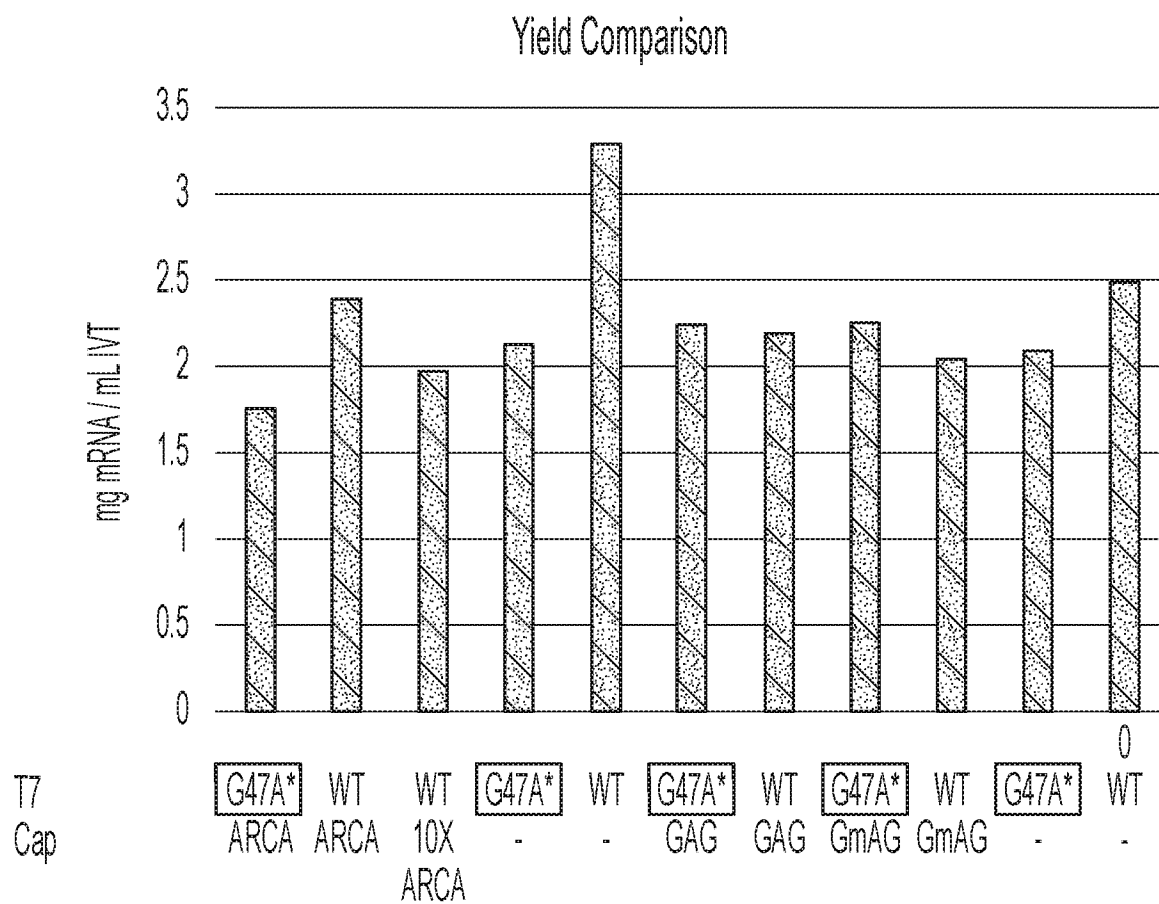


Figure 14B

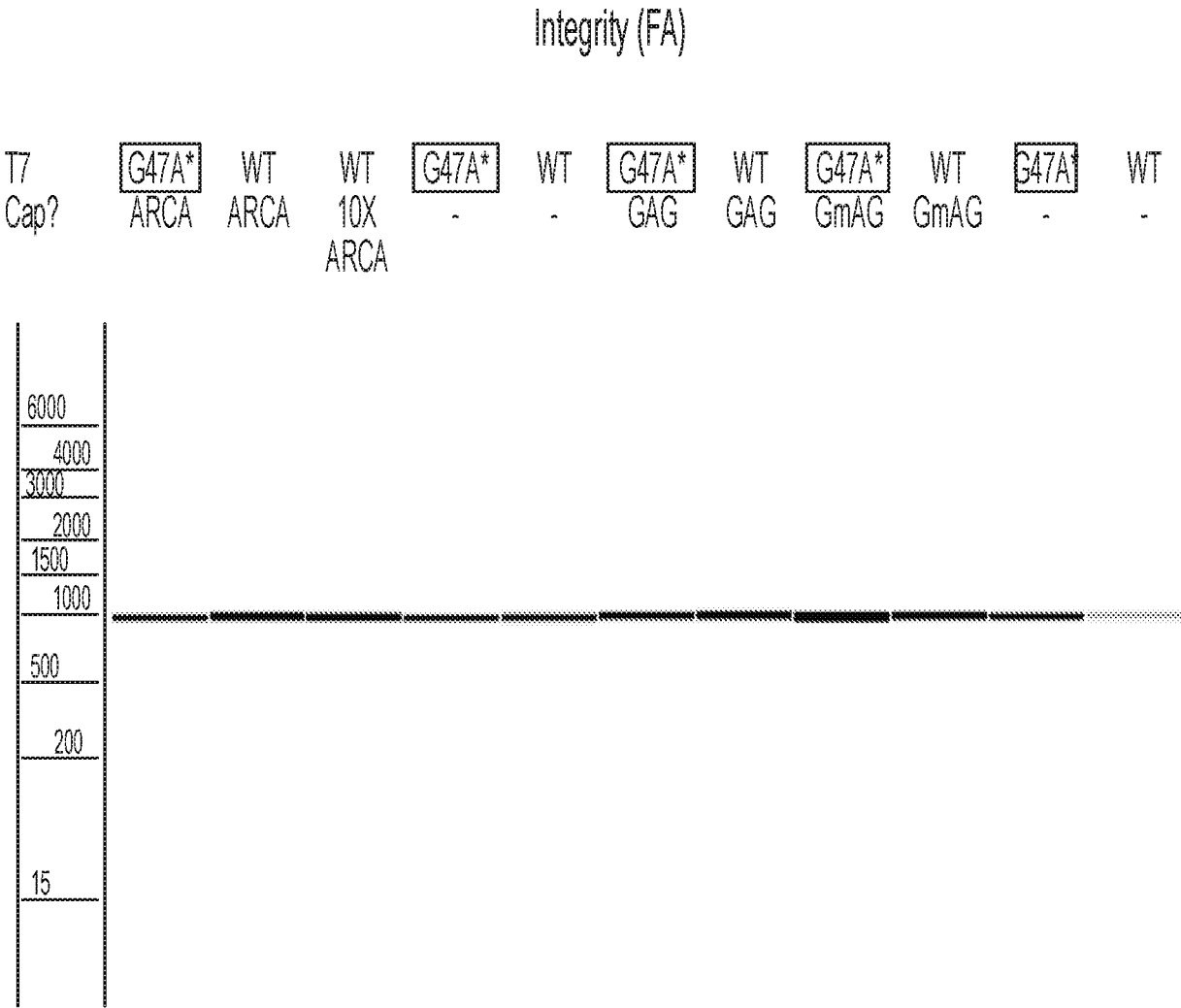


Figure 14C

18/49

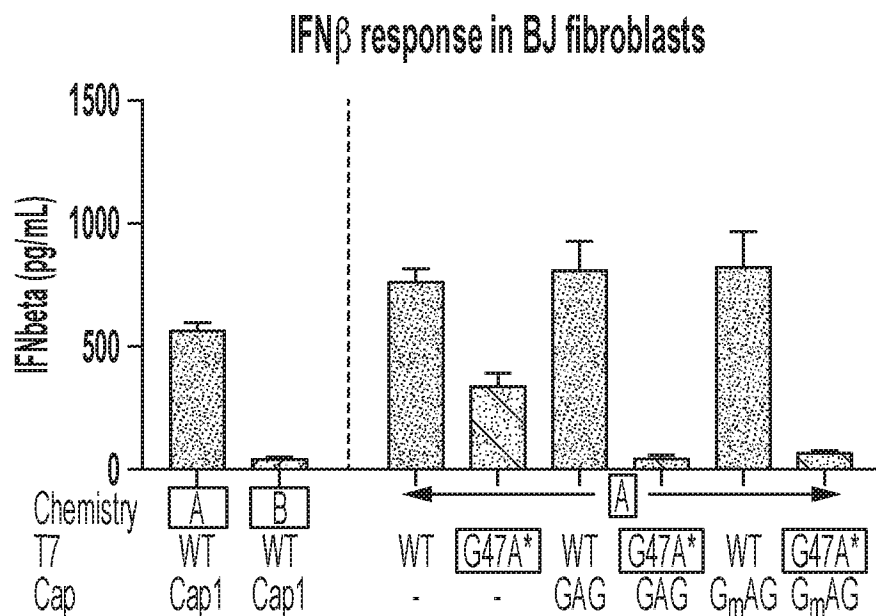


Figure 14D

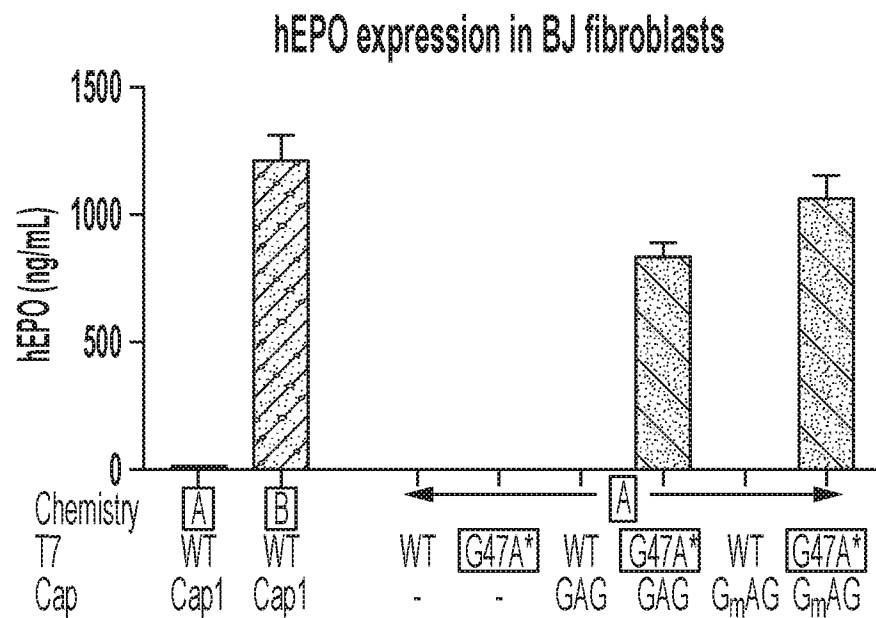


Figure 14E

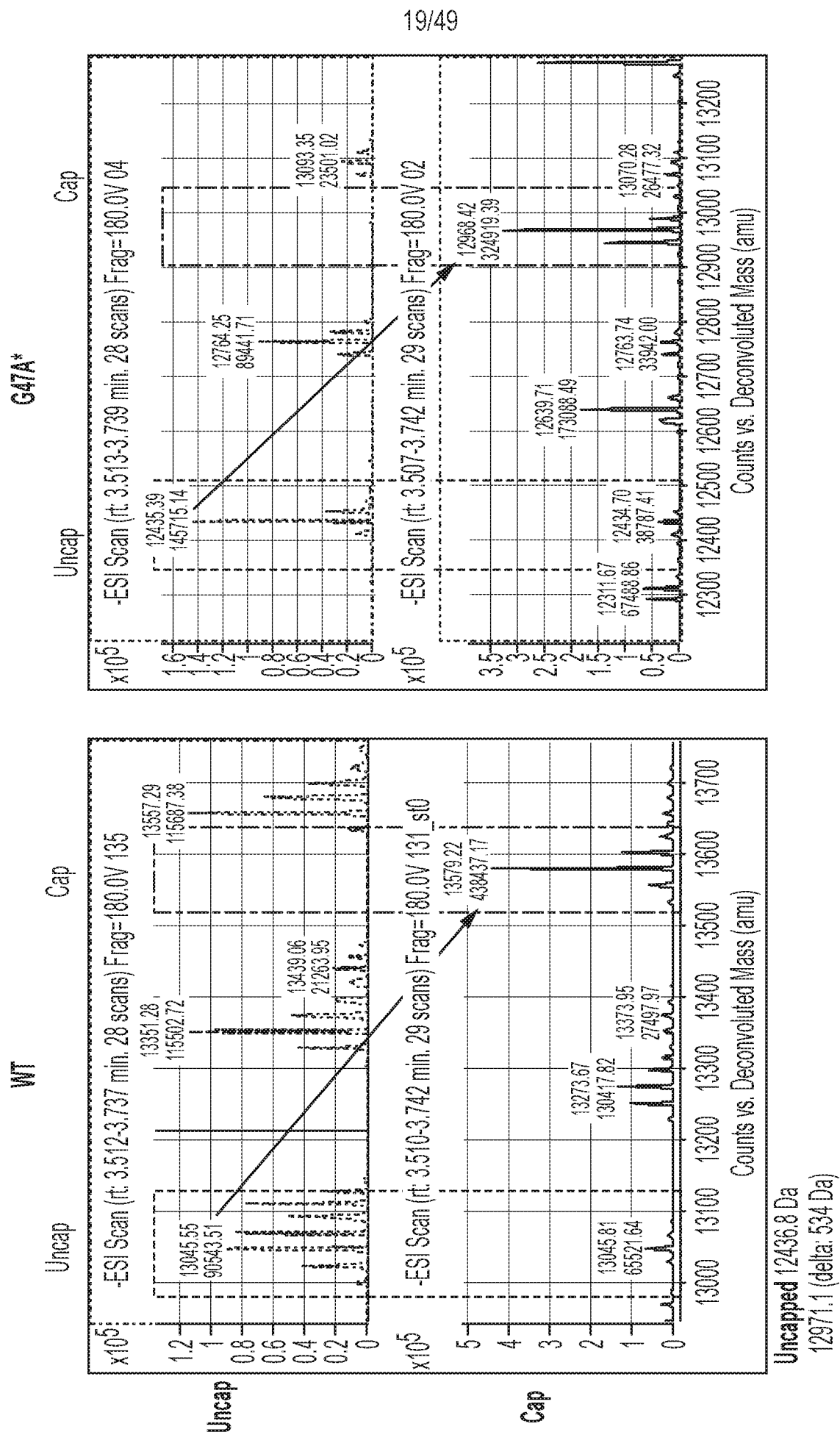


Figure 15

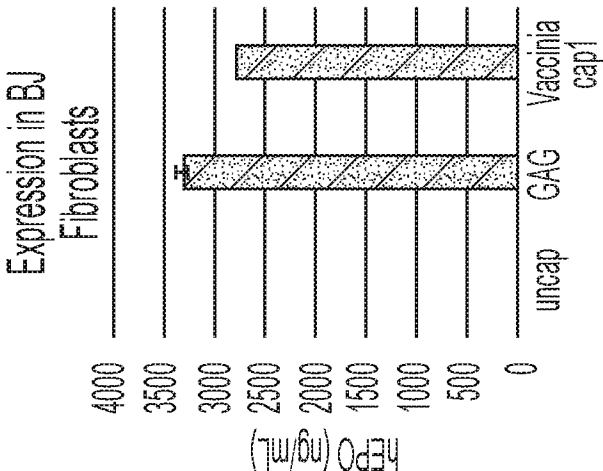


Figure 16B

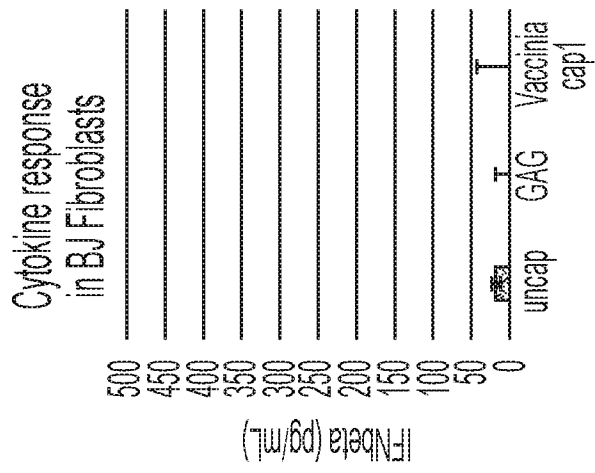


