



(51) International Patent Classification:

A61K 9/127 (2006.01) A61K 38/17 (2006.01)

A61K 9/51 (2006.01) A61P 35/00 (2006.01)

A61K 31/7105 (2006.01) A61P 35/02 (2006.01)

(21) International Application Number:

PCT/US2021/032438

(22) International Filing Date:

14 May 2021 (14.05.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/024,862 14 May 2020 (14.05.2020) US

63/183,119 03 May 2021 (03.05.2021) US

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

(72) Inventors: JAIN, Ruchi; 200 Technology Square, Cambridge, Massachusetts 02139 (US). BICKNELL, Alicia Anne; 200 Technology Square, Cambridge, Massachusetts 02139 (US).

(74) Agent: GHAI, Kanika et al.; Fish & Richardson, P.O. Box 1022, Minneapolis, Minnesota 55440-1022 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO,

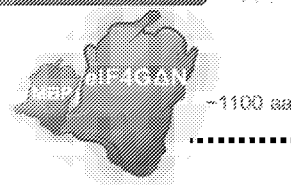
(54) Title: LNP COMPOSITIONS COMPRISING AN MRNA THERAPEUTIC AND AN EFFECTOR MOLECULE

Fig. 1

RNA encoding effector protein:

5' Cap 5' UTR 3' UTR Poly (A) tail

MBP: MS2 binding protein



RNA encoding target protein:

5' Cap 5' UTR Coding sequence 3' UTR Poly (A) tail

19nt MS2 loop
w 20nt spacers

223 nt

v1.1 333 nt

v1.1 333 nt

(57) **Abstract:** The disclosure features LNP compositions and systems comprising a therapeutic payload or prophylactic payload, a binding element, a tether molecule and/or an effector molecule and uses thereof. The LNP compositions or systems of the present disclosure comprise: (a) a first polynucleotide (e.g., mRNA) comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide (e.g., mRNA) comprising a sequence encoding: (1) an effector molecule, and/or (2) a polypeptide that recognizes the binding element (a tether molecule). Such compositions or systems can: increase the level and/or activity of the therapeutic payload or prophylactic payload, e.g., increase the level, stability and/or activity of the mRNA encoding the therapeutic payload or prophylactic payload. Also disclosed herein are methods of treating a disorder, or for modulating an immune response in a subject using the disclosed LNP compositions or systems.



NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW,
SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

LNP COMPOSITIONS COMPRISING AN MRNA THERAPEUTIC AND AN EFFECTOR MOLECULE

5 CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of priority of U.S. Provisional Appl. No. 63/024,862, filed May 14, 2020, and U.S. Provisional Appl. No. 63/183,119, filed May 3, 2021, the contents of which are incorporated herein by reference in their entirety.

SEQUENCE LISTING

10 The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on May 14, 2021, is named 45817-0103WO1_SL.txt and is 475,291 bytes in size.

BACKGROUND OF THE DISCLOSURE

15 Efforts to increase mRNA potency have focused on generating canonical linear mRNAs with optimal sequence design for the untranslated regions (UTRs) and open reading frame (ORFs). Recent advances have added end-protection where modified caps and tails render mRNAs more resistant to the cellular degradation machinery. However, there is no indication that the maximum potency possible has been achieved
20 with these efforts, with respect to peak or duration of expression. This is particularly true in the case of mRNA therapeutics. Current approaches are focused on modifying the mRNAs themselves. Therefore, there is a need to further improve peak and/or duration of mRNA expression by exploiting RNA biology.

25 SUMMARY OF THE DISCLOSURE

The present disclosure provides, *inter alia*, lipid nanoparticle (LNP) compositions or systems comprising a therapeutic payload or prophylactic payload, a binding element, a tether molecule and/or an effector molecule and uses thereof. The LNP compositions or systems of the present disclosure comprise: (a) a first

polynucleotide (e.g., mRNA) comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide (e.g., mRNA) comprising a sequence encoding: (1) an effector molecule, and/or (2) a polypeptide that binds to, e.g., recognizes the binding element (a tether molecule). In some embodiments, the effector molecule recognizes and binds to the binding element. In an embodiment, the first polynucleotide and the second polynucleotide are disposed in the same or different polynucleotides. In an embodiment, a system disclosed herein is formulated as an LNP. In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are formulated in the same LNP. In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are formulated in different LNPs. In an aspect, the LNP compositions or systems of the present disclosure can: increase the level, duration of expression, and/or activity of the therapeutic payload or prophylactic payload, e.g., increase the level, duration of expression, and/or activity of the mRNA encoding the therapeutic payload or prophylactic payload, or increase the level, duration of expression and/or activity of the therapeutic payload or prophylactic payload. Also disclosed herein are methods of using an LNP composition or system comprising a therapeutic payload or prophylactic payload, a binding element, a tether molecule and/or an effector molecule, for treating a disease or disorder, or for promoting a desired biological effect in a subject, e.g., for modulating an immune response in a subject. Additional aspects of the disclosure are described in further detail below.

In an aspect, disclosed herein is a lipid nanoparticle (LNP) composition comprising: (a) a first polynucleotide comprising: (1) a sequence encoding a polypeptide, and (2) a binding element; and (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and (2) a polypeptide that binds to, e.g., recognizes, the binding element (a tether molecule).

In an embodiment, the polypeptide of (a) does not comprise a reporter protein. In another embodiment, the polypeptide of (a) encodes a peptide or polypeptide having a desirable biologic effect, e.g., a therapeutic protein.

In an aspect, disclosed herein is a lipid nanoparticle (LNP) composition comprising:

(a) a first polynucleotide comprising: (1) a sequence encoding a polypeptide, and (2) a binding element; and (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule. In an embodiment, the effector molecule further comprises a polypeptide that binds to, e.g., recognizes, the binding element (a tether molecule).

In an embodiment, the polypeptide of (a) does not comprise a reporter protein. In another embodiment, the polypeptide of (a) encodes a peptide or polypeptide having a desirable biologic effect, e.g., a therapeutic protein.

In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a first domain which modulates a parameter of, e.g., level and/or activity of: an RNA (e.g., an mRNA); or a protein encoded by the RNA. In an embodiment, the parameter comprises one, two, three or all of: (1) mRNA level and/or activity and/or subcellular localization (e.g., half-life and/or expression); (2) protein level and/or activity (e.g., half-life and/or expression); (3) protein translation rate or (4) protein localization, e.g., location. In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a second domain which binds to, e.g., recognizes, the binding element (a tether molecule).

In an embodiment the effector molecule comprising the tether molecule comprises a polypeptide comprising the first domain and the second domain. In an embodiment, the first and second domains are operatively linked.

In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the effector molecule is upstream of the nucleotide sequence encoding the tether molecule. In an embodiment, the nucleotide sequence encoding the effector molecule is downstream of the nucleotide sequence encoding the tether molecule. In an embodiment, the nucleotide sequence encoding the effector molecule is separated from the nucleotide sequence encoding the tether molecule by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the effector molecule is adjacent to the nucleotide sequence encoding the tether molecule.

5 In another aspect, disclosed herein is a lipid nanoparticle (LNP) composition comprising: (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and (2) a polypeptide that binds to, e.g., recognizes, the binding element (a tether molecule). In
10 an embodiment, (a) and (b) each comprise an mRNA.

In an embodiment, the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

In an embodiment, the first polynucleotide and the second polynucleotide, in no particular order, are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or
15 TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

In an embodiment, the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

In an embodiment, (a) and (b) are formulated as LNPs, e.g., formulated as the same LNP. In an embodiment, (a) and (b) are formulated as different LNPs.
20

In another aspect, disclosed herein is a lipid nanoparticle (LNP) composition comprising: (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule. In an
25 embodiment, the effector molecule further comprises a polypeptide that binds to, e.g., recognizes, the binding element (a tether molecule).

In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a first domain which modulates a parameter of, e.g., level and/or activity of: an RNA (e.g., an mRNA); or a protein encoded by the RNA. In an
30 embodiment, the parameter comprises one, two, three or all of: (1) mRNA level and/or activity and/or subcellular localization (e.g., half-life and/or expression); (2) protein

level and/or activity (e.g., half-life and/or expression); (3) protein translation rate or (4) protein localization, e.g., location. In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a second domain which binds to, e.g., recognizes, the binding element (a tether molecule).

5 In an embodiment the effector molecule comprising the tether molecule comprises a polypeptide comprising the first domain and the second domain. In an embodiment, the first and second domains are operatively linked.

 In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the
10 effector molecule is upstream of the nucleotide sequence encoding the tether molecule. In an embodiment, the nucleotide sequence encoding the effector molecule is downstream of the nucleotide sequence encoding the tether molecule. In an embodiment, the nucleotide sequence encoding the effector molecule is separated from the nucleotide sequence encoding the tether molecule by a protease cleavage site (e.g., a
15 P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

 In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the effector molecule is adjacent to the nucleotide sequence encoding the tether molecule.

 In an embodiment, (a) and (b) each comprise an mRNA.

20 In an embodiment, the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

 In an embodiment, the first polynucleotide and the second polynucleotide are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

25 In an embodiment, the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

 In an embodiment, (a) and (b) are formulated as LNPs, e.g., formulated as the same LNP. In an embodiment, (a) and (b) are formulated as different LNPs.

30 In an aspect, the disclosure provides a system comprising: (a) a first polynucleotide comprising: (1) a sequence encoding a polypeptide, and (2) a binding

element; and (b) a second polynucleotide comprising a sequence encoding: (1) an effector, and a (2) polypeptide that binds to, e.g., recognizes, the binding element (a tether molecule).

In an embodiment, the polypeptide of (a) does not comprise a reporter protein.

5

In yet another aspect, disclosed herein is a system comprising: (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and/or (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and a (2) polypeptide that recognizes the binding element (a tether molecule). In an embodiment, (a) and (b) each comprise an mRNA.

In an embodiment, the system comprises (a).

In an embodiment, the system comprises (b).

In an embodiment, the system comprises (a) and (b).

15 In an embodiment, the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

In an embodiment, the first polynucleotide and the second polynucleotide are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

20 In an embodiment, the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

In an embodiment, the therapeutic payload or prophylactic payload is not a reporter protein.

25 In another aspect, disclosed herein is a system comprising: (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and/or (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule In an embodiment, the effector molecule further comprises a polypeptide that binds to, e.g., recognizes, the binding element (a tether molecule).

30

In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a first domain which modulates a parameter of, e.g., level and/or activity of: an RNA (e.g., an mRNA); or a protein encoded by the RNA. In an embodiment, the parameter comprises one, two, three or all of: (1) mRNA level and/or activity and/or subcellular localization (e.g., half-life and/or expression); (2) protein level and/or activity (e.g., half-life and/or expression); (3) protein translation rate or (4) protein localization, e.g., location. In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a second domain which binds to, e.g., recognizes, the binding element (a tether molecule).

10 In an embodiment the effector molecule comprising the tether molecule comprises a polypeptide comprising the first domain and the second domain. In an embodiment, the first and second domains are operatively linked.

In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the effector molecule is upstream of the nucleotide sequence encoding the tether molecule. In an embodiment, the nucleotide sequence encoding the effector molecule is downstream of the nucleotide sequence encoding the tether molecule. In an embodiment, the nucleotide sequence encoding the effector molecule is separated from the nucleotide sequence encoding the tether molecule by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

20 In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the effector molecule is adjacent to the nucleotide sequence encoding the tether molecule.

In an embodiment, (a) and (b) each comprise an mRNA.

25 In an embodiment, the system comprises (a).

In an embodiment, the system comprises (b).

In an embodiment, the system comprises (a) and (b).

In an embodiment, the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

In an embodiment, the first polynucleotide and the second polynucleotide are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

5 In an embodiment, the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

In an embodiment, the therapeutic payload or prophylactic payload is not a reporter protein.

10 In an embodiment of any of the systems disclosed herein, the system comprises less than 5%, 10%, 15%, 20%, 25%, or 50% of a cellular impurity, e.g., a cellular component, e.g., a membrane, protein or lipid derived from a cellular extract.

In an embodiment of any of the systems disclosed herein, at least one of (a) or (b) is formulated as an LNP.

In an embodiment of any of the systems disclosed herein, (a) is formulated as an LNP.

15 In an embodiment of any of the systems disclosed herein, (b) is formulated as an LNP.

In an embodiment of any of the systems disclosed herein, (a) and (b) both are formulated as LNPs, e.g., the same LNP or different LNPs.

20 In an embodiment of any of the systems disclosed herein, (a) formulated as an LNP is in a first composition. In an embodiment, (b) formulated as an LNP is in a second composition. In an embodiment, (a) formulated as an LNP and (b) formulated as an LNP are in separate compositions. In an embodiment, (a) formulated as an LNP and (b) formulated as an LNP are in the same composition.

25 In an aspect, provided herein is a system comprising: (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element comprising an MS2 sequence, e.g., 6 MS2 sequences of 19 nucleotides separated by spacers of 20 nucleotides in length; (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule comprising
30 eIF4G, e.g., wildtype eIF4G, a variant or a fragment thereof; and (2) a tether molecule comprising MBP, e.g., wildtype MBP, a variant or fragment thereof.

In another aspect, provided herein is a pharmaceutical composition comprising a system, or LNP composition of disclosed herein.

5 In an aspect, the disclosure provides, a cell comprising a system, or LNP composition disclosed herein. In an embodiment, the cell has been contacted with the system, or LNP composition. In an embodiment, the cell is contacted with the system, or LNP composition in vivo. In an embodiment, the cell is contacted with the system, or LNP composition in vitro. In an embodiment, the cell is contacted with the system, or
10 LNP composition ex vivo. In an embodiment, the cell is maintained under conditions sufficient to allow for expression of one or both polynucleotides of the system, or LNP composition.

In an aspect, provided herein is a method of increasing expression of a
15 therapeutic payload or prophylactic payload in a cell, comprising administering to the cell a system, or LNP composition disclosed herein. In one embodiment, the cell is present in a subject.

In a related aspect, provided herein is a system or LNP composition for use in a method of increasing expression of a therapeutic payload or prophylactic payload in a
20 cell.

In another aspect, the disclosure provides a method of increasing expression of a therapeutic payload or prophylactic payload, in a subject, comprising administering to the subject an effective amount of a system or LNP composition disclosed herein.

25 In a related aspect, provided herein is a system or LNP composition for use in a method of increasing expression of a therapeutic payload or prophylactic payload in a subject.

In yet another aspect, provided herein is a method of delivering a system, or
30 LNP composition disclosed herein.

In a related aspect, provided herein is a system or LNP composition for use in a method of delivering the system or LNP composition to a cell.

In an embodiment, the method or use, comprises contacting the cell in vitro, in vivo or ex vivo with the system or LNP composition.

5

In an aspect, the disclosure provides a method of delivering a system or LNP composition disclosed herein to a subject having a disease or disorder, e.g., as described herein.

In a related aspect, provided herein is a system or LNP composition for use in a method of delivering the system or LNP composition to a subject having a disease or disorder, e.g., as described herein.

10

In another aspect, provided herein is a method of modulating an immune response in a subject, comprising administering to the subject in need thereof an effective amount of a system, or LNP composition disclosed herein.

15

In a related aspect, provided herein is a system or LNP composition for use in a method of modulating an immune response in a subject, comprising administering to the subject an effective amount of the system, or LNP composition.

In an aspect, provided herein is a method of treating, preventing, or preventing a symptom of, a disease or disorder comprising administering to a subject in need thereof an effective amount of a system, or LNP composition disclosed herein.

20

In a related aspect, provided herein is a system or LNP composition for use in a method of treating, preventing, or preventing a symptom of, a disease or disorder in a subject, comprising administering to the subject in need thereof an effective amount of the system, or LNP composition.

25

In an embodiment, the LNP composition comprises: (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and (2) a polypeptide that recognizes the binding element (a tether molecule). In an embodiment, (a) and (b) each comprise an mRNA.

30

In an embodiment, the system comprises: (a) a first polynucleotide comprising:
(1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a
binding element; and (b) a second polynucleotide comprising a sequence encoding: (1)
an effector molecule, and (2) a polypeptide that recognizes the binding element (a tether
5 molecule). In an embodiment, (a) and (b) each comprise an mRNA.

In an embodiment, the first polynucleotide and/or the second polynucleotide of
the system is formulated as an LNP. In an embodiment, the first polynucleotide of the
system is formulated as an LNP. In an embodiment, the second polynucleotide of the
system is formulated as an LNP. In an embodiment, both the first and the second
10 polynucleotides of the system are formulated as LNPs.

In an embodiment, the LNP comprising the first polynucleotide is the same as
the LNP comprising the second polynucleotide. In an embodiment, the LNP comprising
the first polynucleotide is different from the LNP comprising the second polynucleotide.

In an embodiment, the LNP comprising the first polynucleotide is in a first
15 composition. In an embodiment, the LNP comprising the second polynucleotide is in a
second composition. In an embodiment, the LNP comprising the first polynucleotide
and the LNP comprising the second polynucleotide are in the same composition. In an
embodiment, the LNP comprising the first polynucleotide and the LNP comprising the
second polynucleotide are in different compositions.

20

In an embodiment, the LNP comprising (a) and the LNP comprising (b) are
administered simultaneously, e.g., substantially simultaneously.

In an embodiment, the LNP comprising (a) and the LNP comprising (b) are
administered sequentially.

25 In an embodiment, the LNP comprising (a) is administered first.

In an embodiment, the LNP comprising (a) is administered first followed by
administration of the LNP comprising (b).

In an embodiment, the LNP comprising (b) is administered first.

In an embodiment, the LNP comprising (b) is administered first followed by
30 administration of the LNP comprising (a).

In some embodiments of any of the methods disclosed herein, the LNP composition comprises: (i) an ionizable lipid, e.g., an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-lipid. In some embodiments, the ionizable lipid comprises a compound of Formula (IIa). In some
5 embodiments, the ionizable lipid comprises a compound of Formula (IIe).