Figure 16A

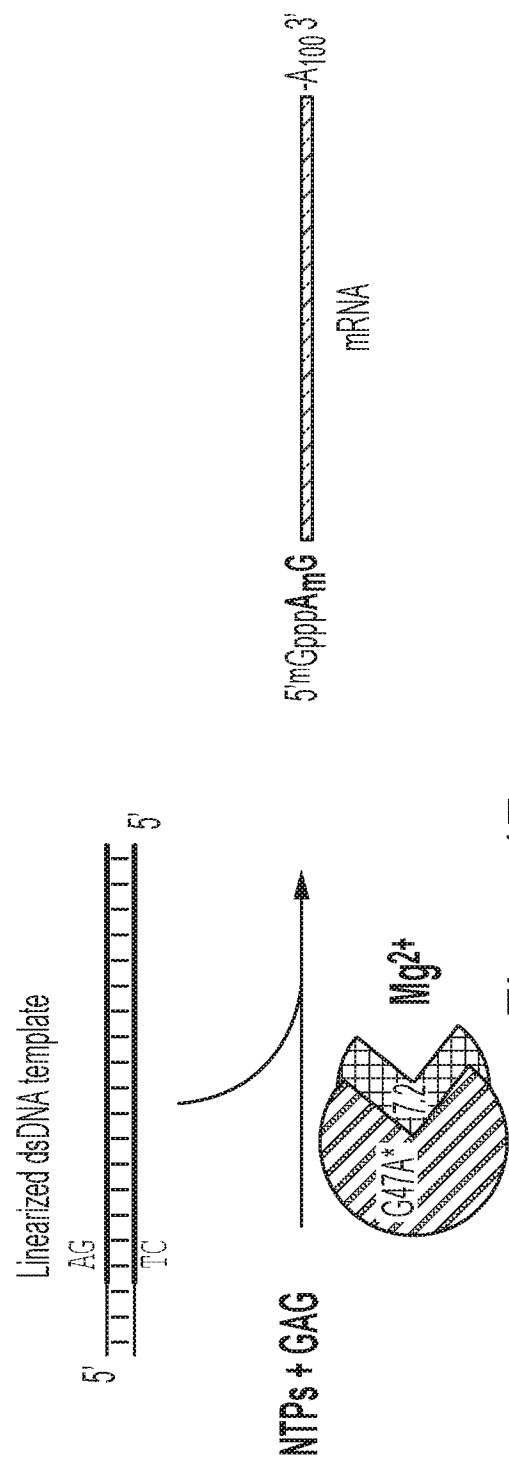


Figure 17

21/49

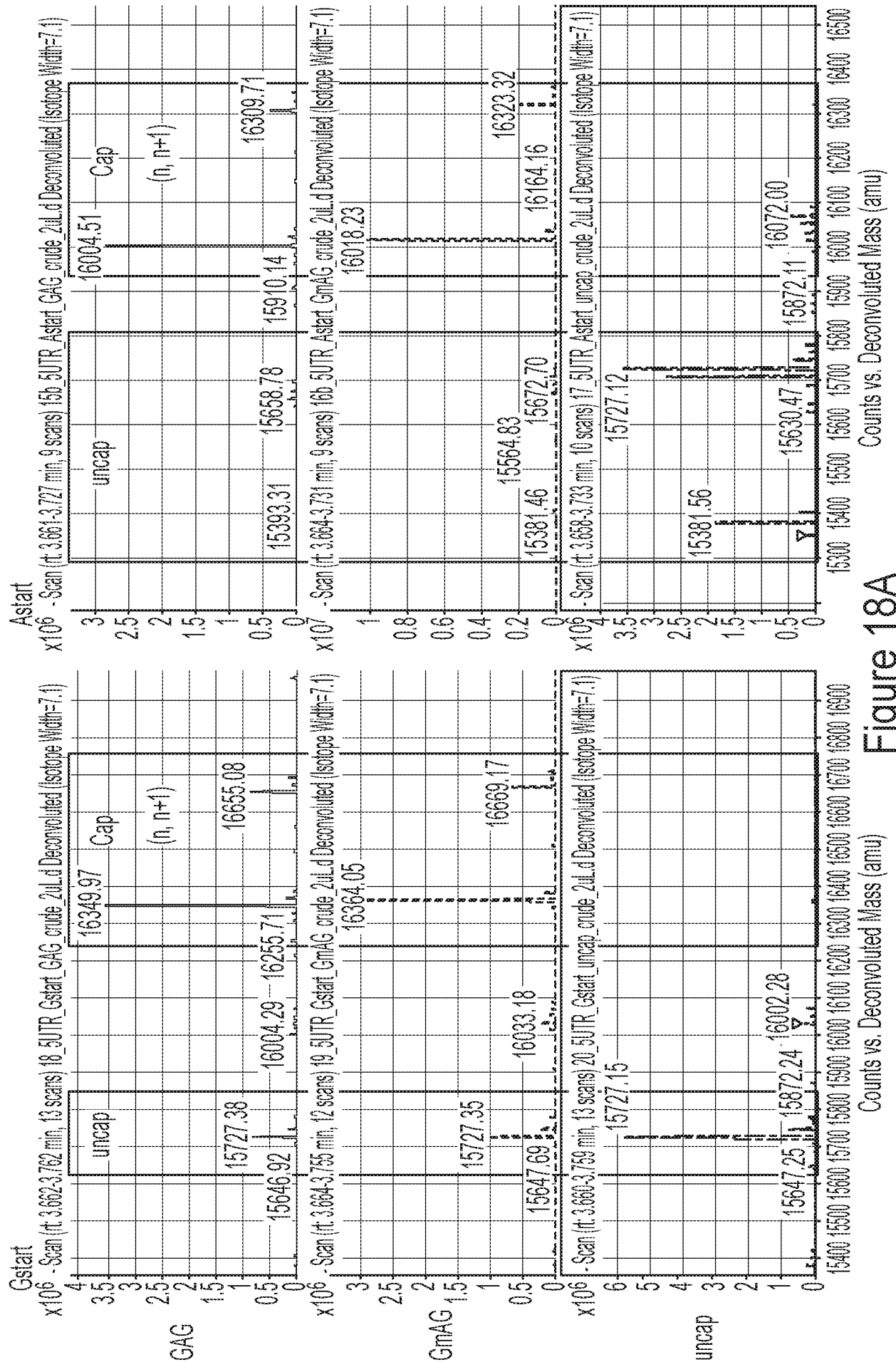
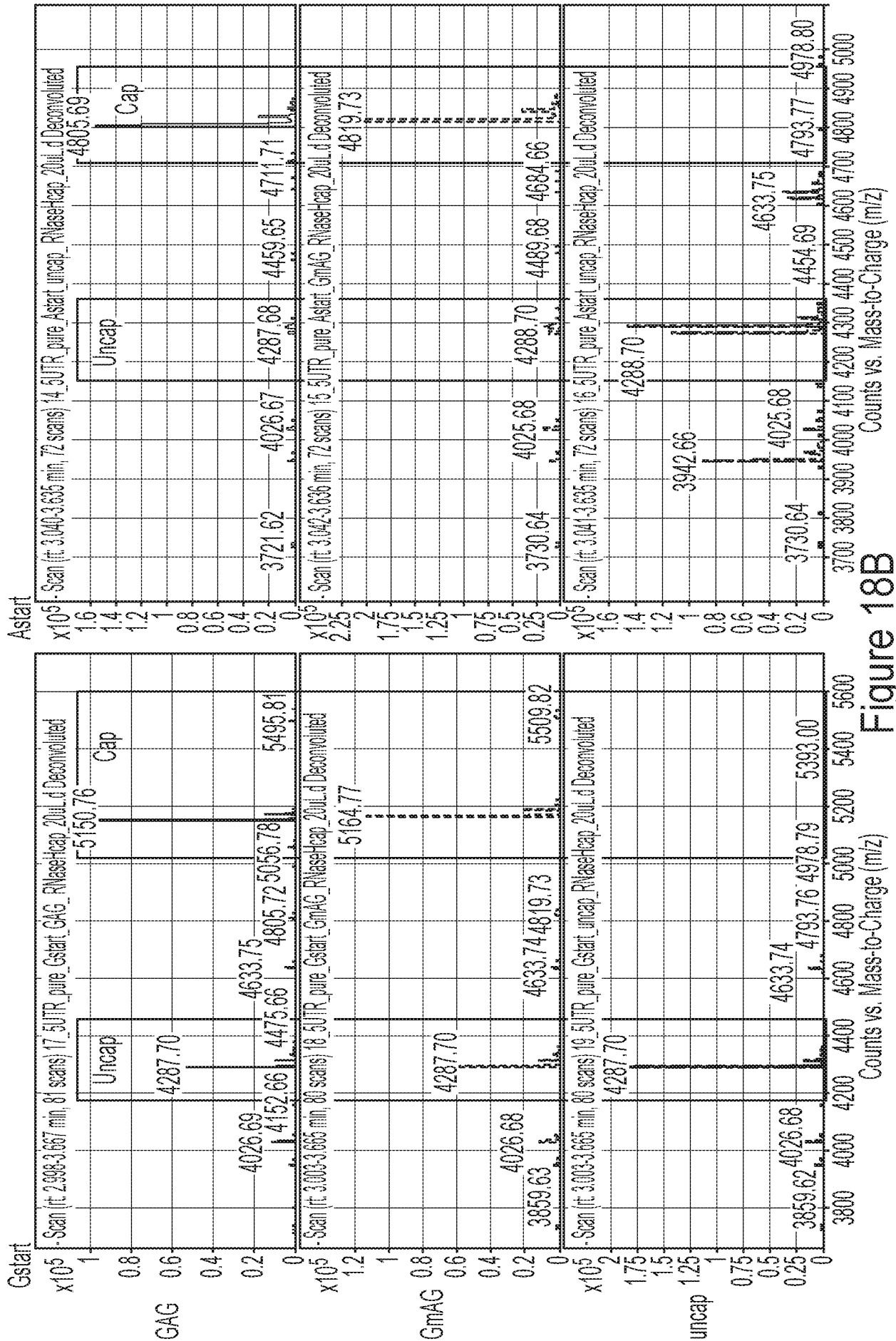


Figure 18A

22/49



23/49

--- Uncap
---- Production cap1
— GAG
— GmAG

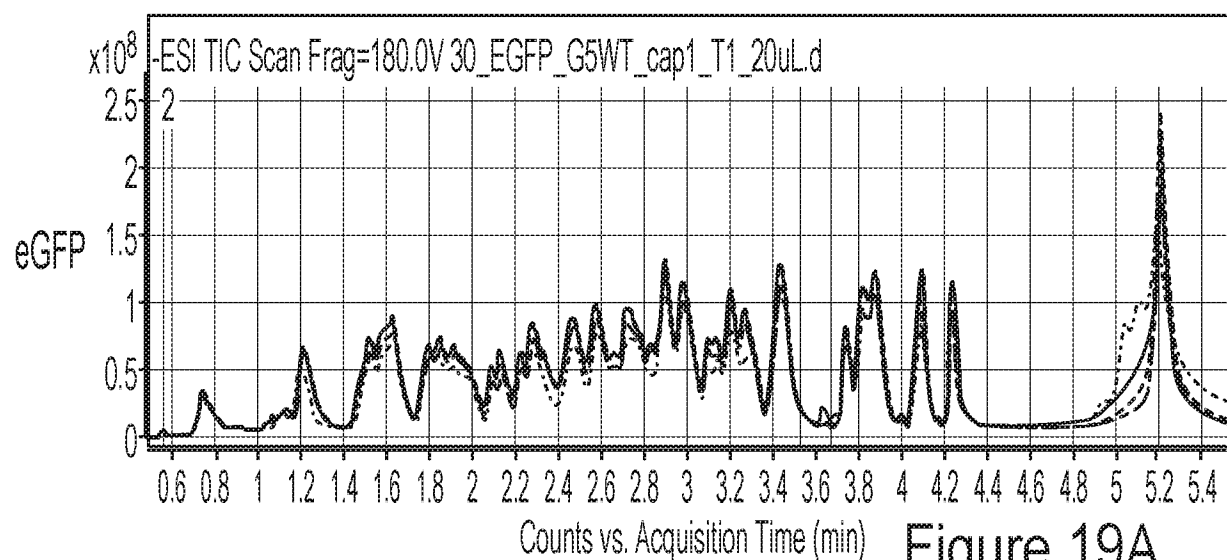
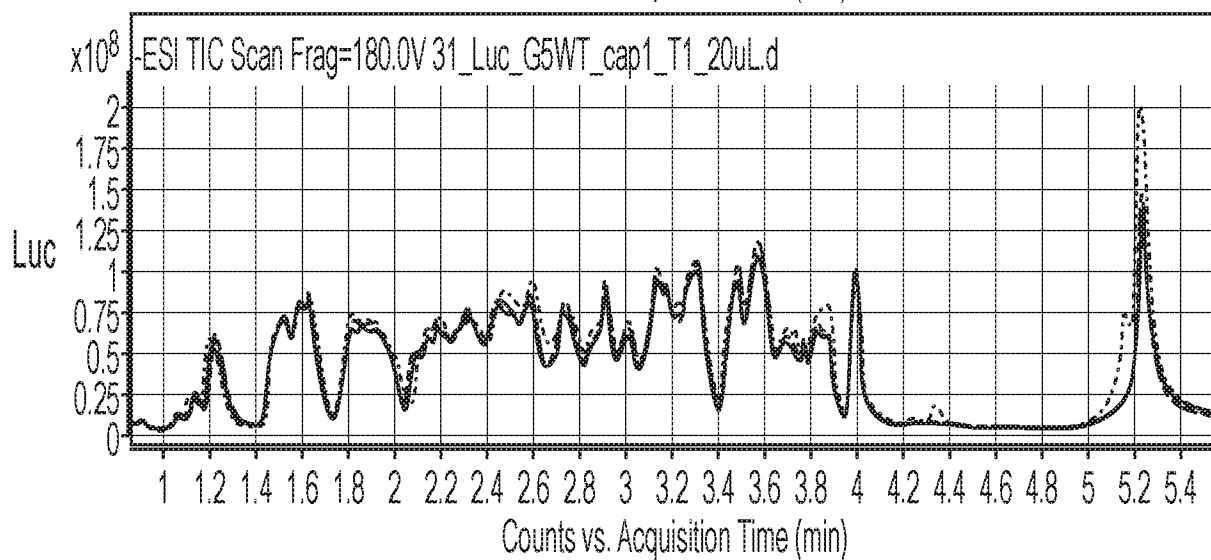
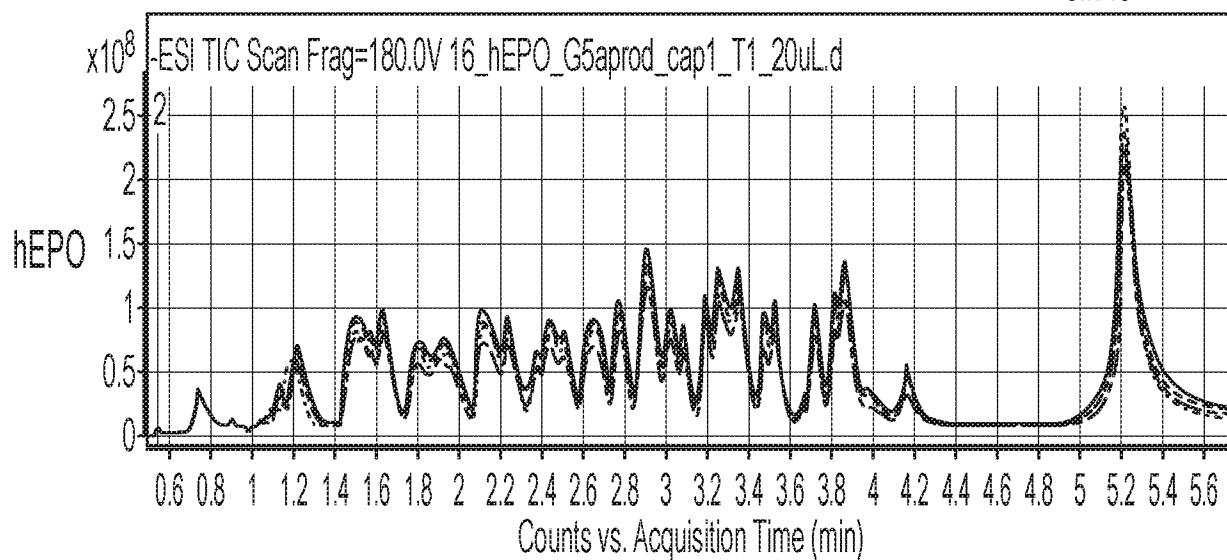