Additional features of any of the LNP compositions, pharmaceutical composition comprising said LNPs, methods or compositions for use disclosed herein include the following embodiments.

10

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the tether molecule of the second polynucleotide comprises an RNA binding protein or a fragment thereof, which binds to, e.g., recognizes, the binding element of the first polynucleotide. In some embodiments, the binding element of the
15 first polynucleotide is situated upstream (5') or downstream (3') or in the open reading frame of the sequence encoding the therapeutic payload or prophylactic payload. In some embodiments, the binding element of the first polynucleotide is situated upstream (5') or downstream (3') of a 5' UTR of the first polynucleotide. In some embodiments, the binding element of the first polynucleotide is situated upstream (5') or downstream
20 (3') of a 3' UTR of the first polynucleotide. In some embodiments, the binding element of the first polynucleotide is situated downstream of a 3' UTR of the first polynucleotide.

In some embodiments, the binding element of the first polynucleotide is bound
25 by the tether molecule of the second polynucleotide, e.g., a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, or 15.5kd or a variant or fragment thereof. In some embodiments, the binding element comprises a sequence which is bound, e.g., recognized, by the tether molecule. In some embodiments, the binding element comprises a sequence comprising a structure that is bound, e.g., recognized, by
30 the tether molecule.

In some embodiments, the binding element is chosen from a binding element provided in Table 1, e.g., MS2, PP7, BoxB, U1A hairpin or PRE or kink-turn, or a variant or fragment thereof. In some embodiments, the binding element is MS2. In some embodiments, the binding element is PP7. In some embodiments, the binding element is BoxB. In some embodiments, the binding element is U1A hairpin. In some
5 embodiments, the binding element is PRE. In some embodiments, the binding element is kink-turn.

In some embodiments, when the binding element is MS2 (e.g., wildtype MS2, or a variant or fragment thereof) the tether molecule is MBP (e.g., wildtype MBP, a variant
10 or fragment thereof).

In some embodiments, when the binding element is PP7 (e.g., wildtype PP7, or a variant or fragment thereof) the tether molecule is PCP (e.g., wildtype PCP, or a variant or fragment thereof).

In some embodiments, when the binding element is BoxB (e.g., wildtype BoxB,
15 or a variant or fragment thereof) the tether molecule is Lambda N (e.g., wildtype Lambda N, or a variant or fragment thereof).

In some embodiments, when the binding element is U1A hairpin (e.g., wildtype U1A hairpin, or a variant or fragment thereof) the tether molecule is U1A (e.g., wildtype U1A, or a variant or fragment thereof).

In some embodiments, when the binding element is PRE (e.g., wildtype PRE, or a variant or fragment thereof) the tether molecule is PUF (e.g., wildtype PUF, or a variant or fragment thereof).
20

In some embodiments, when the binding element is a kink-turn forming sequence and the tether molecule is 15.5kd (e.g., wildtype 15.5kd, or a variant or
25 fragment thereof).

In some embodiments, the binding element comprises a sequence comprising 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90 or 100 nucleotides. In some embodiments, the binding element comprises a sequence comprising about 5-100, about 5-90, about 5-80, about 5-70, about 5-60,
30 about 5-50, about 5-40, about 5-30, about 5-25, about 5-20, about 5-19, about 5-18, about 5-17, about 5-16, about 5-15, about 5-14, about 5-13, about 5-12, about 5-11,

about 5-10, about 5-9, about 5-8, about 5-7 or about 5-6 nucleotides. In some embodiments, the binding element comprises a sequence comprising about 5-100, about 6-100, about 7-100, about 8-100, about 9-100, about 10-100, about 11-100, about 12-100, about 13-100, about 14-100, about 15-100, about 16-100, about 17-100, about 18-100, about 19-100, about 20-100, about 21-100, about 22-100, about 23-100, about 24-100, about 25-100, about 30-100, about 40-100, about 50-100, about 60-100, about 70-100, about 80-100, or about 90-100 nucleotides. In some embodiments, the binding element comprises a sequence comprising about 5-100, about 6-90, about 7-80, about 8-70, about 9-60, about 10-50, about 11-40, about 12-30, about 13-25, about 14-24, about 15-23, about 16-22, about 17-21, or about 18-20 nucleotides. In some embodiments, the binding element comprises a sequence comprising 19 nucleotides, e.g., a binding element nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the binding element comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20 or 30 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises no more than 80, 70, 60, 50, 40 or 30 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises about 1-30, about 1-20, about 1-10, about 1-9, about 1-8, about 1-7, about 1-6, about 1-5, about 1-4, about 1-3, or about 1-2 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises about 1-30, about 2-30, about 3-30, about 4-30 about, 5-30 about, 6-30, about 7-30, about 8-30, about 9-30, about 10-30, about 11-30, about 12-30, about 13-30, about 14-30, about 15-30, or about 20-30 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises about 1-30, about 2-20, about 3-15, about 4-14, about 5-13, about 6-12, about 7-11, or about 8-10 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises 6 repeats of the sequence bound by the tether molecule of the second polynucleotide.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, each repeat is separated by a spacer sequence comprising 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90 or 100 nucleotides. In some embodiments, the spacer sequence comprises

5 about 1-100, about 1-90, about 1-80, about 1-70, about 1-60, about 1-50, about 1-40, about 1-30, about 1-25, about 1-20, about 1-19, about 1-18, about 1-17, about 1-16, about 1-15, about 1-14, about 1-13, about 1-12, about 1-11, about 1-10, about 1-9, about 1-8, about 1-7, about 1-6, about 1-5, about 1-4, about 1-3, or about 1-2 nucleotides. In some embodiments, the spacer sequence comprises about 1-100, about 2-100, about 3-

10 100, about 4-100, about 5-100, about 6-100, about 7-100, about 8-100, about 9-100, about 10-100, about 11-100, about 12-100, about 13-100, about 14-100, about 15-100, about 16-100, about 17-100, about 18-100, about 19-100, about 20-100, about 21-100, about 22-100, about 23-100, about 24-100, about 25-100, about 30-100, about 40-100, about 50-100, about 60-100, about 70-100, about 80-100, or about 90-100 nucleotides.

15 In some embodiments, the spacer sequence comprises about 1-100, about 2-90, about 3-80, about 4-70, about 5-60, about 6-50, about 7-40, about 8-40, about 9-30, about 10-25, about 11-24, about 12-23, about 13-22, about 14-21, about 15-20, about 16-19, about 17-18 nucleotides. In some embodiment, the spacer sequence comprises 20 nucleotides, e.g., a spacer sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%,

20 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the effector molecule is chosen from: a translation factor, a splicing factor, an RNA stabilizing factor, an RNA editing factor, an RNA-binding

25 factor, an RNA localizing factor, or a combination thereof. In some embodiments, the effector molecule is a translation factor, e.g., a translation factor provided in Table 4, e.g., eIF4G; Poly A binding protein (PABP); eIF3d or a component thereof; Dazl, or a fragment, or variant or combination thereof.

In some embodiments, the effector molecule is a translation factor which

30 modulates, e.g., facilitates, ribosome binding, e.g., recruitment, pre-initiation complex

formation, or RNA unwinding. In some embodiments, the effector molecule comprises eIF4G, e.g., wildtype eIF4G, a variant of eIF4G, or a fragment thereof.

In some embodiments, the effector molecule comprises wildtype eIF4G. In some embodiments, wildtype eIF4G comprises a sequence of about 1600 amino acids.

5 In some embodiments, the effector molecule comprises a fragment of eIF4G, e.g., as disclosed herein. In some embodiments, the eIF4G fragment retains ribosome binding, e.g., recruitment.

In some embodiments, the eIF4G fragment is about 1,500-200 amino acids, about 1,400-300 amino acids, about 1,300-350 amino acids, about 1,200-400 amino acids, about 1,100-450 amino acids, about 1,000-500 amino acids, about 900-550 amino acids, about 800-600 amino acids, about 1,500-300 amino acids, 1,500-400 amino acids, 1,500-500 amino acids, about 1,500-600 amino acids, amino acids, about 1,500-700 amino acids, about 1,500-800 amino acids, about 1,500-900 amino acids, about 1,500-1000 amino acids, about 1,500-1,100 amino acids, about 1,500-1,200 amino acids, about 1,500-1,300 amino acids, about 1,500-1,400 amino acids, about 1,400-200 amino acids, about 1,300-200 amino acids, about 1,200-200 amino acids, about 1,100-200 amino acids, about 1,000-200 amino acids, about 900-200 amino acids, about 800-200 amino acids, about 700-200 amino acids, about 600-200 amino acids, or about 500-200 amino acids in length.

20 In some embodiments, the eIF4G fragment is about 500 amino acids in length. In some embodiments, the eIF4G fragment is about 600 amino acids in length. In some embodiments, the eIF4G fragment is about 700 amino acids in length. In some embodiments, the eIF4G fragment is about 800 amino acids in length. In some embodiments, the eIF4G fragment is about 900 amino acids in length. In some
25 embodiments, the eIF4G fragment is about 1000 amino acids in length. In some embodiments, the eIF4G fragment is about 1100 amino acids in length. In some embodiments, the eIF4G fragment is about 1200 amino acids in length. In some embodiments, the eIF4G fragment is about 1300 amino acids in length. In some embodiments, the eIF4G fragment is about 1400 amino acids in length. In some
30 embodiments, the eIF4G fragment is about 1500 amino acids in length.