Figure 19A

24/49

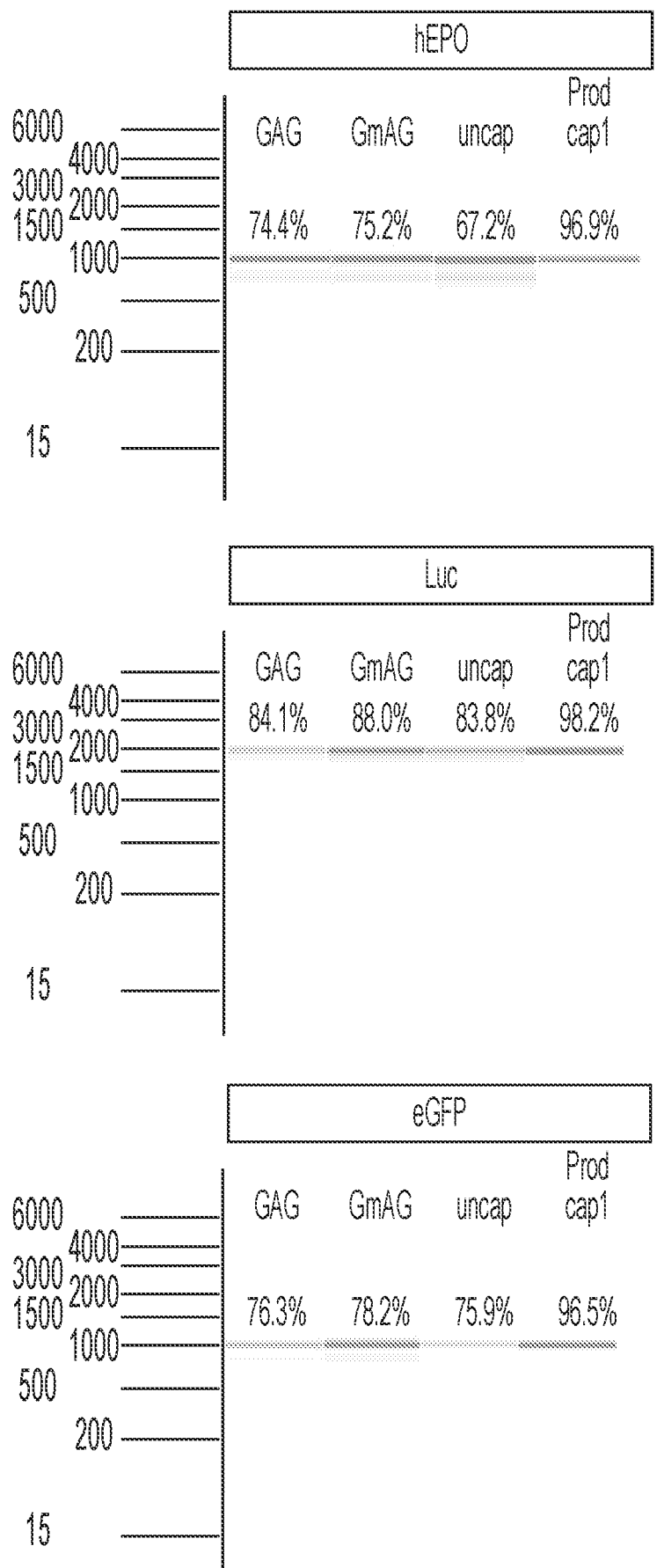


Figure 19B

25/49

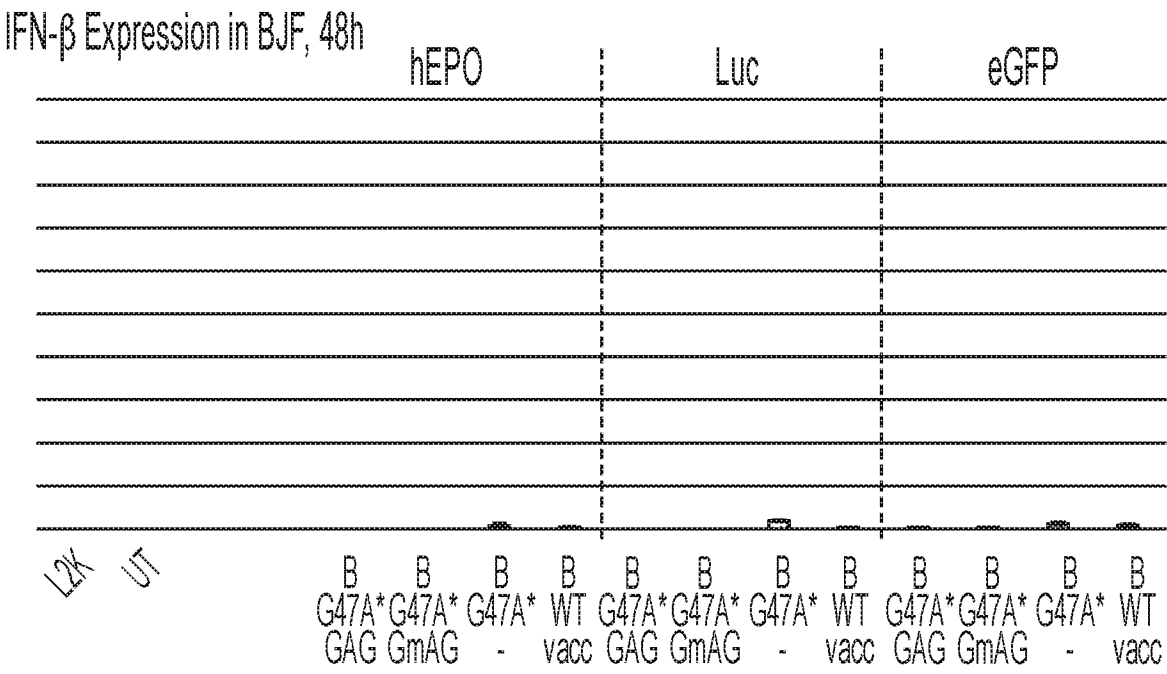
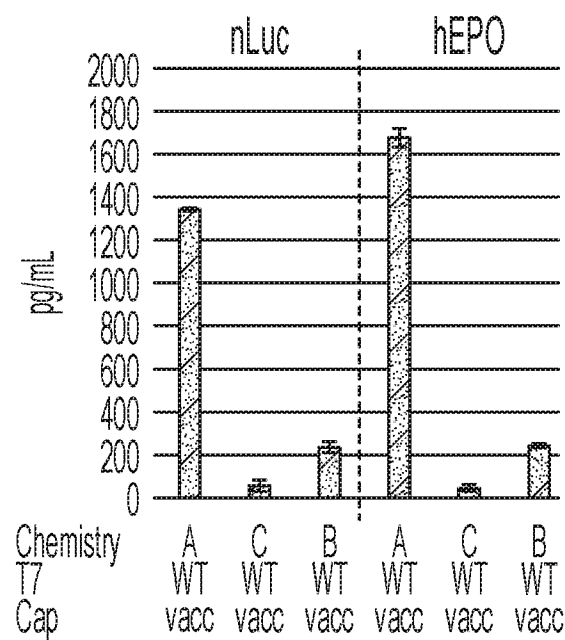