In some embodiments, the effector molecule comprises a variant of eIF4G, e.g., as disclosed herein. In some embodiments, the eIF4G variant retains ribosome binding, e.g., recruitment. In some embodiments, the eIF4G variant comprises a mutation (e.g., substitution) in the eIF4G polypeptide sequence at any one, two, all or a combination of the following positions: amino acid 768, amino acid 771, or amino acid 776. In some
5 embodiments, the eIF4G variant comprises a mutation, e.g., substitution, at position 768 of the eIF4G polypeptide sequence, e.g., a Leucine to Alanine substitution at position 768. In some embodiments, the eIF4G variant comprises a mutation, e.g., substitution, at position 771 of the eIF4G polypeptide sequence, e.g., a Leucine to Alanine
10 substitution at position 771. In some embodiments, the eIF4G variant comprises a mutation, e.g., substitution, at position 776 of the eIF4G polypeptide sequence, e.g., a Phenylalanine to Alanine at position 776. In some embodiments, the eIF4G variant comprises a mutation, e.g., substitution, at position 768 of the eIF4G polypeptide sequence, e.g., an Alanine at position 768; and a mutation, e.g., substitution, at position
15 771 of the eIF4G polypeptide sequence, e.g., an Alanine at position 771. In some embodiments, the eIF4G variant comprises a mutation, e.g., substitution, at position 768 of the eIF4G polypeptide sequence, e.g., an Alanine at position 768; and a mutation, e.g., substitution, at position 776 of the eIF4G polypeptide sequence, e.g., an Alanine at position 776. In some embodiments, the eIF4G variant comprises a mutation, e.g.,
20 substitution, at position 771 of the eIF4G polypeptide sequence, e.g., an Alanine at position 771; and a mutation, e.g., substitution, at position 776 of the eIF4G polypeptide sequence, e.g., an Alanine at position 776. In some embodiments, the eIF4G variant comprises a mutation, e.g., substitution, at position 771 of the eIF4G polypeptide sequence, e.g., an Alanine at position 771; a mutation, e.g., substitution, at position 771
25 of the eIF4G polypeptide sequence, e.g., an Alanine at position 771; and a mutation, e.g., substitution, at position 776 of the eIF4G polypeptide sequence, e.g., an Alanine at position 776.

In some embodiments, the effector molecule is a part of the eIF3 complex, e.g., which can recruit the ribosome. In some embodiments, the eIF3 complex comprises
30 eIF3d, eIF3c, eIF3e, or eIF3i, or a fragment thereof, or any combination thereof.

In some embodiments, the effector molecule comprises an amino acid sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof. In some embodiments, the effector molecule is encoded by a nucleotide sequence provided in Table 2, or a sequence with at least 80%, 85%,
5 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the effector molecule is a splicing factor, e.g., a splicing factor provided in Table 4, e.g., Rnps1, Magoh, Y14, or a fragment or variant, or combination
10 thereof.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the effector molecule is an RNA stabilizing factor, e.g., a splicing factor provided in Table 4, e.g., HuR or a fragment, or variant thereof.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the tether molecule binds to a binding element in the first polynucleotide. In some embodiments, the tether molecule binds to a sequence of the binding element or to a structure comprising the sequence of the binding element. In some embodiments, the tether molecule comprises an RNA binding protein or a
15 20 fragment thereof.

In some embodiments, the tether molecule comprises a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, or 15.5kd or a variant or fragment thereof.

In some embodiments, when the tether molecule is MBP (e.g., wildtype MBP, a variant or fragment thereof) the binding element is MS2 (e.g., wildtype MS2, or a variant or fragment thereof).
25

In some embodiments, when the tether molecule is PCP (e.g., wildtype PCP, or a variant or fragment thereof) the binding element is PP7 (e.g., wildtype PP7, or a
30 variant or fragment thereof).

In some embodiments, when the tether molecule is Lambda N (e.g., wildtype Lambda N, or a variant or fragment thereof) the binding element is BoxB (e.g., wildtype BoxB, or a variant or fragment thereof).

5 In some embodiments, when the tether molecule is U1A (e.g., wildtype U1A, or a variant or fragment thereof) the binding element is U1A hairpin (e.g., wildtype U1A hairpin, or a variant or fragment thereof).

In some embodiments, when the tether molecule is 15.5kd (e.g., wildtype 15.5kd, or a variant or fragment thereof) the binding element is a kink-turn forming sequence (e.g., wildtype kink-turn forming sequence, or a variant or fragment thereof).

10 In some embodiments, when the tether molecule is PUF (e.g., wildtype PUF, or a variant or fragment thereof) the binding element is PRE (e.g., wildtype PRE, or a variant or fragment thereof).

In some embodiments, the tether molecule comprises MBP. In some embodiments, the tether molecule comprises an amino acid sequence provided in Table 15 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof. In some embodiments, the tether molecule comprises the tether molecule is encoded by a nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

20 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the therapeutic payload or prophylactic payload comprises an mRNA encoding: a secreted protein, a membrane-bound protein; or an intercellular protein.

In some embodiments, the therapeutic payload or prophylactic payload is chosen 25 from a cytokine, an antibody, a vaccine (e.g., an antigen, an immunogenic epitope), a receptor, an enzyme, a hormone, a transcription factor, a ligand, a membrane transporter, a structural protein, a nuclease, or a component, variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload 30 comprises a cytokine, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises an antibody or a variant or fragment (e.g., a biologically active fragment) thereof.

5 In some embodiments, the therapeutic payload or prophylactic payload comprises a vaccine (e.g., an antigen, an immunogenic epitope), or a component, variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a protein or peptide.

10 In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in one, two, three, four, five, six or all, or any combination thereof, of the following in a cell (e.g., in a cell contacted with the system or LNP composition):

- 15 (i) increased expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (ii) sustained expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (iii) increased expression and/or level of therapeutic payload or prophylactic payload;
- 20 (iv) sustained expression and/or level of therapeutic payload or prophylactic payload;
- (v) increased stability of mRNA encoding the therapeutic payload or prophylactic payload;
- (vi) increased resistance of translation of therapeutic payload or prophylactic payload to cellular environment, e.g., stress or nutrient deprivation or translation factor availability;
- 25 (vii) reduced dosing of the therapeutic payload or prophylactic payload; or
- (viii) reduced toxicity, e.g., reduced modulation of a protein translated from endogenous mRNA in a cell.

In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in increased expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload.

5 In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in sustained expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload.

In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in increased expression and/or level of therapeutic payload or prophylactic payload.

10 In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in sustained expression and/or level of therapeutic payload or prophylactic payload.

In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in increased stability of mRNA encoding the therapeutic
15 payload or prophylactic payload. In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in increased resistance of translation of therapeutic payload or prophylactic payload to cellular environment, e.g., stress or nutrient deprivation or translation factor availability.

In some embodiments, any of the systems, LNP compositions, methods or uses
20 disclosed herein, results in reduced dosing of the therapeutic payload or prophylactic payload.

In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in reduced toxicity, e.g., reduced modulation of a protein translated from endogenous mRNA in a cell.

25 In some embodiments, any one, or all of (i)-(vii) is compared to a cell which:
(a) has not been contacted with the system disclosed herein;
(b) has not been contacted with the LNP composition disclosed herein;
(c) has not been contacted with an LNP comprising the first polynucleotide; or
(d) has been contacted with an LNP comprising the first polynucleotide but has
30 not been contacted with the second polynucleotide, e.g., an LNP comprising the second polynucleotide.

In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, result in increased expression, duration of expression, and/or level of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., as measured
5 by an assay in Example 2 or 3. In some embodiments, the increased expression, duration of expression, and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding element of the first polynucleotide. In some embodiments, the increase
10 in expression, duration of expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 30, 35, 40 or 50 fold.

In some embodiments, the increase in expression and/or level of the mRNA comprises an increase in stability (e.g., half-life) of the mRNA encoding the therapeutic
15 payload or prophylactic payload, e.g., about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 20 fold increase in stability of the mRNA encoding the therapeutic payload or prophylactic payload. In some embodiments, the mRNA encoding the therapeutic payload or prophylactic payload has a half-life of about 3-25 hours, about 4-20 hours, about 4-15 hours, about 5-10 hours, about 6-9 hours or about 7-8 hours.

20

In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in sustained expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., as measured by an assay in Example
25 4. In some embodiments, at least 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more of the expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is sustained for a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 24 or 36 hours. In some embodiments, the sustained expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic
30 payload or prophylactic payload, which mRNA lacks a binding element of the first polynucleotide. As used herein, the term “sustained expression” refers to a longer

duration of expression and/or longer maintainence of mRNA levels compared to mRNA expression in an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload, which mRNA lacks a binding element of the first polynucleotide.

5 In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in a decreased loss, e.g., about a 1.2-fold, 2-fold, 3-fold, 4-fold or 5-fold decrease in loss, of mRNA encoding the therapeutic payload or prophylactic payload. In some embodiments, the decrease in loss of mRNA encoding the therapeutic payload or prophylactic payload occurs over a period of time, e.g., about 1, 2, 3, 4, 5, 6,
10 7, 8, 9, 10, 12, or 24 hours.

 In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in a decreased loss, e.g., about a 1.2-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold decrease in loss, of translating mRNA. In some embodiments, the decrease in loss of mRNA encoding the therapeutic payload or
15 prophylactic payload occurs over a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, or 24 hours.

 In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in a sustained, e.g., maintained, level of translation of an mRNA encoding the therapeutic payload or prophylactic payload.

20 In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in increased expression, duration of expression, and/or level of the therapeutic payload or prophylactic payload, e.g., increased protein level, translation, or half-life, e.g., as measured by an assay of Example 4. In some
embodiments, the increased expression and/or level of the therapeutic payload or
25 prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding element of the first polynucleotide. In some embodiments, the increase in expression and/or level of the therapeutic payload or prophylactic payload is about
1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 30, 35, 40 or 50 fold.

30

In some embodiments, any of the systems, LNP compositions, methods or uses disclosed herein, results in sustained expression and/or level of the therapeutic payload or prophylactic payload.

5 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the first polynucleotide and the second polynucleotide each comprises an mRNA.

 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the tether molecule of (b)(2) binds to, e.g., recognizes, the
10 binding element of (a)(2).

 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the first polynucleotide is formulated as an LNP or the second polynucleotide is formulated as an LNP.

 In some embodiments of any of the systems, LNP compositions, methods or
15 uses disclosed herein, the first polynucleotide is formulated as an LNP and the second polynucleotide is formulated as an LNP, e.g., the same LNP or different LNPs.

 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide is the same LNP.

20 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are different LNPs.

 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the binding element comprises a sequence comprising 19
25 nucleotides, e.g., a binding element nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the spacer sequence comprises 20 nucleotides, e.g., a spacer
30 sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the effector molecule comprises an amino acid sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

5 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the effector molecule is encoded by a nucleotide sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

10 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the tether molecule comprises an amino acid sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

15 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the tether molecule is encoded by a nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

20 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the first polynucleotide comprises an mRNA comprising at least one chemical modification.

25 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the LNP is formulated for intravenous, subcutaneous, intramuscular, intranasal, intraocular, rectal, pulmonary or oral delivery. In some embodiments, the LNP is formulated for intravenous delivery. In some embodiments, the LNP is formulated for subcutaneous delivery. In some embodiments, the LNP is formulated for intramuscular delivery. In some embodiments, the LNP is formulated for intranasal delivery. In some embodiments, the LNP is formulated for intraocular delivery. In some embodiments, the LNP is formulated for rectal delivery. In some
30 embodiments, the LNP is formulated for pulmonary delivery. In some embodiments, the LNP is formulated for oral delivery.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the LNP further comprising a pharmaceutically acceptable carrier or excipient.

5

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the polynucleotide, e.g., the first and/or second polynucleotide comprises a cap, a 3' UTR, a 5' UTR, a Poly A tail and/or a micro RNA (miRNA) binding site. In some embodiments, the cap comprises a cap disclosed herein. In some
10 embodiments, the polynucleotide, e.g., the first and/or second polynucleotide does not comprise a cap. In some embodiments, the 3' UTR comprises a 3' UTR disclosed herein, e.g., a v1.1 3' UTR or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto. In some embodiments, the 5' UTR comprises a 5' UTR disclosed herein. In some embodiments, the Poly A tail comprises
15 a Poly A tail sequence disclosed herein or a fragment thereof. In some embodiments, the polynucleotide, e.g., the first and/or second polynucleotide does not comprise a Poly A tail. In some embodiments, the miRNA binding site comprises a miRNA binding site disclosed herein.

In some embodiments, the polynucleotide, e.g., the first and/or second
20 polynucleotide is a circular polynucleotide. In some embodiments, the first polynucleotide is a circular polynucleotide. In some embodiments, the second polynucleotide is a circular polynucleotide.

In some embodiments of any of the methods or uses disclosed herein, the LNP
25 comprising the first polynucleotide is administered at a lower dose compared to a reference LNP. In some embodiments, the LNP comprising the first polynucleotide is administered at a dose that is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% lower compared to the dose of a reference LNP. In some embodiments, the reference LNP is chosen from: an otherwise similar LNP comprising a polynucleotide
30 which does not have the binding element of the first polynucleotide; or an LNP that does not comprise the second polynucleotide.

In some embodiments, the LNP comprising the first polynucleotide is administered at a higher dose compared to the LNP comprising the second polynucleotide. In some embodiments, the LNP comprising the first polynucleotide is administered at a dose that is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% higher compared to the dose of the LNP comprising the second polynucleotide. In some embodiments, the LNP comprising the first polynucleotide is in molar excess compared to the LNP comprising the second polynucleotide.

In some embodiments, the LNP comprising the first polynucleotide is in about 1-800X molar excess, compared to the LNP comprising the second polynucleotide. In some embodiments, the LNP comprising the first polynucleotide is in about 1-750x, about 2-700x, about 3-650 x, about 4-600 x, about 5-550 x, about 6-500 x, about 7-450 x, about 8-400 x, about 10-350 x, about 15-300 x, about 20-250 x, about 25-200 x, about 30-150 x, about 35-100 x, about 40-90 x, about 45-80 x, about 50-75 x, about 60-70x molar excess compared to the LNP comprising the second polynucleotide. In some embodiments, the LNP comprising the first polynucleotide is in about 2x, about 3x, about 4x, about 5 x, about 6x, about 7x, about 8x, about 9 x, about 10x, about 11 x, about 12 x, about 13 x, about 14 x, about 15 x, about 20 x, about 25 x, about 30 x, about 35 x, about 40 x, about 50 x, about 60 x, about 70 x, about 80 x, about 90 x, about 100 x, about 150 x, about 200 x, about 250 x, about 300 x, about 350 x, about 400 x, about 450 x, about 500 x, about 600 x, about 650 x, about 700 x, about 750 x, or about 800x molar excess compared to the LNP comprising the second polynucleotide.

In some embodiments, the LNP comprising the first polynucleotide is in about 9x molar excess compared to the LNP comprising the second polynucleotide.

In some embodiments, the LNP comprising the first polynucleotide is in about 10x molar excess compared to the LNP comprising the second polynucleotide.

In some embodiments, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are at the same molar amounts. In an embodiment of any of the LNP compositions, systems, methods or compositions for use disclosed herein, the first and second polynucleotides are formulated at an (a):(b) mass ratio of 10:1, 8:1, 6:1, 4:1, 3:1, 2:1, 1.5:1, or 1:1. In an embodiment, the first and second polynucleotides are formulated at an (a):(b) mass ratio of 1:1, 1.1:1, 1:2, 1:3, 1:4, 1:6,

1:8, or 1:10. In an embodiment, the first and second polynucleotides are formulated at an (a):(b) mass ratio of 1:1.

In an embodiment of any of the LNP compositions, systems, methods or compositions for use disclosed herein, the first polynucleotide, the second
 5 polynucleotide, or both, comprises at least one chemical modification. In an embodiment, the chemical modification is selected from the group consisting of pseudouridine, N1-methylpseudouridine, 2-thiouridine, 4'-thiouridine, 5-methylcytosine, 2-thio-1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-pseudouridine, 2-thio-5-aza-uridine, 2-thio-dihydropseudouridine, 2-thio-
 10 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-methyluridine, 5-methoxyuridine, and 2'-O-methyl uridine. In an embodiment, the chemical modification is selected from the group consisting of pseudouridine, N1-methylpseudouridine, 5-methylcytosine, 5-
 15 methoxyuridine, and a combination thereof. In an embodiment, the chemical modification is N1-methylpseudouridine. In an embodiment, each mRNA in the lipid nanoparticle comprises fully modified N1-methylpseudouridine. In some embodiments, the first polynucleotide, the second polynucleotide, or both, do not comprise any chemical modification.

20 In an embodiment of any of the LNP compositions, systems, methods or compositions for use disclosed herein, the LNP composition comprises: (i) an ionizable lipid, e.g., an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and, optionally, (iv) a PEG-lipid.

In an embodiment of any of the LNP compositions, systems, methods or
 25 compositions for use disclosed herein, the LNP composition comprises an ionizable lipid comprising an amino lipid. In an embodiment, the ionizable lipid comprises a compound of any of Formulae (I), (IA), (IB), (II), (IIa), (IIb), (IIc), (IIc), (IIe), (IIe), (IIg), (III), (IIIa1), (IIIa2), (IIIa3), (IIIa4), (IIIa5), (IIIa6), (IIIa7), or (IIIa8). In an embodiment, the ionizable lipid comprises a compound of Formula (I). In an
 30 embodiment, the ionizable lipid comprises a compound of Formula (IIa). In an embodiment, the ionizable lipid comprises a compound of Formula (IIe).