Figure 19C

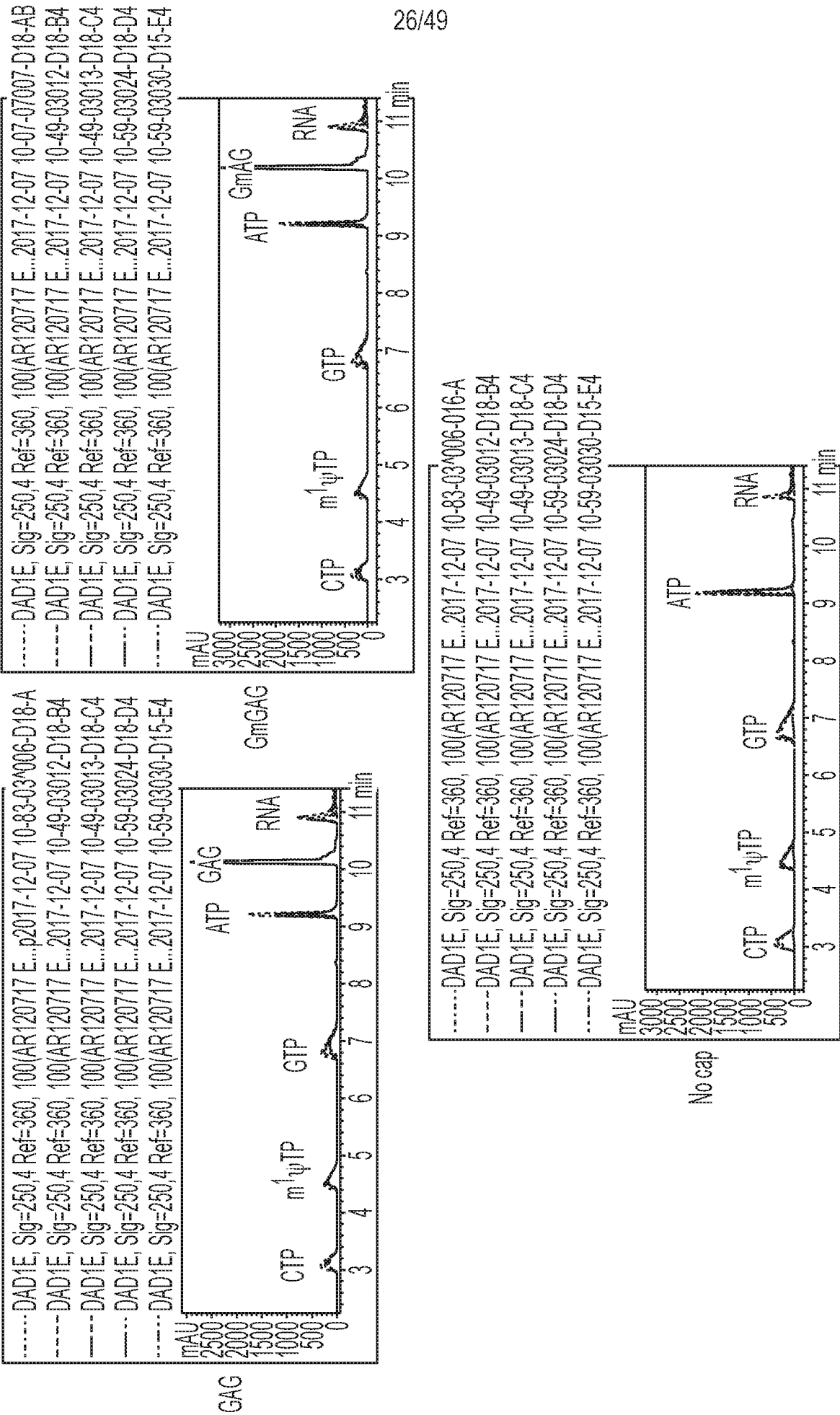


Figure 20A

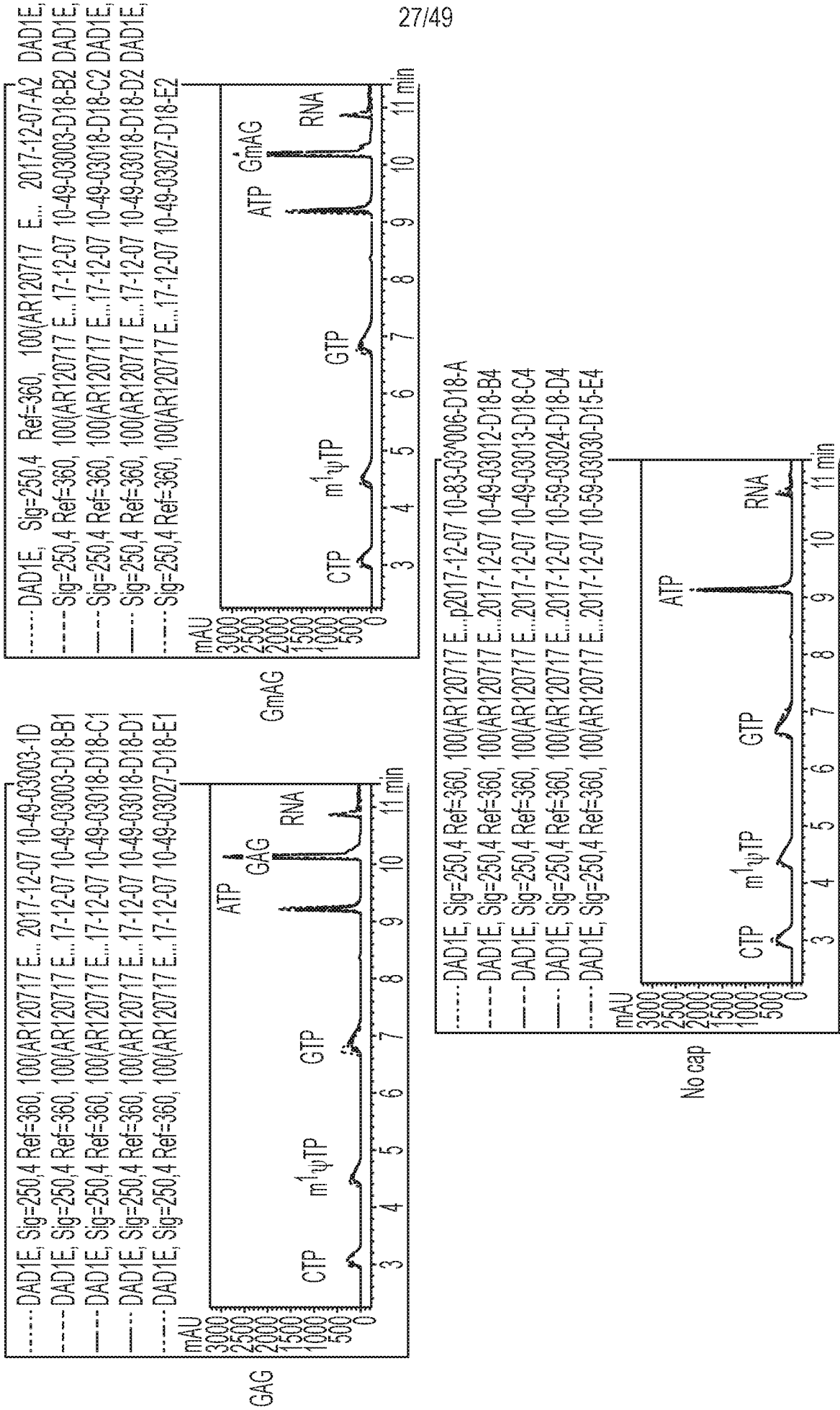


Figure 20B

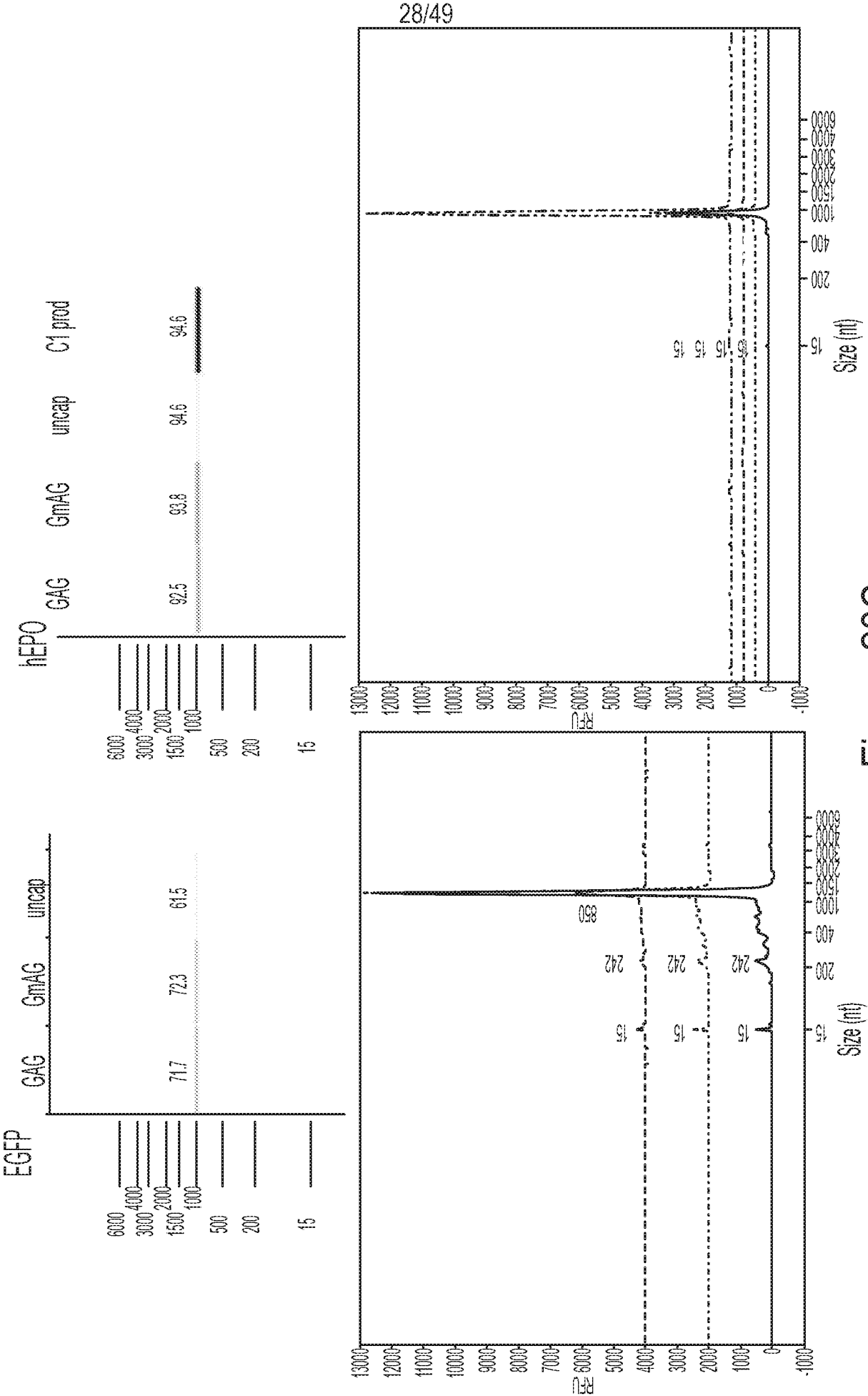


Figure 20C

29/49

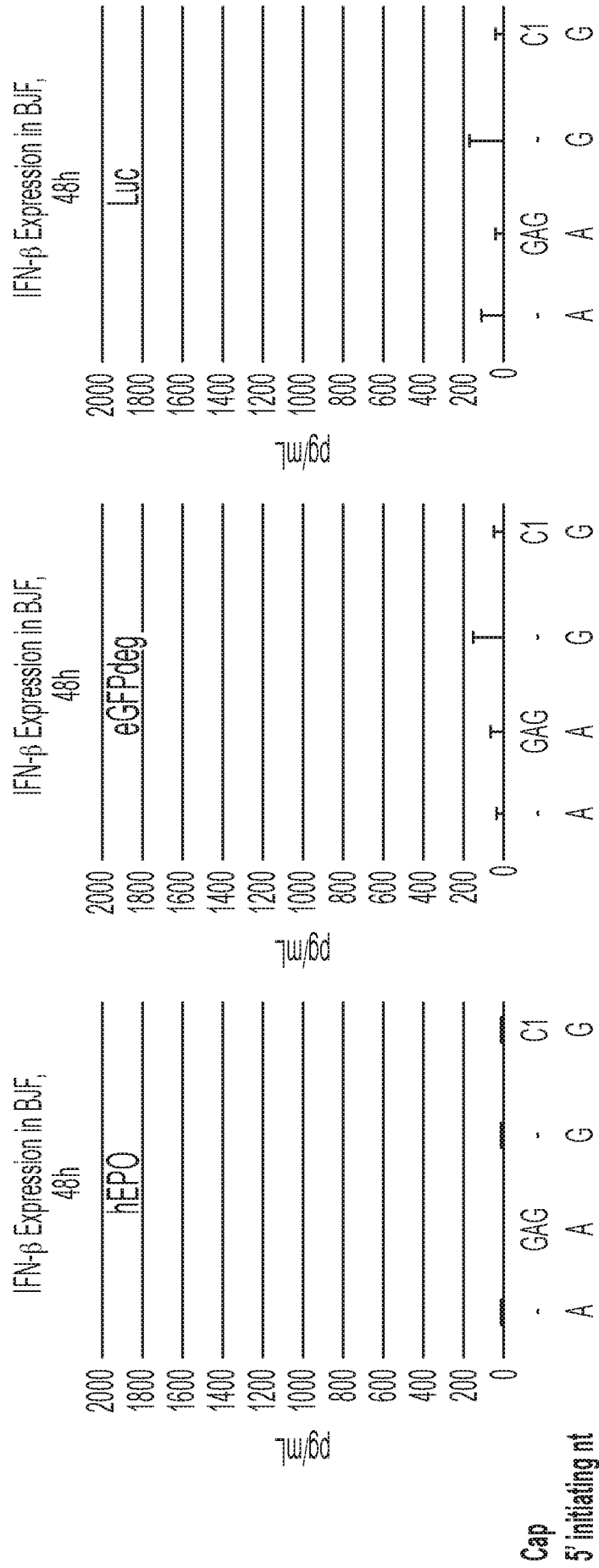


Figure 20D

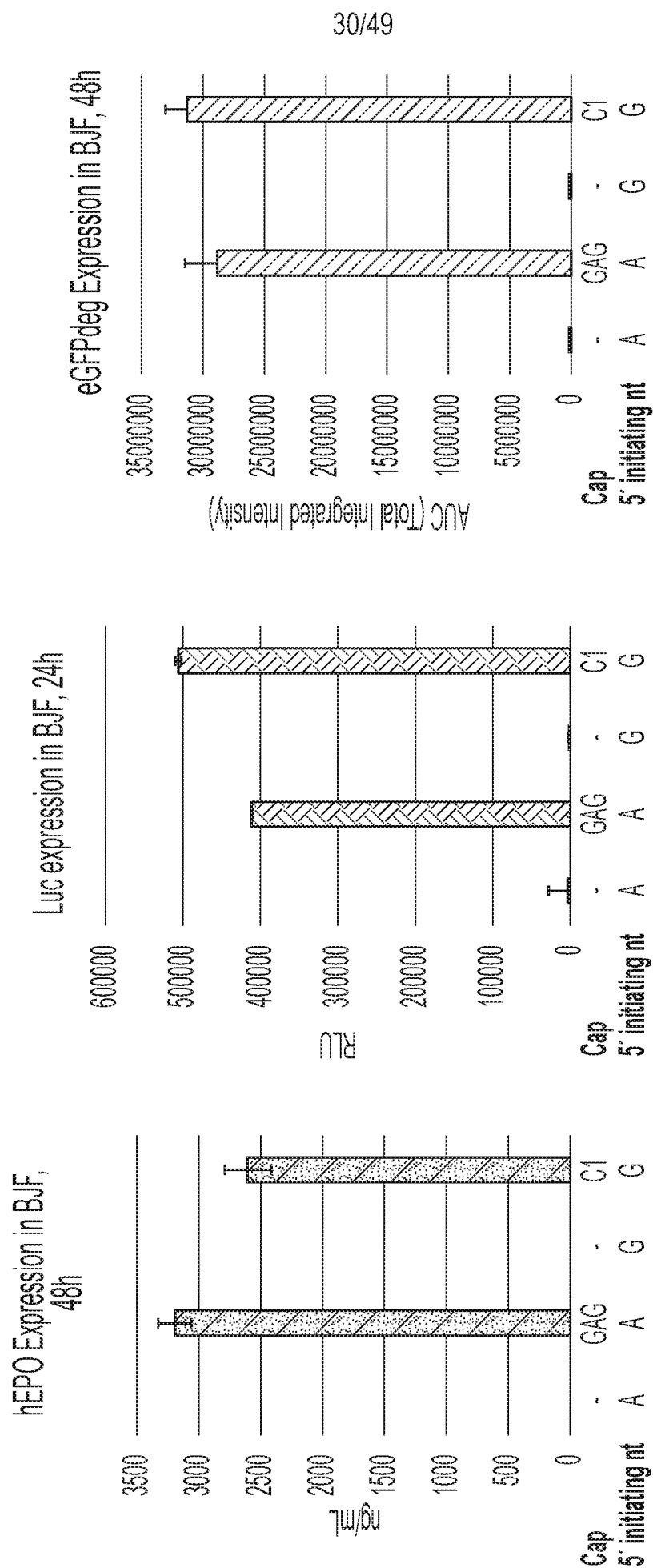


Figure 20E

Figure 20F

Figure 20G

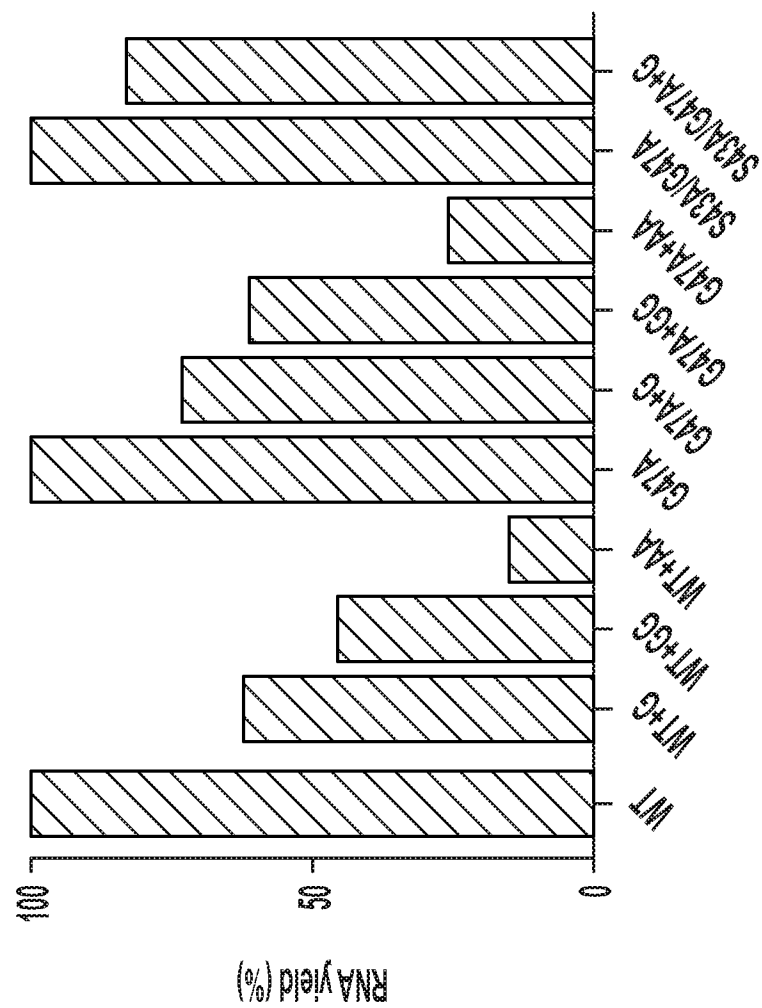


Figure 22

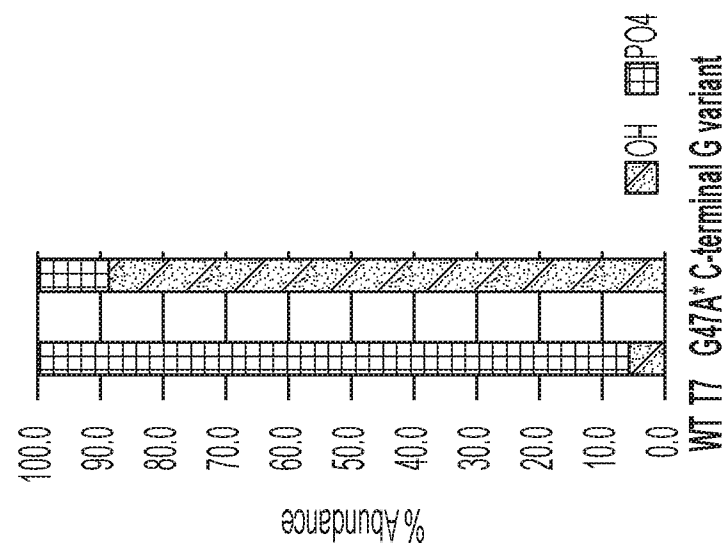


Figure 21

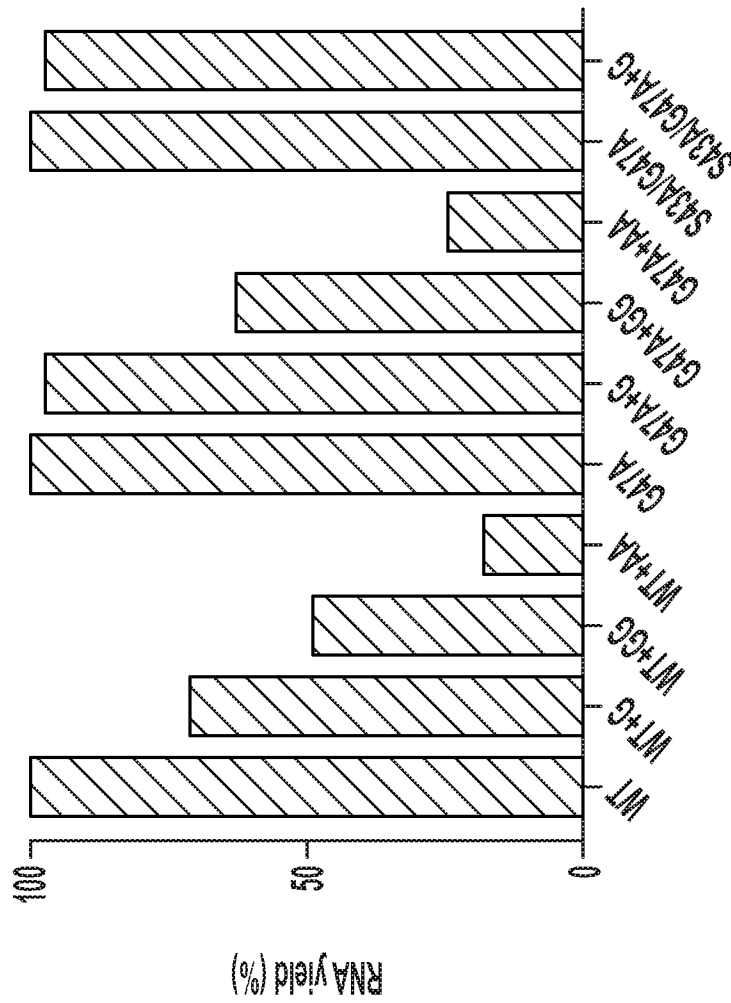


Figure 23