In an embodiment of any of the LNP compositions, systems, methods or compositions for use disclosed herein, the LNP composition comprises a non-cationic helper lipid or phospholipid comprising a compound selected from the group consisting of 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-Distearoyl-sn-glycero-3-phosphoethanolamine (DSPE), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine, 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16.0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), sphingomyelin, and mixtures thereof. In an embodiment, the phospholipid is DSPC, e.g., a variant of DSPC, e.g., a compound of Formula (IV).

In an embodiment of any of the LNP compositions, systems, methods or compositions for use disclosed herein, the LNP composition comprises a structural lipid. In one embodiment, the structural lipid is a phytosterol or a combination of a phytosterol and cholesterol. In one embodiment, the phytosterol is selected from the group consisting of β -sitosterol, stigmasterol, β -sitostanol, campesterol, brassicasterol, and combinations thereof.

In one embodiment, the structural lipid can be selected from the group including but not limited to, cholesterol, fecosterol, sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol, tomatidine, tomatine, ursolic acid, alpha-tocopherol, hopanoids, phytosterols, steroids, and mixtures thereof. In some embodiments, the structural lipid is a sterol. As defined herein, "sterols" are a subgroup of steroids consisting of steroid alcohols. In certain embodiments, the structural lipid is a steroid. In certain embodiments, the structural lipid is cholesterol. In certain embodiments, the structural lipid is an analog of cholesterol. In certain embodiments, the structural lipid is alpha-tocopherol.

In one embodiment, the structural lipid is selected from selected from β -sitosterol and cholesterol. In an embodiment, the structural lipid is β -sitosterol. In an embodiment, the structural lipid is cholesterol.

In an embodiment of any of the LNP compositions, systems, methods or compositions for use disclosed herein, the LNP composition comprises a PEG lipid. In one embodiment, the PEG-lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof.

In an embodiment, the PEG lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof. In an embodiment, the PEG lipid
5 is selected from the group consisting of PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC and PEG-DSPE lipid. In an embodiment, the PEG-lipid is PEG-DMG.

In an embodiment, the PEG lipid is chosen from a compound of: Formula (V), Formula (VI-A), Formula (VI-B), Formula (VI-C) or Formula (VI-D). In an
10 embodiment, the PEG-lipid is a compound of Formula (VI-A). In an embodiment, the PEG-lipid is a compound of Formula (VI-B). In an embodiment, the PEG-lipid is a compound of Formula (VI-C). In an embodiment, the PEG-lipid is a compound of Formula (VI-D).

In an embodiment of any of the LNP compositions, systems, methods or
15 compositions for use disclosed herein, the LNP comprises about 20 mol % to about 60 mol % ionizable lipid, about 5 mol % to about 25 mol % non-cationic helper lipid or phospholipid, about 25 mol %
to about 55 mol % sterol or other structural lipid, and about 0.5 mol % to about 15 mol % PEG lipid. In one embodiment of the LNPs or methods of the disclosure, the LNP
20 comprises about 35 mol % to about 55 mol % ionizable lipid, about 5 mol % to about 25 mol % non-cationic helper lipid or phospholipid, about 30 mol % to about 40 mol % sterol or other structural lipid, and about 0 mol % to about 10 mol % PEG lipid. In one
embodiment of the LNPs or methods of the disclosure, the LNP comprises about 50 mol % ionizable lipid, about 10 mol % non-cationic helper lipid or phospholipid, about 38.5
25 mol % sterol or other structural lipid, and about 1.5 mol % PEG lipid. In one
embodiment of the LNPs or methods of the disclosure, the LNP comprises about 49.83 mol % ionizable lipid, about 9.83 mol % non-cationic helper lipid or phospholipid,
about 30.33 mol % sterol or other structural lipid, and about 2.0 mol % PEG lipid.

In one embodiment, the mol % sterol or other structural lipid is 18.5% phytosterol and the total mol % structural lipid is 38.5%. In one embodiment, the mol % sterol or other structural lipid is 28.5% phytosterol and the total mol % structural lipid is 38.5%.

5 In one embodiment of the LNPs, systems or methods of the disclosure, the LNP comprises about 50 mol % a compound of Formula (IIa) and about 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs or methods of the disclosure, the LNP comprises 50 mol % a compound of Formula (IIa) and about 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs, systems, or methods of the disclosure, the LNP comprises about 50 mol % a compound of Formula (IIa) and 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs or methods of the disclosure, the LNP comprises 50 mol % a compound of Formula (IIa) and 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs or methods of the disclosure, the LNP comprises about 15 49.83 mol % a compound of Formula (IIa), about 9.83 mol % non-cationic helper lipid or phospholipid, about 30.33 mol % sterol or other structural lipid, and about 2.0 mol % PEG lipid.

In one embodiment of the LNPs, systems, or methods of the disclosure, the LNP comprises about 50 mol % a compound of Formula (IIe) and about 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs or methods of the disclosure, the LNP comprises 50 mol % a compound of Formula (IIe) and about 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs or methods of the disclosure, the LNP comprises about 50 mol % a compound of Formula (IIe) and 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs or methods of the disclosure, the LNP 25 comprises 50 mol % a compound of Formula (IIe) and 10 mol % non-cationic helper lipid or phospholipid. In one embodiment of the LNPs or methods of the disclosure, the LNP comprises about 49.83 mol % a compound of Formula (IIe), about 9.83 mol % non-cationic helper lipid or phospholipid, about 30.33 mol % sterol or other structural lipid, and about 2.0 mol % PEG lipid. 30

In an embodiment of any of the LNP compositions, systems, methods or compositions for use disclosed herein, the LNP is formulated for intravenous, subcutaneous, intramuscular, intraocular, intranasal, rectal, pulmonary or oral delivery. In an embodiment, the LNP is formulated for intravenous delivery. In an embodiment, the LNP is formulated for subcutaneous delivery. In an embodiment, the LNP is formulated for intramuscular delivery. In an embodiment, the LNP is formulated for intraocular delivery. In an embodiment, the LNP is formulated for intranasal delivery. In an embodiment, the LNP is formulated for rectal delivery. In an embodiment, the LNP is formulated for pulmonary delivery. In an embodiment, the LNP is formulated for oral delivery.

In an embodiment of any of the methods or compositions for use disclosed herein, the subject is a mammal, *e.g.*, a human.

Additional features of any of the aforesaid LNP compositions or methods of using said LNP compositions, include one or more of the following enumerated embodiments. Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following enumerated embodiments.

Other embodiments of the Disclosure

E1. A lipid nanoparticle (LNP) composition comprising:

(a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and

(b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and/or (2) a polypeptide that recognizes the binding element (a tether molecule),

optionally wherein, (a) and (b) each comprise an mRNA.

E2. The LNP composition of embodiment E1, wherein the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

E3. The LNP composition of embodiment E1, wherein the first polynucleotide and the second polynucleotide are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

- 5 E4. The LNP composition of embodiment E1, wherein the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

E5. The LNP composition of embodiment E1, wherein (a) and (b) each are formulated as LNPs, e.g., the same LNP.

10

E6. The LNP composition of embodiment E1, wherein (a) and (b) are formulated as different LNPs, optionally wherein:

- (i) (a) formulated as an LNP is in a first composition and (b) formulated as an LNP is in a second composition, e.g., (a) and (b) are in different compositions; or
- 15 (ii) (a) formulated as an LNP and (b) formulated as an LNP are in the same composition.

E7. A system comprising:

- (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic
- 20 payload or prophylactic payload, and (2) a binding element; and/or
- (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and/or a (2) polypeptide that recognizes the binding element (a tether molecule),
- optionally wherein, (a) and (b) each comprise an mRNA.

25

E8. The system of embodiment E7, wherein the system comprises (a).

E9. The system of embodiment E7, wherein the system comprises (b).

- 30 E10. The system of embodiment E7, wherein the system comprises (a) and (b).

E11. The system of any one of embodiments E7-E10, wherein the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

5 E12. The system of embodiment E11, wherein the first polynucleotide and the second polynucleotide are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

E13. The system of any one of embodiments E7-E10, wherein the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

10

E14. The system of any one of embodiments E7-E13, wherein the therapeutic payload or prophylactic payload is not a reporter protein.

15 E15. The system of any one of embodiments E7-E13, wherein the system comprises less than 5%, 10%, 15%, 20%, 25%, or 50% of a cellular impurity, e.g., a cellular component, e.g., a membrane, protein or lipid from a cell.

E16. The system of any one of embodiments E7-E15, wherein at least one of (a) or (b) is formulated as a lipid nanoparticle (LNP).

20

E17. The system, or LNP composition of embodiment E16, wherein (a) is formulated as an LNP.

25 E18. The system, or LNP composition of embodiment E16, wherein (b) is formulated as an LNP.

E19. The system, or LNP composition of embodiment E16, wherein (a) and (b) both are formulated as LNPs, e.g., the same LNP or different LNPs, optionally wherein:

30 (i) (a) formulated as an LNP is in a first composition and (b) formulated as an LNP is in a second composition, e.g., (b) and (b) are in different compositions; or

(ii) (a) formulated as an LNP and (b) formulated as an LNP are in the same composition.

5 E20. The method or system, or LNP composition of any one of the preceding embodiments, wherein the tether molecule of the second polynucleotide comprises an RNA binding protein or a fragment thereof which binds to, e.g., recognizes, the binding element of the first polynucleotide.

10 E21. The method, or system, or LNP composition of any one of the preceding embodiments, wherein the binding element of the first polynucleotide is situated upstream (5') or downstream (3') or in the open reading frame of the sequence encoding the therapeutic payload or prophylactic payload.

15 E22. The method, or system, or LNP composition of any one of the preceding embodiments, wherein the binding element of the first polynucleotide is situated upstream (5') or downstream (3') of a 5' UTR of the first polynucleotide.

20 E23. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element of the first polynucleotide is situated upstream (5') or downstream (3') of a 3' UTR of the first polynucleotide.

25 E24. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element of the first polynucleotide is situated downstream of a 3' UTR of the first polynucleotide.

E25. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element of the first polynucleotide is bound by the tether molecule of the second polynucleotide, e.g., a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, or 15.5kd or a variant or fragment thereof.

30

E26. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises a sequence which is bound, e.g., recognized, by the tether molecule.

- 5 E27. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises a sequence comprising a structure that is bound, e.g., recognized, by the tether molecule.

- E28. The system, or LNP composition of any one of the preceding embodiments,
10 wherein the binding element is chosen from a binding element provided in Table 1, e.g., MS2, PP7, BoxB, U1A hairpin or PRE or kink-turn, or a variant or fragment thereof.

- E29. The system, or LNP composition of any one of the preceding embodiments,
wherein when the binding element is MS2 (e.g., wildtype MS2, or a variant or fragment
15 thereof) the tether molecule is MBP (e.g., wildtype MBP, a variant or fragment thereof).

- E30. The system, or LNP composition of any one of the preceding embodiments,
wherein when the binding element is PP7 (e.g., wildtype PP7, or a variant or fragment
thereof) the tether molecule is PCP (e.g., wildtype PCP, or a variant or fragment
20 thereof).

- E31 The system, or LNP composition of any one of the preceding embodiments,
wherein when the binding element is BoxB (e.g., wildtype BoxB, or a variant or
fragment thereof) the tether molecule is Lambda N (e.g., wildtype Lambda N, or a
25 variant or fragment thereof).

- E32. The system, or LNP composition of any one of the preceding embodiments,
wherein when the binding element is U1A hairpin (e.g., wildtype U1A hairpin, or a
variant or fragment thereof) the tether molecule is U1A (e.g., wildtype U1A, or a
30 variant or fragment thereof).

E33. The system, or LNP composition of any one of the preceding embodiments, wherein when the binding element is PRE (e.g., wildtype PRE, or a variant or fragment thereof) the tether molecule is PUF (e.g., wildtype PUF, or a variant or fragment thereof).

5

E34. The system, or LNP composition of any one of the preceding embodiments, wherein when the binding element is a kink-turn forming sequence and the tether molecule is 15.5kd (e.g., wildtype 15.5kd, or a variant or fragment thereof).

10 E35. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises a sequence comprising 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90 or 100 nucleotides.

15 E36. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises a sequence comprising about 5-100, about 5-90, about 5-80, about 5-70, about 5-60, about 5-50, about 5-40, about 5-30, about 5-25, about 5-20, about 5-19, about 5-18, about 5-17, about 5-16, about 5-15, about 5-14, about 5-13, about 5-12, about 5-11, about 5-10, about 5-9, about 5-8, about 5-7 or about
20 5-6 nucleotides.

E37. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises a sequence comprising about 5-100, about 6-100, about 7-100, about 8-100, about 9-100, about 10-100, about 11-100, about 12-100,
25 about 13-100, about 14-100, about 15-100, about 16-100, about 17-100, about 18-100, about 19-100, about 20-100, about 21-100, about 22-100, about 23-100, about 24-100, about 25-100, about 30-100, about 40-100, about 50-100, about 60-100, about 70-100, about 80-100, or about 90-100 nucleotides.

E38. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises a sequence comprising about 5-100, about 6-90, about 7-80, about 8-70, about 9-60, about 10-50, about 11-40, about 12-30, about 13-25, about 14-24, about 15-23, about 16-22, about 17-21, or about 18-20 nucleotides.

5

E39. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises a sequence comprising 19 nucleotides, e.g., a binding element nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

10

E40. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20 or 30 repeats of the sequence bound by the tether molecule of the second polynucleotide.

15

E41. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises no more than 80, 70, 60, 50, 40 or 30 repeats of the sequence bound by the tether molecule of the second polynucleotide.

20

E42. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises about 1-30, about 1-20, about 1-10, about 1-9, about 1-8, about 1-7, about 1-6, about 1-5, about 1-4, about 1-3, or about 1-2 repeats of the sequence bound by the tether molecule of the second polynucleotide.

25

E43. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises about 1-30, about 2-30, about 3-30, about 4-30, about 5-30, about 6-30, about 7-30, about 8-30, about 9-30, about 10-30, about 11-30, about 12-30, about 13-30, about 14-30, about 15-30, or about 20-30 repeats of the sequence bound by the tether molecule of the second polynucleotide.

30

E44. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises about 1-30, about 2-20, about 3-15, about 4-14, about 5-13, about 6-12, about 7-11, or about 8-10 repeats of the sequence bound by the tether molecule of the second polynucleotide.

5

E45. The system, or LNP composition of any one of the preceding embodiments, wherein the binding element comprises 6 repeats of the sequence bound by the tether molecule of the second polynucleotide.

10 E46. The system, or LNP composition of any one of embodiments 40-45, wherein each repeat is separated by a spacer sequence comprising 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90 or 100 nucleotides.

15 E47. The system, or LNP composition of embodiment 46, wherein the spacer sequence comprises about 1-100, about 1-90, about 1-80, about 1-70, about 1-60, about 1-50, about 1-40, about 1-30, about 1-25, about 1-20, about 1-19, about 1-18, about 1-17, about 1-16, about 1-15, about 1-14, about 1-13, about 1-12, about 1-11, about 1-10, about 1-9, about 1-8, about 1-7, about 1-6, about 1-5, about 1-4, about 1-3, or about 1-2
20 nucleotides.

E48. The system, or LNP composition of embodiment 46 or 47, wherein the spacer sequence comprises about 1-100, about 2-100, about 3-100, about 4-100, about 5-100, about 6-100, about 7-100, about 8-100, about 9-100, about 10-100, about 11-100, about
25 12-100, about 13-100, about 14-100, about 15-100, about 16-100, about 17-100, about 18-100, about 19-100, about 20-100, about 21-100, about 22-100, about 23-100, about 24-100, about 25-100, about 30-100, about 40-100, about 50-100, about 60-100, about 70-100, about 80-100, or about 90-100 nucleotides.

E49. The system, or LNP composition of any one of embodiments 46-48, wherein the spacer sequence comprises about 1-100, about 2-90, about 3-80, about 4-70, about 5-60, about 6-50, about 7-40, about 8-40, about 9-30, about 10-25, about 11-24, about 12-23, about 13-22, about 14-21, about 15-20, about 16-19, about 17-18 nucleotides.

5

E50. The system, or LNP composition of any one of embodiments 46-49, wherein the spacer sequence comprises 20 nucleotides, e.g., a spacer sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

10

E51. The system, or LNP composition of any one of the preceding embodiments, wherein the effector molecule is chosen from: a translation factor, a splicing factor, an RNA stabilizing factor, an RNA editing factor, an RNA-binding factor, an RNA localizing factor, or a combination thereof.

15

E52. The system, or LNP composition of any one of the preceding embodiments, wherein the effector molecule is a translation factor, e.g., a translation factor provided in Table 4, e.g., eIF4G; Poly A binding protein (PABP); eIF3d or a component thereof; Dazl, or a fragment, or variant or combination thereof.

20

E53. The system, or LNP composition of any one of the preceding embodiments, wherein the effector molecule is a translation factor which modulates, e.g., facilitates, ribosome binding, e.g., recruitment, pre-initiation complex formation, or RNA unwinding.

25

E54. The system, or LNP composition of any one of the preceding embodiments, wherein the effector molecule comprises eIF4G, e.g., wildtype eIF4G, a variant of eIF4G, or a fragment thereof.

30

E55. The system, or LNP composition of any one of the preceding embodiments, wherein the effector molecule comprises a fragment of eIF4G, e.g., as disclosed herein.

E56. The system, or LNP composition of any one of the preceding embodiments, wherein the eIF4G fragment retains ribosome binding, e.g., recruitment.

- 5 E57. The system, or LNP composition of any one of embodiments 54-56, wherein the eIF4G fragment is about 1,200-200 amino acids, about 1,100-300 amino acids, 1,000-400 amino acids, 900-450 amino acids, 800-500 amino acids, 700-550 amino acids, 600-650 amino acids, 1,200-300 amino acids, 1,200-400 amino acids, 1,200-500 amino acids, 1,200-600 amino acids, 1,100-200 amino acids, 1,000-200 amino acids, 900-200
10 amino acids, 800-200 amino acids, 700-200 amino acids, 600-200 amino acids, or 500-200 amino acids, amino acids in length.