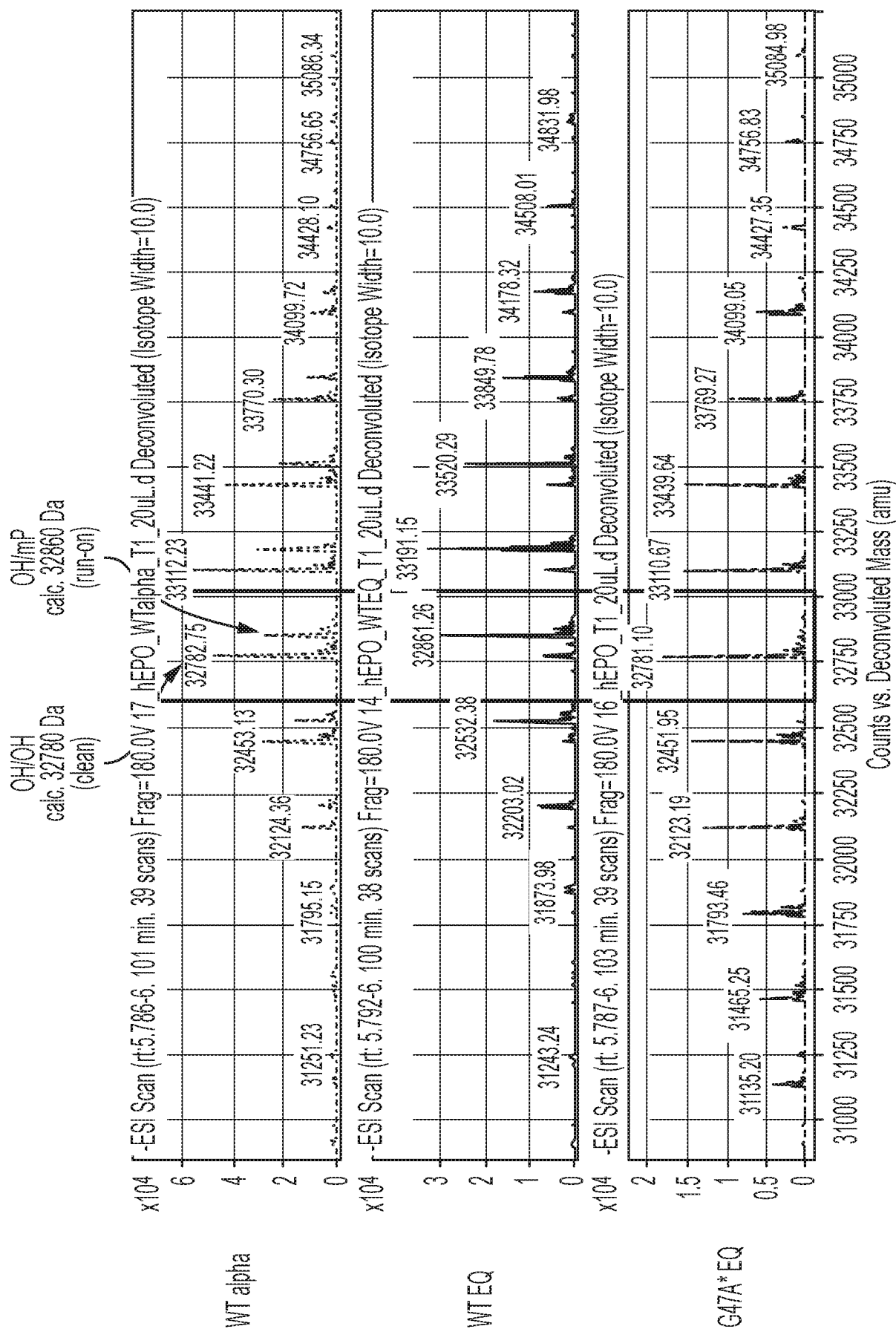


Figure 24A

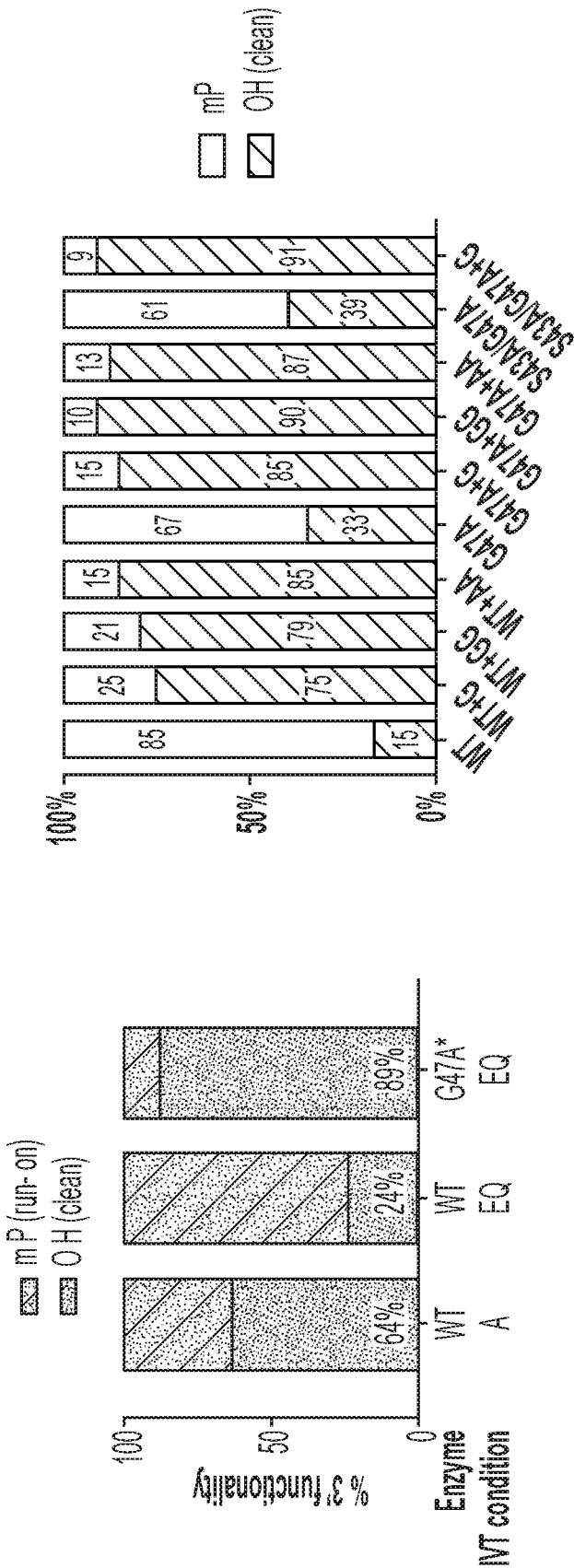


Figure 24C

Figure 24B

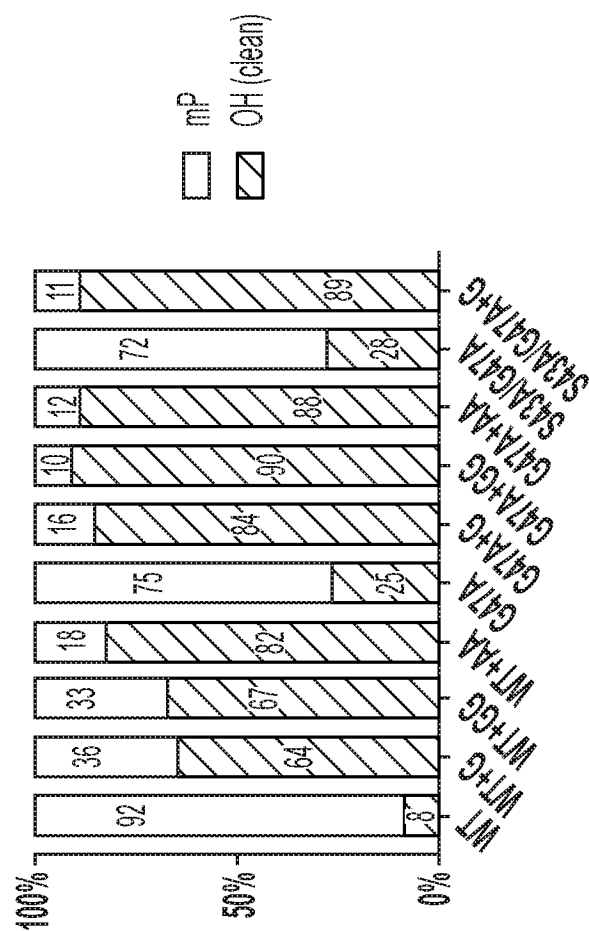


Figure 24D

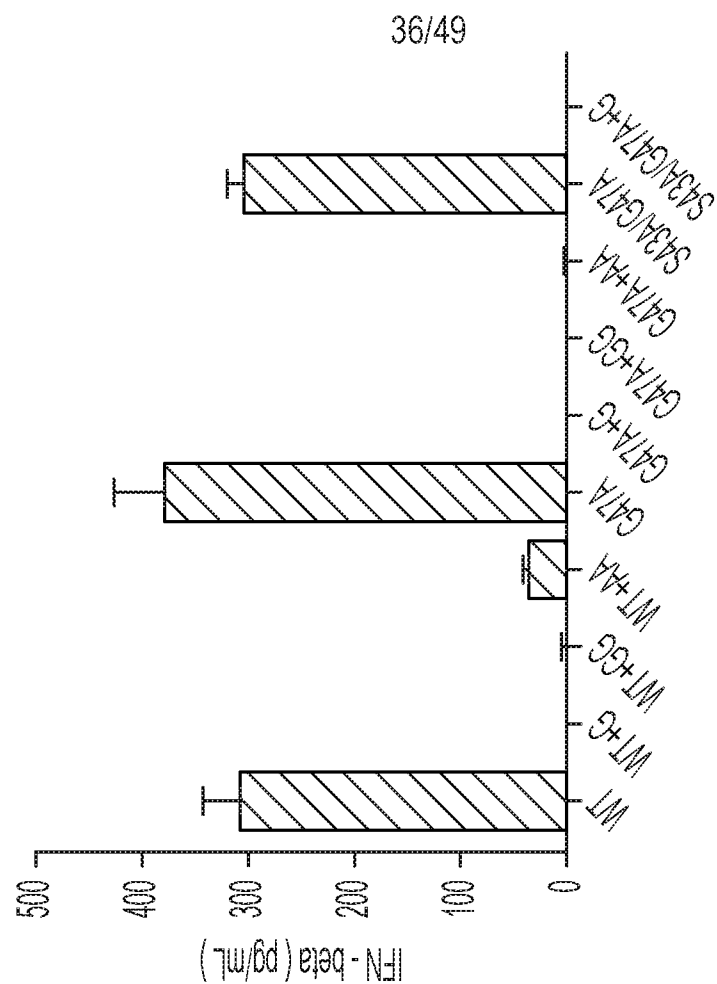


Figure 25B

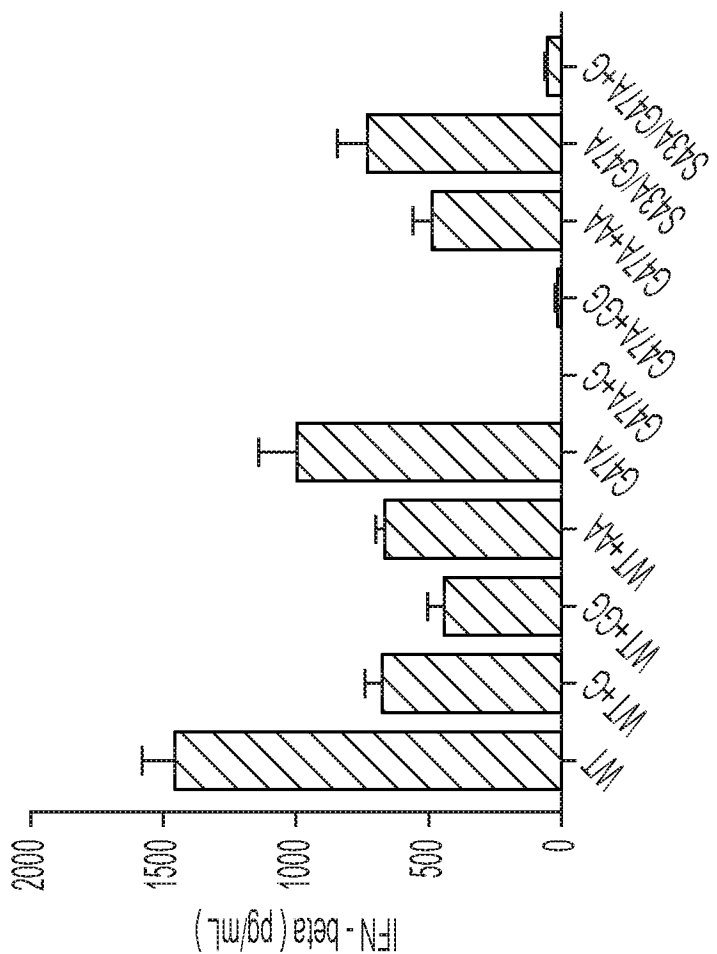


Figure 25A

37/49

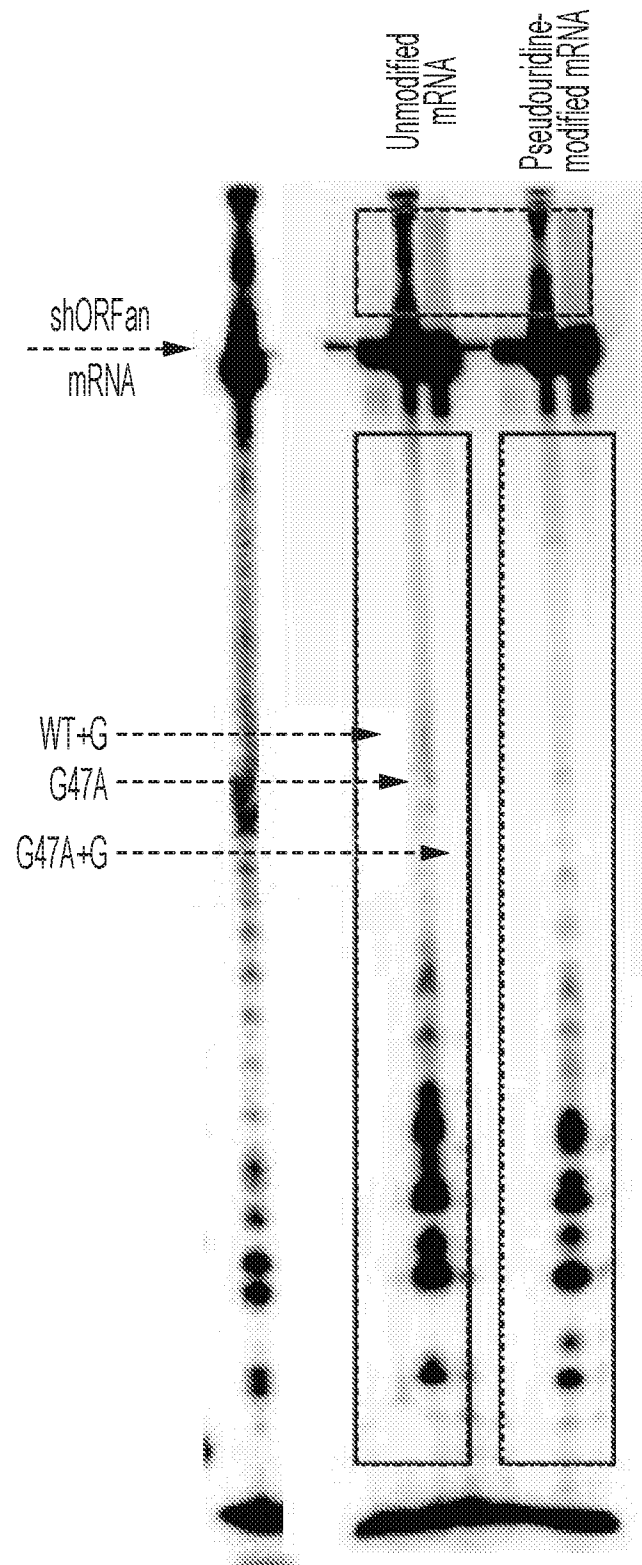


Figure 26A

38/49

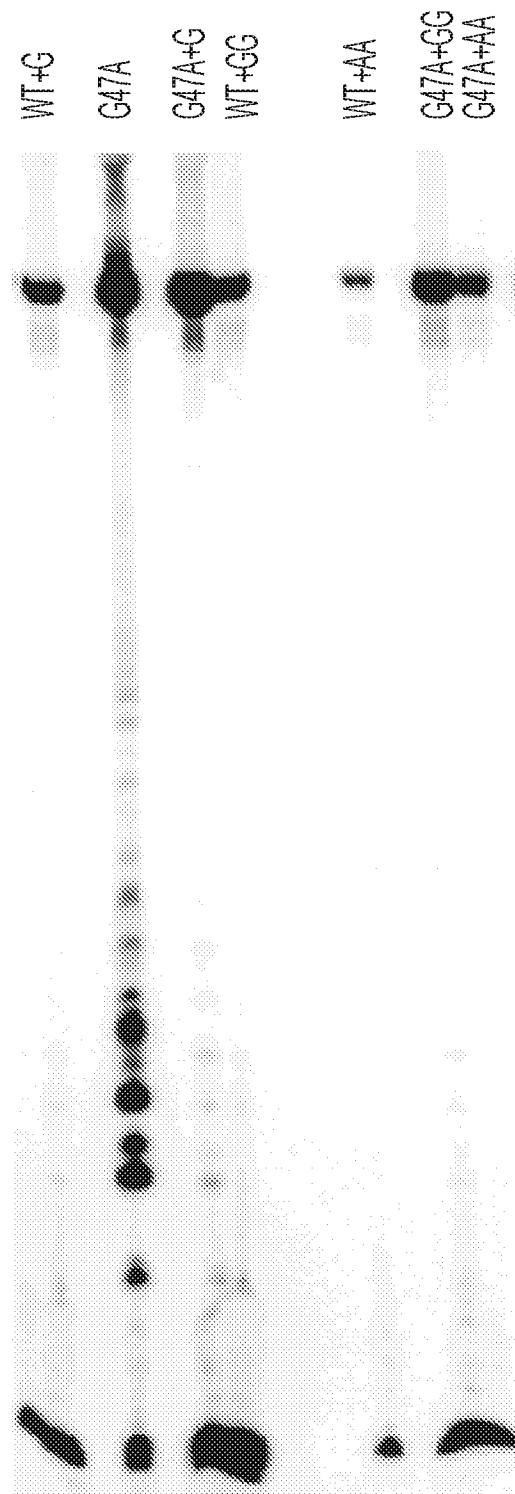
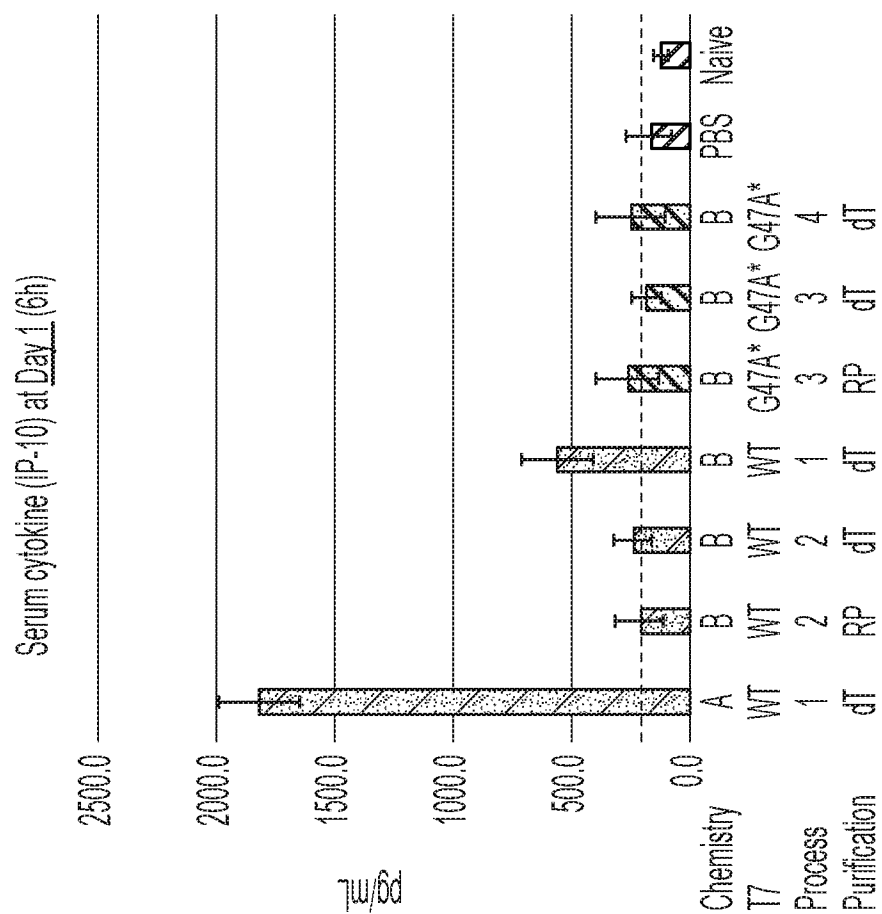
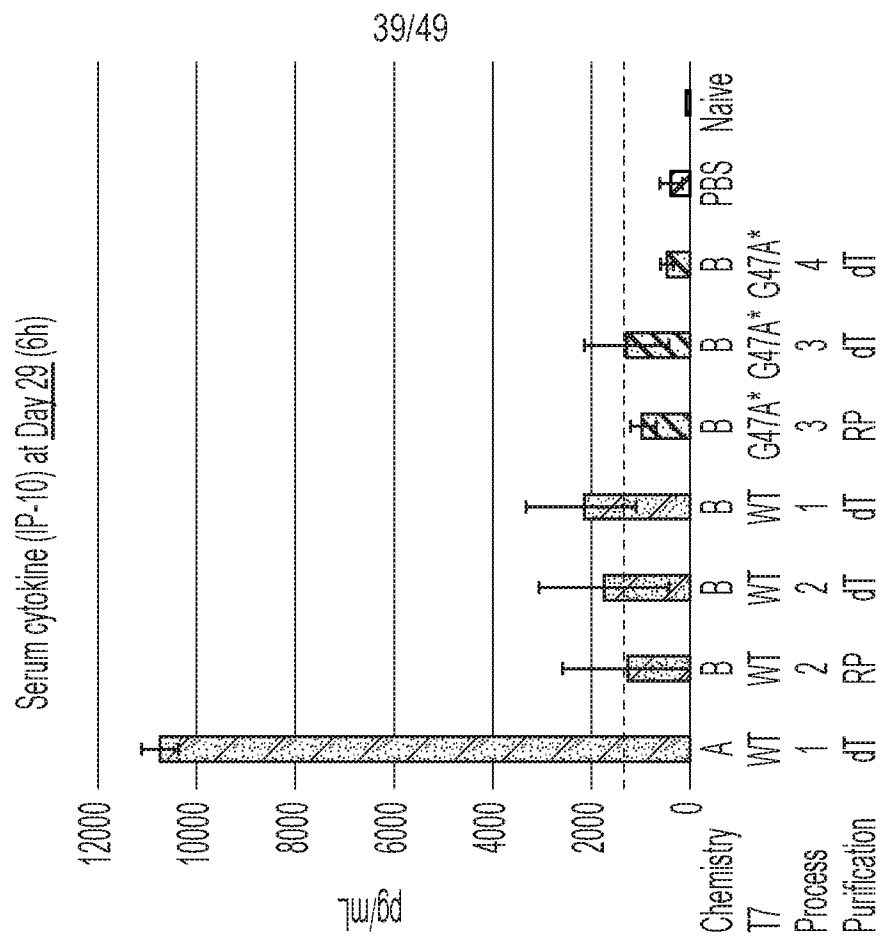


Figure 26B



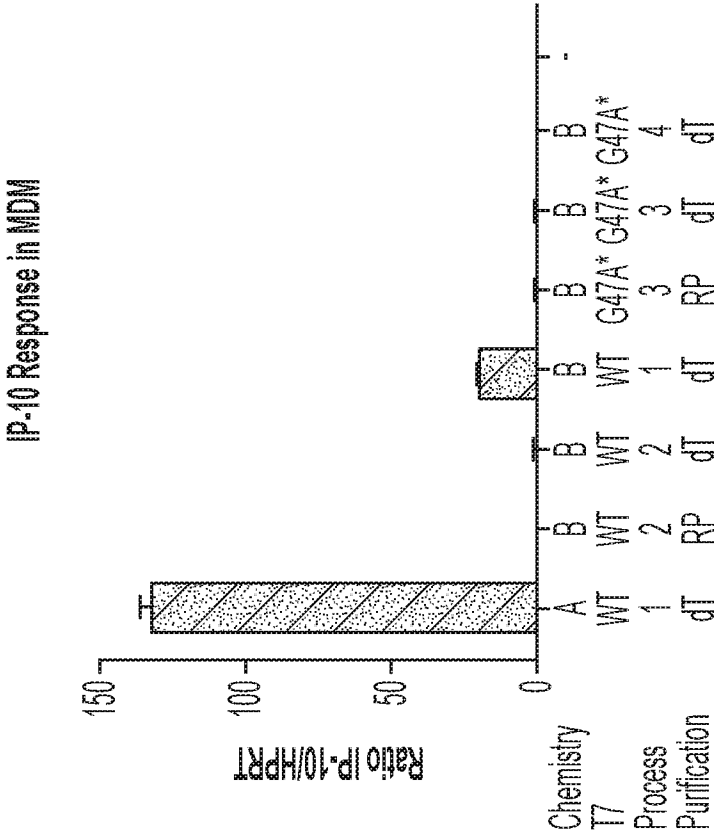


Figure 28B

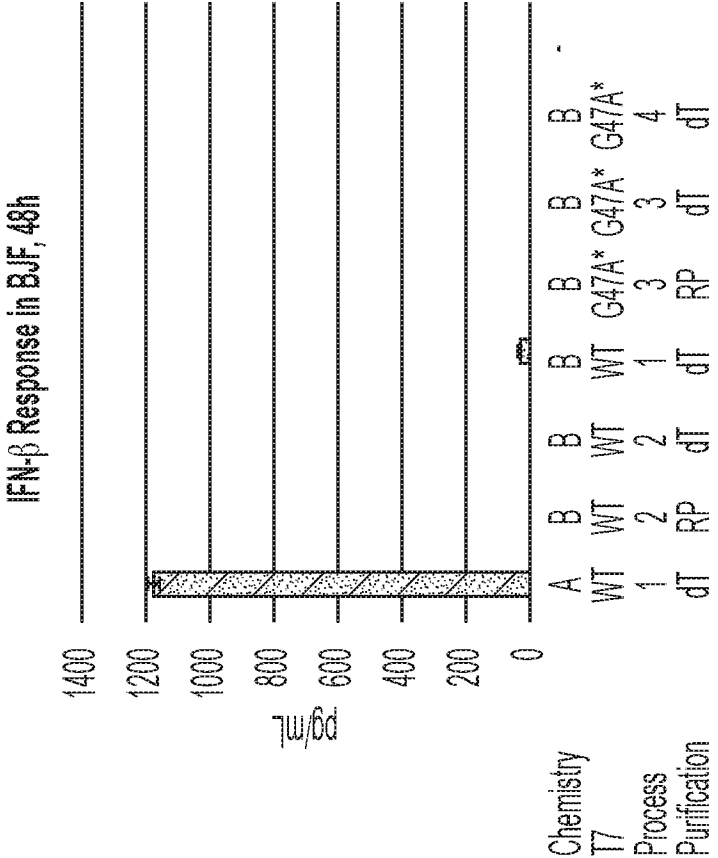


Figure 28A

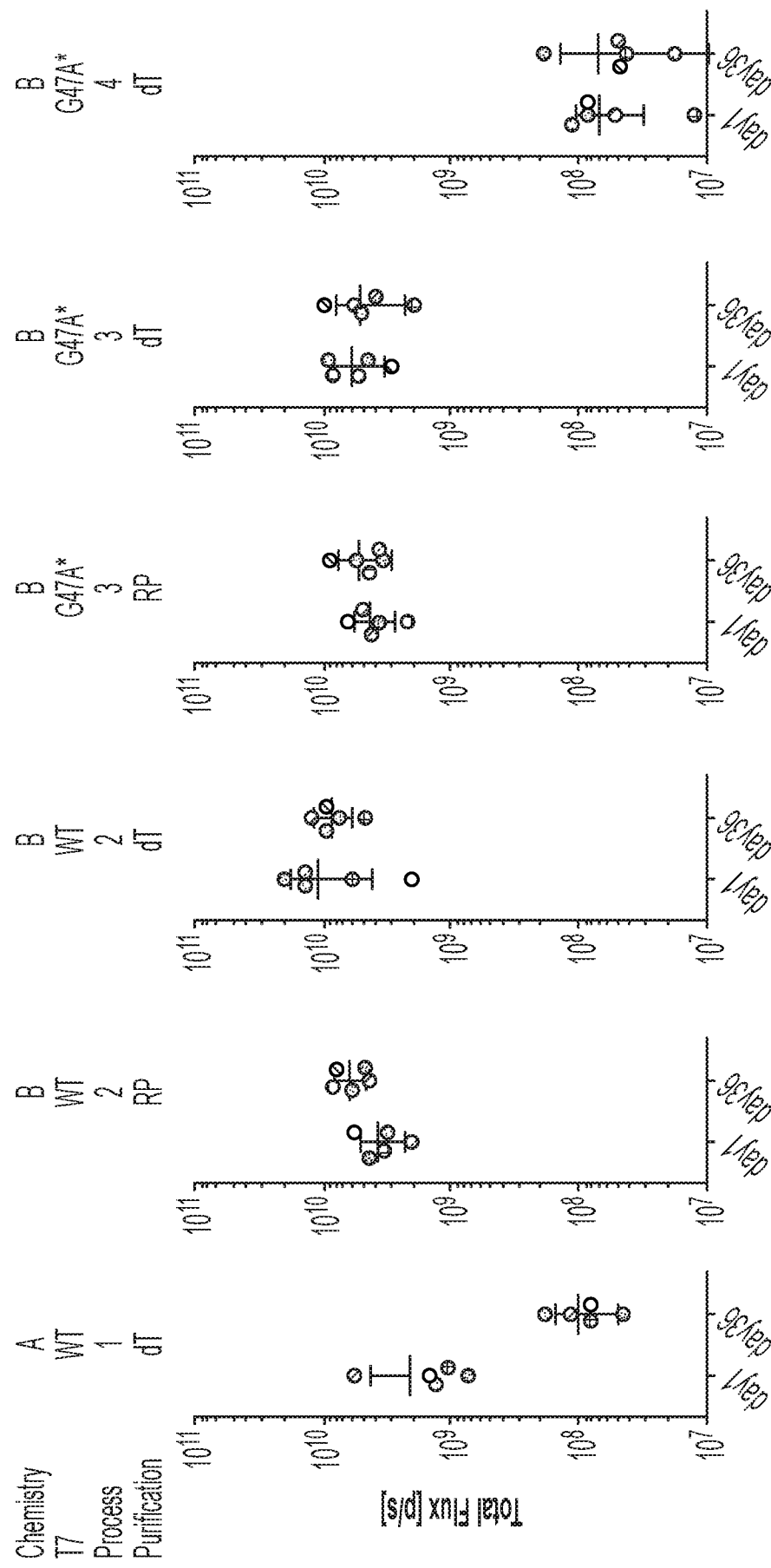


Figure 29

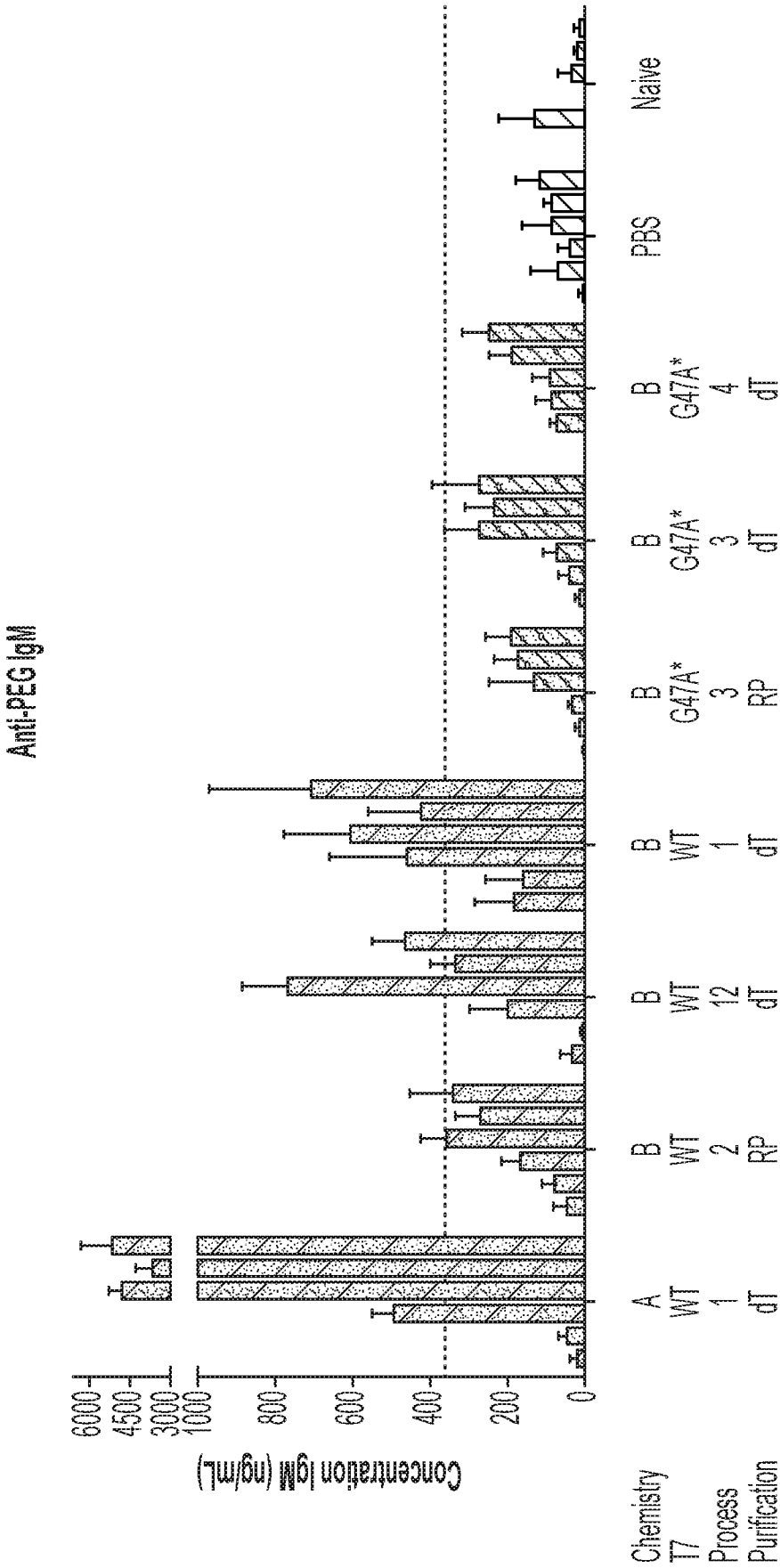


Figure 30

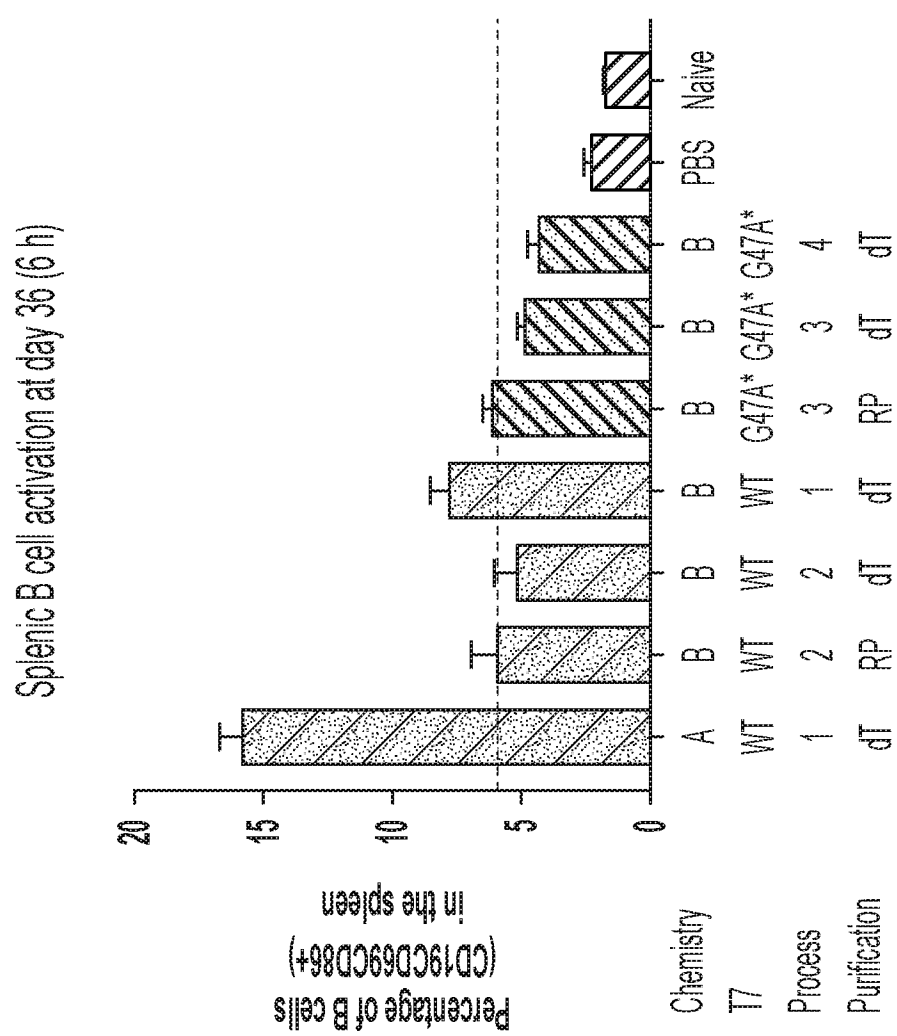


Figure 3

44/49

Serum cytokine (IP-10) at Day 1

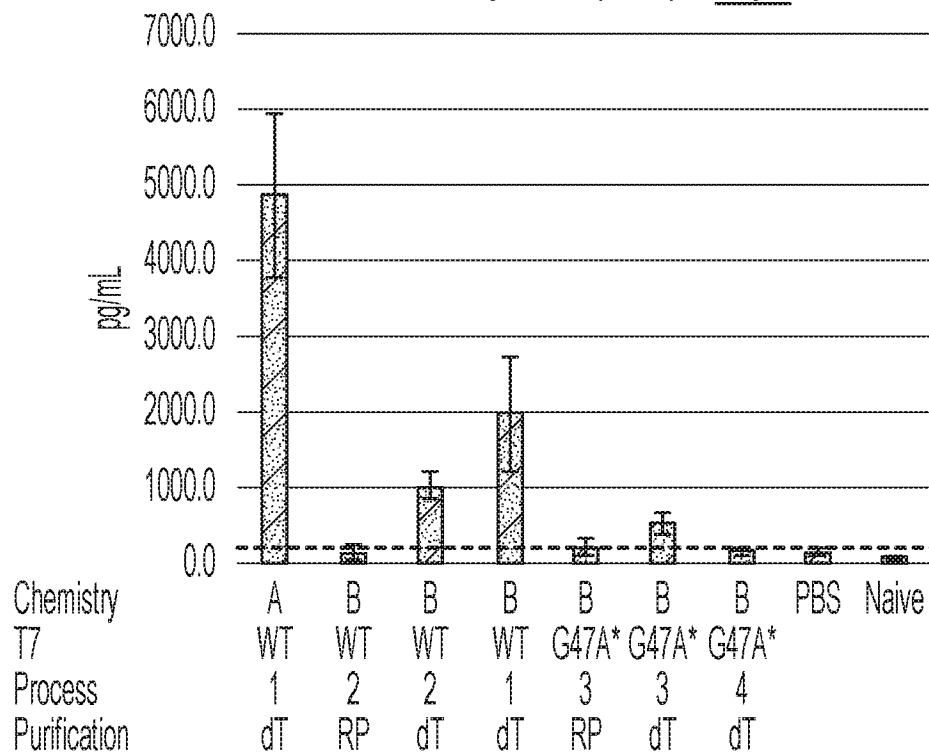


Figure 32A

Serum cytokine (IP-10) at Day 22

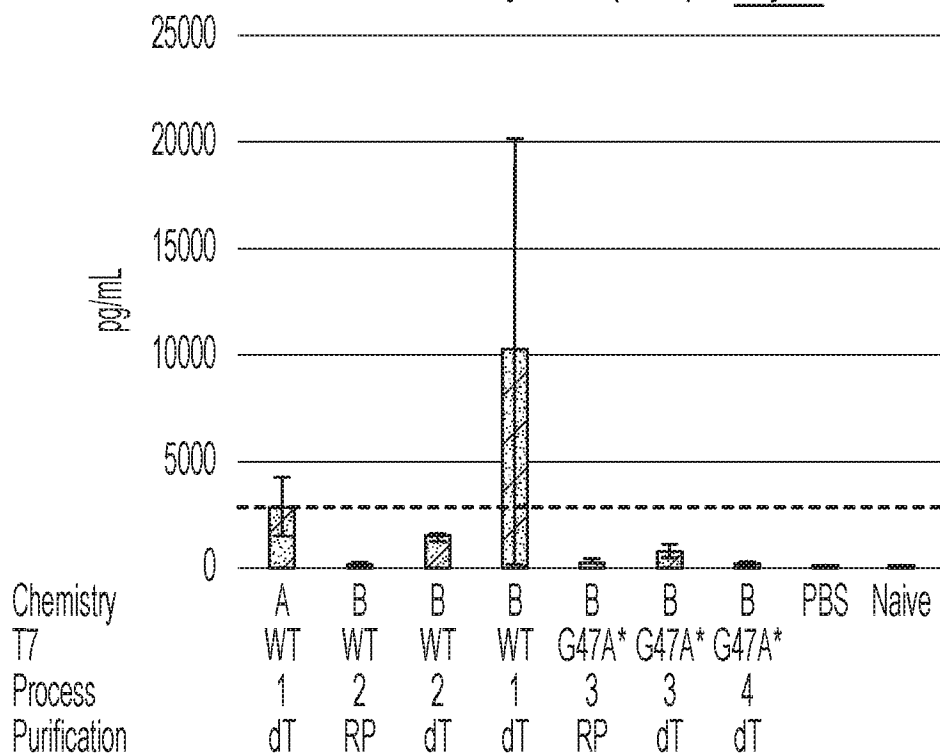


Figure 32B

45/49

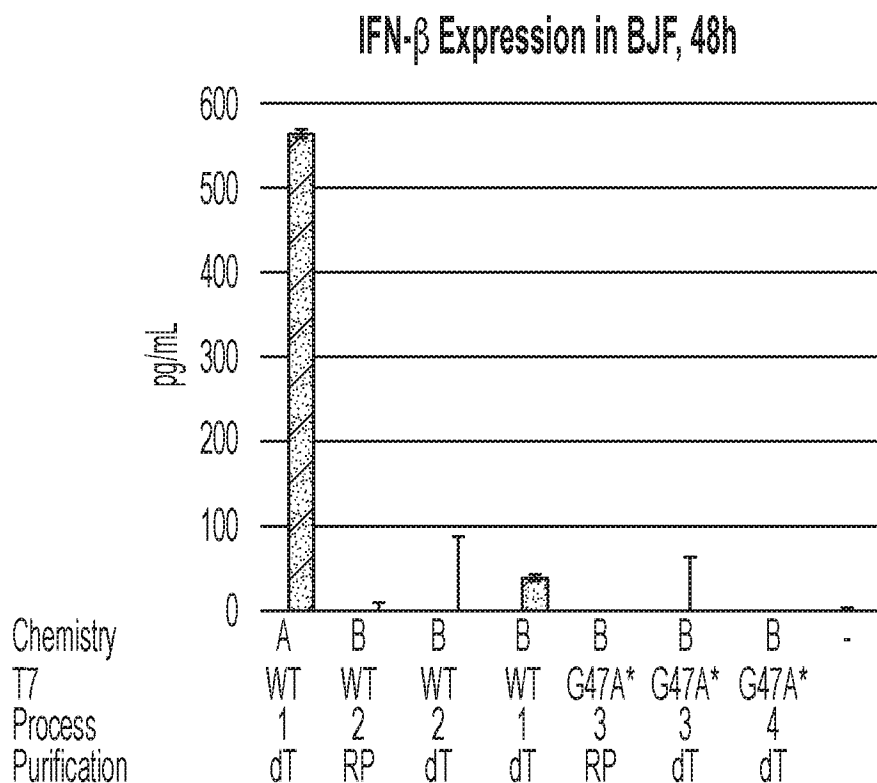


Figure 33A

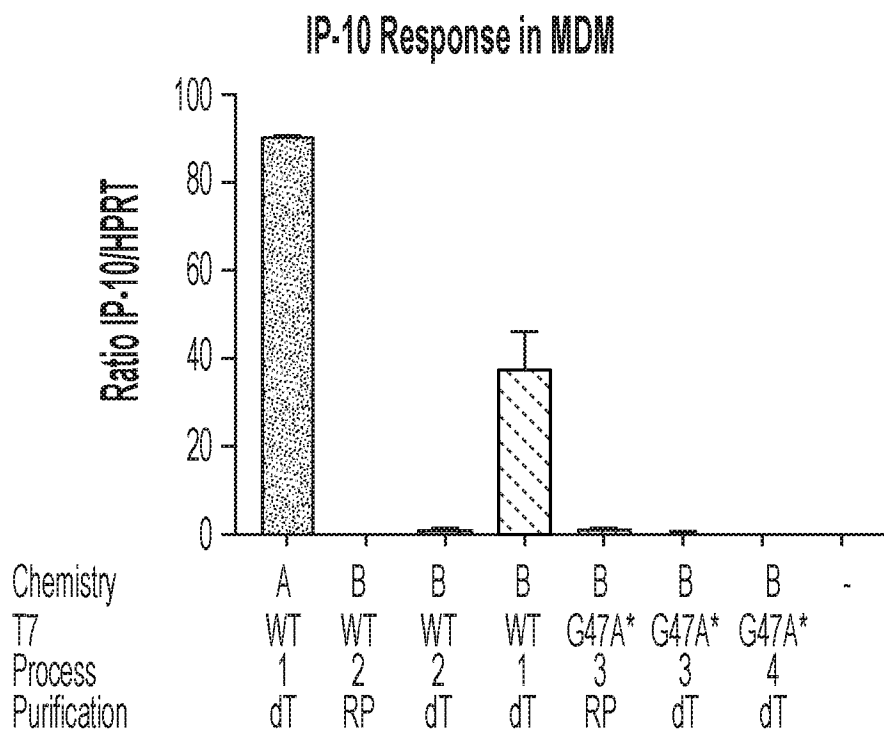


Figure 33B

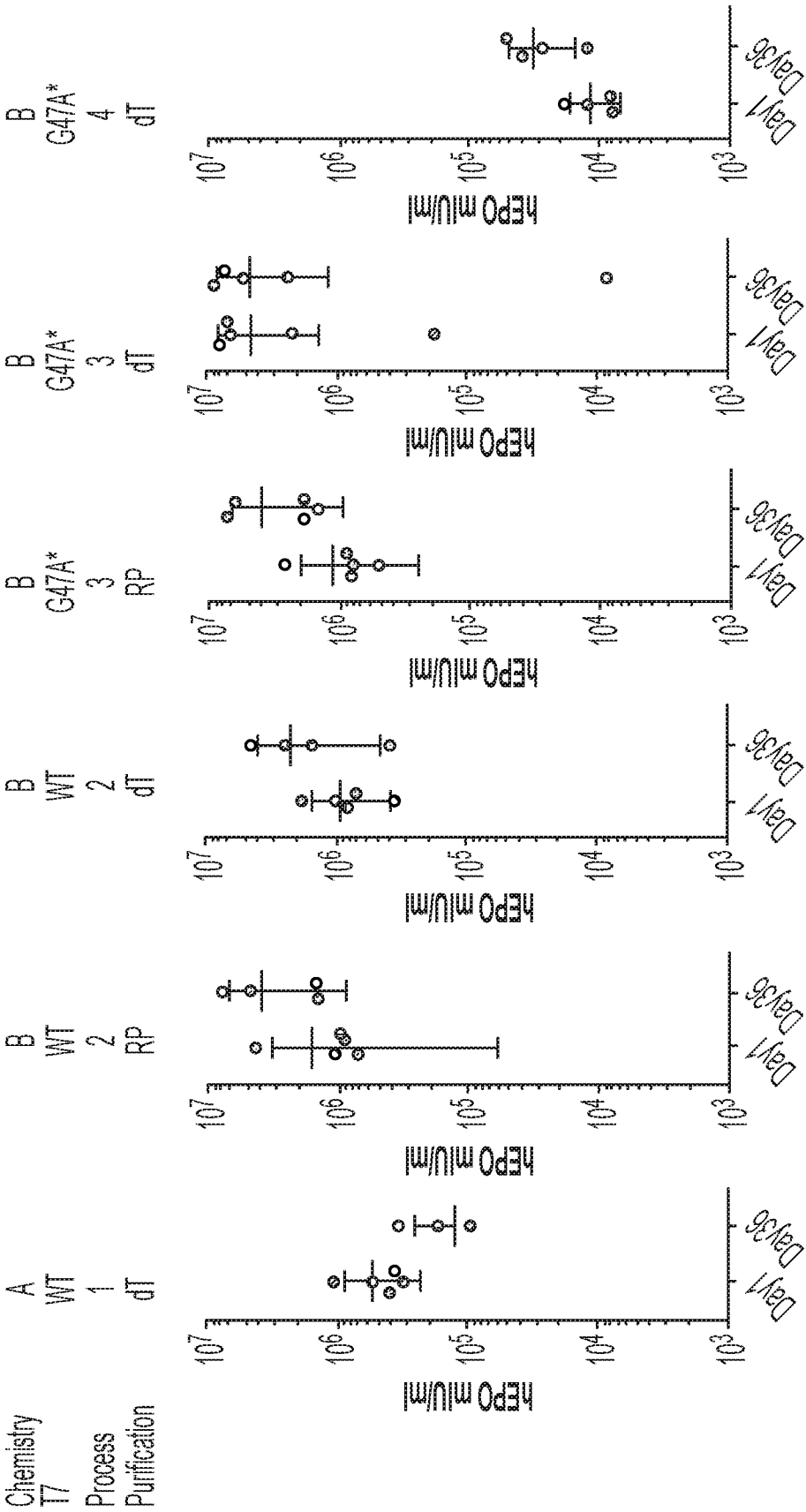


Figure 34

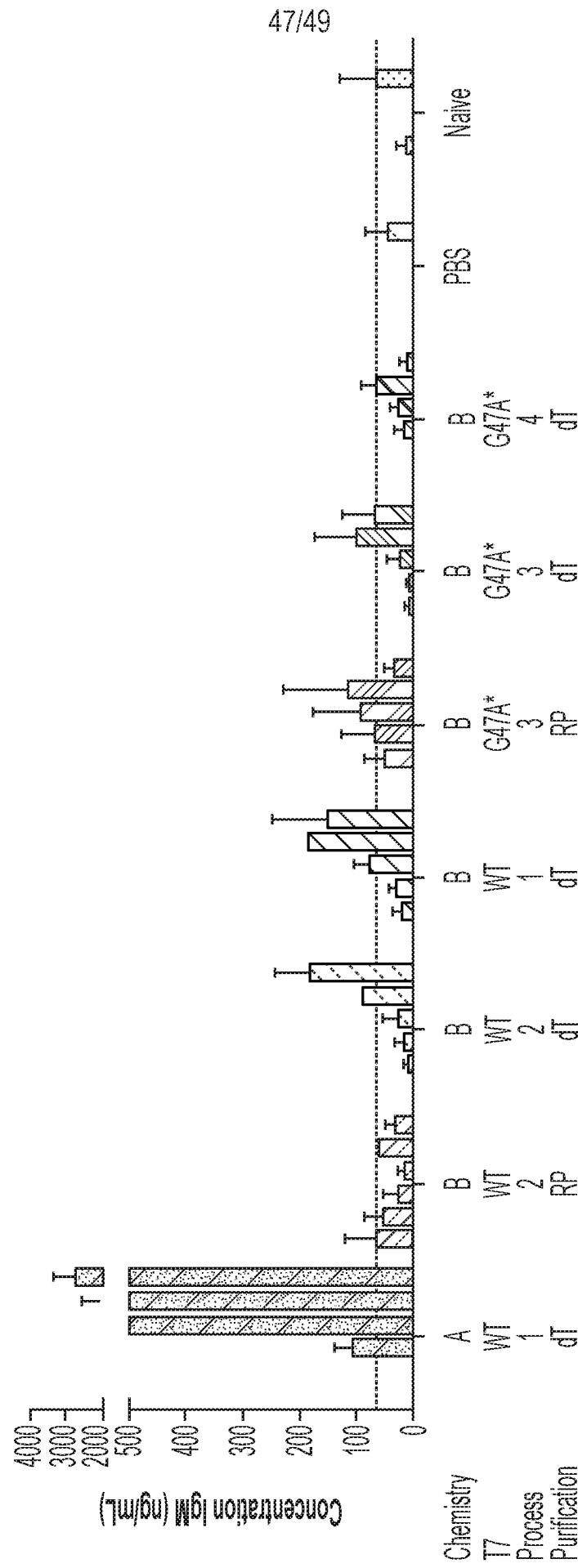


Figure 35

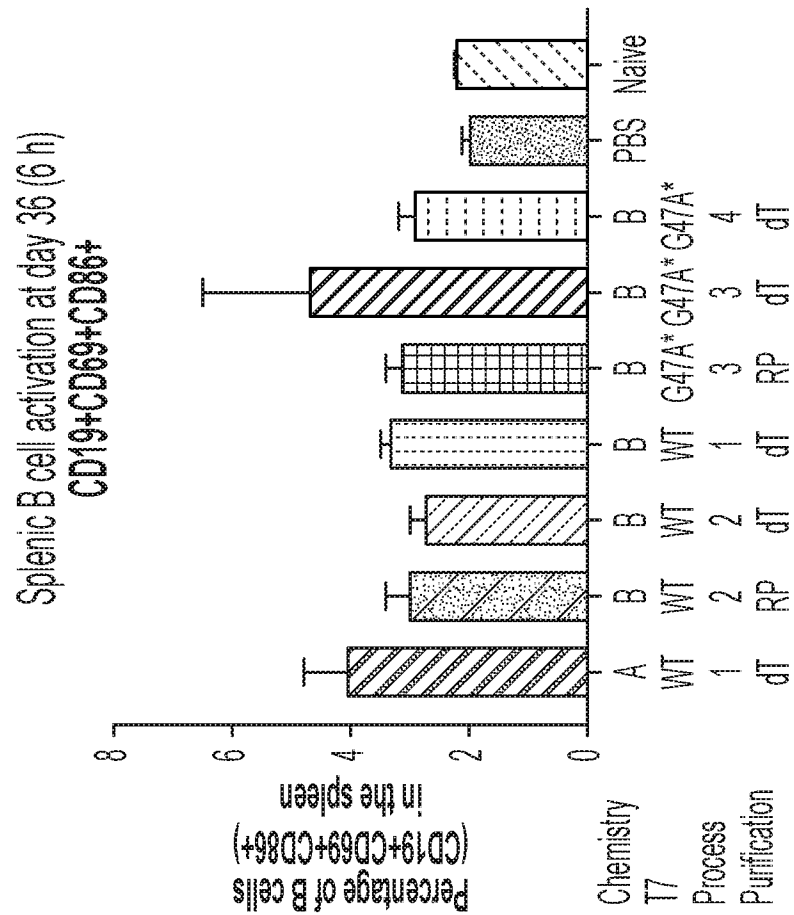


Figure 36

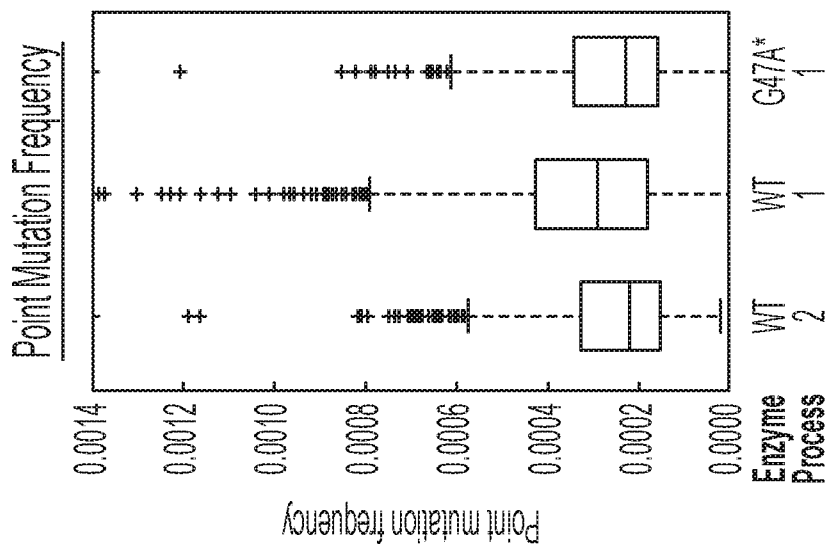


Figure 37B

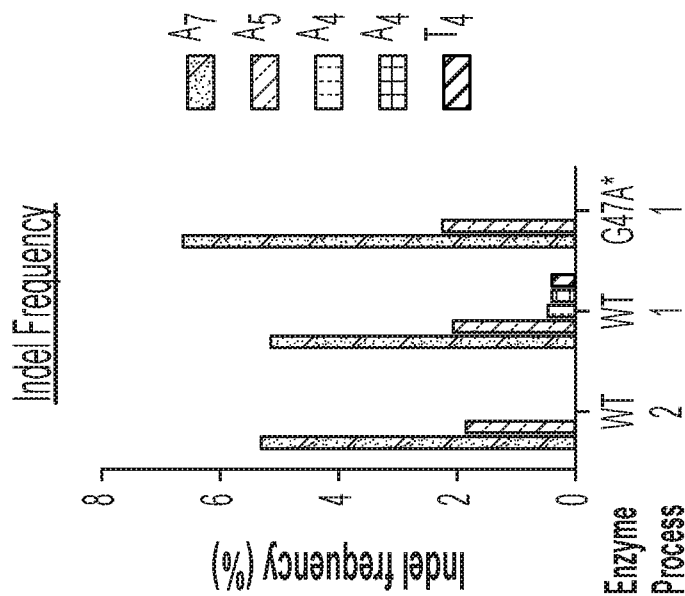


Figure 37A

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2018/046989

A. CLASSIFICATION OF SUBJECT MATTER

C12N 9/12 (2006.01) C12N 15/10 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PATENW, CAPLUS, BIOSIS, MEDLINE. Keywords: C12N9/1247, C12Y207/07006, T7 RNA Polymerase, loop, helix, C-helix, C-linker, mutation, substitution, variant, modify, conformation, structure, initiation, elongation, cap, capping, capped, S43, G47, cytokine, immune, pro-inflammatory, immunogen, immunestimulatory, and similar terms.

GENOMEQUEST: SEQ ID NOS: 1, 6, 7, 8, 108 & 110

Applicant/Inventors name searched in Google, Medline, ESP@CE and internal databases provided by IP Australia.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	



Further documents are listed in the continuation of Box C



See patent family annex

* "A"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E"	earlier application or patent but published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means	"&"	document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 18 December 2018		Date of mailing of the international search report 18 December 2018	
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaustalia.gov.au		Authorised officer Anita Cochrane AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. +61262832118	

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☐ in the form of an Annex C/ST.25 text file.
 - ☒ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

A sequence listing compiling with C/ST.25 was not filed, however the sequences depicted in Tables 1 and 4 (SEQ ID NOs: 1, 6, 7, 8, 108 and 110) were used for the purposes of the search and opinion.

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/US2018/046989
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	WO 2017/123748 A1 (NEW ENGLAND BIOLABS, INC.) 20 July 2017 Abstract, Figure 1, page 1, page 16, SEQ ID NO: 59 As above	47 1-14 (in part), 15-16 (in full) and 43-46 (in part)
X A	TANG G-Q et al, "Relaxed Rotational and Scrunching Changes in P266L Mutant of T7 RNA Polymerase Reduce Short Abortive RNAs while Delaying Transition into Elongation", PLOS One, 2014, 9: e91859 Abstract As above	47 1-14 (in part), 15-16 (in full) and 43-46 (in part)
X A	BANDWAR RP et al, "The Transition to an Elongation Complex by T7 RNA Polymerase Is a Multistep Process", The Journal of Biological Chemistry, 2007, 282: 22879-22886 Abstract, Table 1 As above	47 1-14 (in part), 15-16 (in full) and 43-46 (in part)
Y A	GARDNER LP et al, "Initiation, Elongation, and Processivity of Carboxyl-Terminal Mutants of T7 RNA Polymerase", Biochemistry, 1997, 36: 2908-2918 Abstract As above	47 1-14 (in part), 15-16 (in full) and 43-46 (in part)
X Y	JP 2011-223982 A (TOSOH CORPORATION et al) 10 November 2011 Abstract, SEQ ID NO: 18, 21, 23, 27, 62, 68 paragraphs [0103], [0114], [0131], [0136], [0146], [0162], [0184] As above	17-21, 23-24, 26-29, 33-34, 37-46 52
X	WO 2017/070624 A1 (MODERNATX, INC.) 27 April 2017 pages 23, 145	47-48
X Y	WO 2013/102203 A1 (CELLSCRIPT, INC.) 04 July 2013 pages 23, 33, 107, 109 As above	47-51, 53-62, 64, 66-70 52
X Y	WO 2017/066789 A1 (MODERNATX, INC.) 20 April 2017 pages 28, 64, 122, Table 8 As above	47-51, 53-62, 64, 66-70 52
X A	MA K et al, "Probing conformational changes in T7 RNA polymerase during initiation and termination by using engineered disulphide linkages", PNAS, 2005, 102: 17612-17617 Abstract, page 17612 As above	47 63-64
X A	BANDWAR, RP et al, "Sequential Release of Promoter Contacts during Transcription Initiation to Elongation Transition", J Mol Biol, 2006, 360: 466-482 Abstract As above	47 63-64
X	WO 2017/053297 A1 (TRILINK BIOTECHNOLOGIES, INC.) 30 March 2017 Abstract, pages 27, 53, paragraphs [00142], [00151], [00181], Examples 14, 15, 17, 18, Claim 18	47, 66-70

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/US2018/046989
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	As above	63-64
X	US 2011/0143397 A1 (KARIKO et al) 16 June 2011 paragraphs [0036], [0037], [0094], [0127], [0244], Figure 17A	47-48, 66-70
A	BRIEBA LG et al, "Scanning Mutagenesis Reveals Roles for Helix N of the Bacteriophage T7 RNA Polymerase Thumb Subdomain in Transcription Complex Stability, Pausing, and Termination", The Journal of Biological Chemistry, 2001, 276: 10306-10313 Abstract	
A	GAAL T et al, "DNA-binding determinants of the α subunit of RNA polymerase: novel DNA-binding domain architecture", Genes & Development, 1996, 10: 16-26 Abstract	
P,X	WO 2018/053209 A1 (MODERNATX, INC.) 22 March 2018 pages 29-33, 45-46, 89-90; claims	65, 67-71
P,X	US 10034951 B1 (NEW ENGLAND BIOLABS, INC.) 31 July 2018 Figures 5A, 5B; Claim 1	65, 67-71

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
the subject matter listed in Rule 39 on which, under Article 17(2)(a)(i), an international search is not required to be carried out, including
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See Supplemental Box for Details

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☒ No protest accompanied the payment of additional search fees.

Supplemental Box**Continuation of: Box III**

This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

This Authority has found that there are different inventions based on the following features that separate the claims into the following general concepts:

Concept 1: an RNA polymerase with a mutation that causes conformational change in a loop structure. This reads on independent claim 1

Concept 2: an RNA polymerase with additions to its C-terminus. This reads on independent claims 17, 24, 27, 28.

Concept 3: an RNA produced by the method of claims 33-42. This reads on independent claim 47. Under the claim interpretation standards of the PCT, such a product is not limited to the particular process and therefore does not necessarily import the features of the process.

Concept 4: an in vitro co-transcriptional RNA capping method. This reads on independent claim 49.

Concept 5: the combination of some aspects of concepts 1 and 4. This reads on independent claim 63.

Concept 6: an unpurified pharmaceutical composition comprising in-vitro transcribed RNA that is substantially free of cytokine-inducing RNA contaminant. This reads on independent claim 65.

Concept 7: a pharmaceutical composition comprising in-vitro transcribed RNA that comprises at least 95% capped RNA species. This reads on independent claim 66.

Concept 8: the combination of some aspects of concepts 1 and 2. This reads on independent claims 72 and 75.

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

When there is no special technical feature common to all the claimed inventions there is no unity of invention.

In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a priori*.

It is further noted that concept 1 lacks unity internally - the specification discloses a number of mutations that are said to affect the conformation of T7 RNA polymerase. However the concept of mutations to RNA polymerase is well traversed, see e.g. Briebe et al J. Biol Chem, 2001, Vol. 276, pp. 10306–10313 which discloses alanine scanning mutagenesis in T7 polymerase, and Gaal et al Genes Dev, 1996, Vol. 10 pp. 16-26 which discloses alanine scanning mutagenesis of the C-terminal domain of E. coli RNA polymerase. Therefore each of the mutants disclosed in Table 3 represents a separate invention *a posteriori*.

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
WO 2017/123748 A1	20 July 2017	WO 2017123748 A1	20 Jul 2017
		CN 108779446 A	09 Nov 2018
		EP 3402881 A1	21 Nov 2018
		US 2017247670 A1	31 Aug 2017
JP 2011-223982 A	10 November 2011	JP 2011223982 A	10 Nov 2011
WO 2017/070624 A1	27 April 2017	WO 2017070624 A1	27 Apr 2017
		AR 106463 A1	17 Jan 2018
		AR 106464 A1	17 Jan 2018
		AR 106465 A1	17 Jan 2018
		AU 2016341309 A1	07 Jun 2018
		AU 2016341311 A1	07 Jun 2018
		AU 2016342045 A1	07 Jun 2018
		AU 2016342048 A1	07 Jun 2018
		AU 2016342049 A1	07 Jun 2018
		AU 2016342371 A1	07 Jun 2018
		AU 2016342376 A1	07 Jun 2018
		BR 112018008051 A2	13 Nov 2018
		BR 112018008078 A2	13 Nov 2018
		BR 112018008090 A2	13 Nov 2018
		BR 112018008102 A2	06 Nov 2018
		CA 3002819 A1	27 Apr 2017
		CA 3002820 A1	27 Apr 2017
		CA 3002822 A1	27 Apr 2017
		CA 3002912 A1	27 Apr 2017
		CA 3002922 A1	27 Apr 2017
		CA 3003090 A1	27 Apr 2017
		CA 3003103 A1	27 Apr 2017
		CN 108430456 A	21 Aug 2018
		CN 108472309 A	31 Aug 2018
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			
Form PCT/ISA/210 (Family Annex)(January 2015)			

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
		CN 108472354 A	31 Aug 2018
		CN 108472355 A	31 Aug 2018
		CO 2018005229 A2	30 Nov 2018
		CO 2018005258 A2	22 Nov 2018
		CO 2018005265 A2	22 Nov 2018
		EA 201891001 A1	30 Nov 2018
		EP 3364949 A1	29 Aug 2018
		EP 3364950 A1	29 Aug 2018
		EP 3364980 A1	29 Aug 2018
		EP 3364981 A1	29 Aug 2018
		EP 3364982 A2	29 Aug 2018
		EP 3364983 A2	29 Aug 2018
		EP 3365007 A2	29 Aug 2018
		EP 3365008 A1	29 Aug 2018
		EP 3365009 A1	29 Aug 2018
		JP 2018531290 A	25 Oct 2018
		JP 2018531996 A	01 Nov 2018
		JP 2018532777 A	08 Nov 2018
		JP 2018536022 A	06 Dec 2018
		JP 2018536023 A	06 Dec 2018
		KR 20180094859 A	24 Aug 2018
		KR 20180096591 A	29 Aug 2018
		KR 20180096592 A	29 Aug 2018
		KR 20180096593 A	29 Aug 2018
		PE 15292018 A1	26 Sep 2018
		PE 15302018 A1	26 Sep 2018
		PE 15312018 A1	26 Sep 2018
		PH 12018500855 A1	29 Oct 2018
		PH 12018500856 A1	29 Oct 2018
		PH 12018500858 A1	29 Oct 2018
		SG 11201803360U A	30 May 2018
		SG 11201803363Y A	30 May 2018
		SG 11201803365R A	30 May 2018
		TW 201729835 A	01 Sep 2017
		TW 201729836 A	01 Sep 2017
		TW 201729838 A	01 Sep 2017
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			

INTERNATIONAL SEARCH REPORT Information on patent family members		International application No. PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
		US 2017340725 A1	30 Nov 2017
		US 10064934 B2	04 Sep 2018
		US 2018028645 A1	01 Feb 2018
		US 10064935 B2	04 Sep 2018
		US 2017340724 A1	30 Nov 2017
		US 10124055 B2	13 Nov 2018
		US 2018008694 A1	11 Jan 2018
		US 2018271970 A1	27 Sep 2018
		US 2018271975 A1	27 Sep 2018
		US 2018289792 A1	11 Oct 2018
		US 2018296662 A1	18 Oct 2018
		US 2018303929 A1	25 Oct 2018
		US 2018311336 A1	01 Nov 2018
		US 2018318409 A1	08 Nov 2018
		US 2018326045 A1	15 Nov 2018
		WO 2017070601 A1	27 Apr 2017
		WO 2017070613 A1	27 Apr 2017
		WO 2017070616 A2	27 Apr 2017
		WO 2017070618 A1	27 Apr 2017
		WO 2017070620 A2	27 Apr 2017
		WO 2017070622 A1	27 Apr 2017
		WO 2017070623 A1	27 Apr 2017
		WO 2017070626 A2	27 Apr 2017
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex)(January 2015)			

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
WO 2013/102203 A1	04 July 2013	WO 2013102203 A1	04 Jul 2013
		AU 2010328310 A1	28 Jun 2012
		AU 2010328310 B2	21 May 2015
		AU 2012362113 A1	14 Aug 2014
		AU 2012362113 B2	03 Aug 2017
		AU 2015215938 A1	10 Sep 2015
		AU 2015215938 B2	18 Jan 2018
		AU 2017251767 A1	16 Nov 2017
		AU 2017251767 B2	16 Aug 2018
		AU 2018202479 A1	26 Apr 2018
		AU 2018264152 A1	06 Dec 2018
		BR 112012013875 A2	03 Nov 2015
		CA 2783032 A1	16 Jun 2011
		CA 2862377 A1	04 Jul 2013
		CN 102947450 A	27 Feb 2013
		CN 102947450 B	23 Nov 2016
		CN 107090436 A	25 Aug 2017
		EP 1979364 A2	15 Oct 2008
		EP 2510099 A2	17 Oct 2012
		EP 2510099 B1	09 Aug 2017
		EP 2578685 A2	10 Apr 2013
		EP 2798064 A1	05 Nov 2014
		EP 2798064 B1	31 Aug 2016
		EP 3112467 A1	04 Jan 2017
		EP 3112467 B1	14 Feb 2018
		EP 3144389 A1	22 Mar 2017
		EP 3144389 B1	09 May 2018
		EP 3287525 A1	28 Feb 2018
		HK 1203559 A1	30 Oct 2015
		JP 2013512690 A	18 Apr 2013
		JP 2016171804 A	29 Sep 2016
		KR 20130009944 A	24 Jan 2013
		KR 101878502 B1	13 Jul 2018
		KR 20180081836 A	17 Jul 2018
		SG 181564 A1	30 Jul 2012
		Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.	
Form PCT/ISA/210 (Family Annex)(January 2015)			

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
		SG 10201408129X A	27 Feb 2015
		US 2009286852 A1	19 Nov 2009
		US 8278036 B2	02 Oct 2012
		US 2013111615 A1	02 May 2013
		US 8691966 B2	08 Apr 2014
		US 2013261172 A1	03 Oct 2013
		US 8748089 B2	10 Jun 2014
		US 2011143436 A1	16 Jun 2011
		US 8808982 B2	19 Aug 2014
		US 2013197068 A1	01 Aug 2013
		US 8835108 B2	16 Sep 2014
		US 2011143397 A1	16 Jun 2011
		US 9012219 B2	21 Apr 2015
		US 2015184123 A1	02 Jul 2015
		US 9163213 B2	20 Oct 2015
		US 2015315572 A1	05 Nov 2015
		US 9371511 B2	21 Jun 2016
		US 2014315988 A1	23 Oct 2014
		US 9371544 B2	21 Jun 2016
		US 2015038558 A1	05 Feb 2015
		US 9750824 B2	05 Sep 2017
		US 2016369243 A1	22 Dec 2016
		US 10006007 B2	26 Jun 2018
		US 2013189741 A1	25 Jul 2013
		US 2014328825 A1	06 Nov 2014
		US 2016251629 A1	01 Sep 2016
		US 2017043037 A1	16 Feb 2017
		US 2018265848 A1	20 Sep 2018
		WO 2007024708 A2	01 Mar 2007
		WO 2011071931 A2	16 Jun 2011
		WO 2011071936 A2	16 Jun 2011
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			
Form PCT/ISA/210 (Family Annex)(January 2015)			

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
WO 2017/066789 A1	20 April 2017	WO 2017066789 A1	20 Apr 2017
WO 2017/053297 A1	30 March 2017	WO 2017053297 A1	30 Mar 2017
		AU 2016328645 A1	19 Apr 2018
		CA 2999274 A1	30 Mar 2017
		CN 108366604 A	03 Aug 2018
		EP 3352584 A1	01 Aug 2018
		JP 2018527015 A	20 Sep 2018
		KR 20180050409 A	14 May 2018
		US 2018273576 A1	27 Sep 2018
US 2011/0143397 A1	16 June 2011	US 2011143397 A1	16 Jun 2011
		US 9012219 B2	21 Apr 2015
		AU 2010328310 A1	28 Jun 2012
		AU 2010328310 B2	21 May 2015
		AU 2012362113 A1	14 Aug 2014
		AU 2012362113 B2	03 Aug 2017
		AU 2015215938 A1	10 Sep 2015
		AU 2015215938 B2	18 Jan 2018
		AU 2017251767 A1	16 Nov 2017
		AU 2017251767 B2	16 Aug 2018
		AU 2018202479 A1	26 Apr 2018
		AU 2018264152 A1	06 Dec 2018
		BR 112012013875 A2	03 Nov 2015
		CA 2783032 A1	16 Jun 2011
		CA 2862377 A1	04 Jul 2013
		CN 102947450 A	27 Feb 2013
		CN 102947450 B	23 Nov 2016
		CN 107090436 A	25 Aug 2017
		EP 1979364 A2	15 Oct 2008
		EP 2510099 A2	17 Oct 2012
		EP 2510099 B1	09 Aug 2017
		EP 2578685 A2	10 Apr 2013
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			
Form PCT/ISA/210 (Family Annex)(January 2015)			

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
		EP 2798064 A1	05 Nov 2014
		EP 2798064 B1	31 Aug 2016
		EP 3112467 A1	04 Jan 2017
		EP 3112467 B1	14 Feb 2018
		EP 3144389 A1	22 Mar 2017
		EP 3144389 B1	09 May 2018
		EP 3287525 A1	28 Feb 2018
		HK 1203559 A1	30 Oct 2015
		JP 2013512690 A	18 Apr 2013
		JP 2016171804 A	29 Sep 2016
		KR 20130009944 A	24 Jan 2013
		KR 101878502 B1	13 Jul 2018
		KR 20180081836 A	17 Jul 2018
		SG 181564 A1	30 Jul 2012
		SG 10201408129X A	27 Feb 2015
		US 2009286852 A1	19 Nov 2009
		US 8278036 B2	02 Oct 2012
		US 2013111615 A1	02 May 2013
		US 8691966 B2	08 Apr 2014
		US 2013261172 A1	03 Oct 2013
		US 8748089 B2	10 Jun 2014
		US 2011143436 A1	16 Jun 2011
		US 8808982 B2	19 Aug 2014
		US 2013197068 A1	01 Aug 2013
		US 8835108 B2	16 Sep 2014
		US 2015184123 A1	02 Jul 2015
		US 9163213 B2	20 Oct 2015
		US 2015315572 A1	05 Nov 2015
		US 9371511 B2	21 Jun 2016
		US 2014315988 A1	23 Oct 2014
		US 9371544 B2	21 Jun 2016
		US 2015038558 A1	05 Feb 2015
		US 9750824 B2	05 Sep 2017
		US 2016369243 A1	22 Dec 2016
		US 10006007 B2	26 Jun 2018
		US 2013189741 A1	25 Jul 2013
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			
Form PCT/ISA/210 (Family Annex)(January 2015)			

INTERNATIONAL SEARCH REPORT Information on patent family members		International application No. PCT/US2018/046989	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
		US 2014328825 A1	06 Nov 2014
		US 2016251629 A1	01 Sep 2016
		US 2017043037 A1	16 Feb 2017
		US 2018265848 A1	20 Sep 2018
		WO 2007024708 A2	01 Mar 2007
		WO 2011071931 A2	16 Jun 2011
		WO 2011071936 A2	16 Jun 2011
		WO 2013102203 A1	04 Jul 2013
WO 2018/053209 A1	22 March 2018	WO 2018053209 A1	22 Mar 2018
US 10034951 B1	31 July 2018	US 10034951 B1	31 Jul 2018
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex)(January 2015)			