E58. The system, or LNP composition of any one of embodiments 54-57, wherein the eIF4G fragment is about 500 amino acids in length.

15

E59. The system, or LNP composition of any one of embodiments 54-57, wherein the eIF4G fragment is about 1,100 amino acids in length.

- E60. The system, or LNP composition of any one of the preceding embodiments,
20 wherein the effector molecule comprises a variant of eIF4G, e.g., as disclosed herein.

E61. The system, or LNP composition of any one of embodiments 54-60, wherein the eIF4G variant retains ribosome binding, e.g., recruitment.

- 25 E62. The system, or LNP composition of any one of embodiments 54-61, wherein the eIF4G variant comprises a mutation (*e.g.*, substitution) in the eIF4G polypeptide sequence at any one, two, all or a combination of the following positions: amino acid 768, amino acid 771, or amino acid 776.

E63. The system, or LNP composition of embodiment 62, wherein the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 768 of the eIF4G polypeptide sequence, *e.g.*, a Leucine to Alanine substitution at position 768.

- 5 E64. The system, or LNP composition of embodiment 62, wherein the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, a Leucine to Alanine substitution at position 771.

- 10 E65. The system, or LNP composition of embodiment 62, wherein the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, a Phenylalanine to Alanine at position 776.

- 15 E66. The system, or LNP composition of embodiment 62, wherein the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 768 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 768; and a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771.

- 20 E67. The system, or LNP composition of embodiment 62, wherein the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 768 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 768; and a mutation, *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 776.

- 25 E68. The system, or LNP composition of embodiment 62, wherein the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771; and a mutation, *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 776.

E69. The system, or LNP composition of embodiment 62, wherein the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771; a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771; and a mutation,
5 *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 776.

E70. The system, or LNP composition of any one of embodiments 54-69, wherein the effector molecule is a part of the eIF3 complex, *e.g.*, which can recruit the ribosome.
10

E71. The system, or LNP composition of embodiment 70, wherein the eIF3 complex comprises eIF3d, eIF3c, eIF3e, or eIF3i, or a fragment thereof, or any combination thereof.

15 E72. The system, or LNP composition of any one of embodiments 54-71, wherein the effector molecule comprises an amino acid sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

E73. The system, or LNP composition of any one of embodiments 54-71, wherein the
20 effector molecule is encoded by a nucleotide sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

E74. The system, or LNP composition of any one of embodiments 1-51, wherein the
25 effector molecule is a splicing factor, *e.g.*, a splicing factor provided in Table 4, *e.g.*, Rnps1, Magoh, Y14, or a fragment or variant, or combination thereof.

E75. The system, or LNP composition of any one of embodiments 1-51, wherein the effector molecule is an RNA stabilizing factor, *e.g.*, a splicing factor provided in Table
30 4, *e.g.*, HuR or a fragment, or variant thereof.

- E76. The system, or LNP composition of any one of the preceding embodiments, wherein the tether molecule binds to a binding element in the first polynucleotide.
- E77. The system, or LNP composition of any one of the preceding embodiments,
5 wherein the tether molecule binds to a sequence of the binding element or to a structure comprising the sequence of the binding element.
- E78. The system, or LNP composition of any one of the preceding embodiments,
10 wherein the tether molecule comprises an RNA binding protein or a fragment thereof.
- E79. The system, or LNP composition of any one of the preceding embodiments,
wherein the tether molecule comprises a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, or 15.5kd or a variant or fragment thereof.
- 15 E80. The system, or LNP composition of any one of the preceding embodiments, wherein when the tether molecule is MBP (e.g., wildtype MBP, a variant or fragment thereof) the binding element is MS2 (e.g., wildtype MS2, or a variant or fragment thereof).
- 20 E81. The system, or LNP composition of any one of the preceding embodiments, wherein when the tether molecule is PCP (e.g., wildtype PCP, or a variant or fragment thereof) the binding element is PP7 (e.g., wildtype PP7, or a variant or fragment thereof).
- 25 E82. The system, or LNP composition of any one of the preceding embodiments, wherein when the tether molecule is Lambda N (e.g., wildtype Lambda N, or a variant or fragment thereof) the binding element is BoxB (e.g., wildtype BoxB, or a variant or fragment thereof).

E83. The system, or LNP composition of any one of the preceding embodiments, wherein when the tether molecule is U1A (e.g., wildtype U1A, or a variant or fragment thereof) the binding element is U1A hairpin (e.g., wildtype U1A hairpin, or a variant or fragment thereof).

5

E84. The system, or LNP composition of any one of the preceding embodiments, wherein when the tether molecule is 15.5kd (e.g., wildtype 15.5kd, or a variant or fragment thereof) the binding element is a kink-turn forming sequence (e.g., wildtype kink-turn forming sequence, or a variant or fragment thereof).

10

E85. The system, or LNP composition of any one of the preceding embodiments, wherein when the tether molecule is PUF (e.g., wildtype PUF, or a variant or fragment thereof) the binding element is PRE (e.g., wildtype PRE, or a variant or fragment thereof).

15

E86. The system, or LNP composition of any one of the preceding embodiments, wherein the tether molecule comprises an amino acid sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

20

E87. The system, or LNP composition of any one of the preceding embodiments, wherein the tether molecule is encoded by a nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

25

E88. The system, or LNP composition of any one of the preceding embodiments, wherein the therapeutic payload or prophylactic payload comprises an mRNA encoding: a secreted protein, a membrane-bound protein; or an intercellular protein.

E89. The system, or LNP composition of embodiment 88, wherein the therapeutic payload or prophylactic payload is chosen from a cytokine, an antibody, a vaccine (e.g., an antigen, an immunogenic epitope), a receptor, an enzyme, a hormone, a transcription factor, a ligand, a membrane transporter, a structural protein, a nuclease, or a component, variant or fragment (e.g., a biologically active fragment) thereof.

E90. The system, or LNP composition of embodiment 88, wherein the therapeutic payload or prophylactic payload comprises a cytokine, or a variant or fragment (e.g., a biologically active fragment) thereof.

E91. The system, or LNP composition of embodiment 88, wherein the therapeutic payload or prophylactic payload comprises an antibody or a variant or fragment (e.g., a biologically active fragment) thereof.

E92. The system, or LNP composition of embodiment 88, wherein the therapeutic payload or prophylactic payload comprises a vaccine (e.g., an antigen, an immunogenic epitope), or a component, variant or fragment (e.g., a biologically active fragment) thereof.

E93. The system, or LNP composition of any one of the preceding embodiments, wherein the therapeutic payload or prophylactic payload comprises a protein or peptide.

E94. The system, or LNP composition of any one of the preceding embodiments, which results in one, two, three, four, five, six or all, or any combination thereof, of the following in a cell (e.g., in a cell contacted with the system or LNP composition):

- (i) increased expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (ii) sustained expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (iii) increased expression and/or level of therapeutic payload or prophylactic payload;

- (iv) sustained expression and/or level of therapeutic payload or prophylactic payload;
- (v) increased stability of mRNA encoding the therapeutic payload or prophylactic payload;
- 5 (vi) increased resistance of translation of therapeutic payload or prophylactic payload to cellular environment, e.g., stress or nutrient deprivation or translation factor availability;
- (vii) reduced dosing of the therapeutic payload or prophylactic payload; or
- (viii) reduced toxicity, e.g., reduced modulation of a protein translated from
- 10 endogenous mRNA in a cell.

E95. The system, or LNP composition of embodiment 94, wherein any one, or all of (i)-(vii) is compared to a cell which:

- (a) has not been contacted with the system of embodiment 7;
- 15 (b) has not been contacted with the LNP composition of embodiment 1;
- (c) has not been contacted with an LNP comprising the first polynucleotide; or
- (d) has been contacted with an LNP comprising the first polynucleotide but has not been contacted with the second polynucleotide, e.g., an LNP comprising the second polynucleotide.

20

E96. The system, or LNP composition of embodiment 94 or 95, wherein the system, or LNP composition results in increased expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., as measured by an assay in Example 2 or 3.

25

E97. The system, or LNP composition of embodiment 96, wherein the increased expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding

30 element of the first polynucleotide.

E98. The system, or LNP composition of embodiment 96 or 97, wherein the increase in expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 30, 35, 40 or 50 fold.

- 5 E99. The system, or LNP composition of embodiment 96 or 97, wherein the increase in expression and/or level of the mRNA comprises an increase in stability (e.g., half-life) of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 20 fold increase in stability of the mRNA encoding the therapeutic payload or prophylactic payload.

10

E100. The system, or LNP composition of embodiment 99, wherein the mRNA encoding the therapeutic payload or prophylactic payload has a half-life of about 3-25 hours, about 4-20 hours, about 4-15 hours, about 5-10 hours, about 6-9 hours or about 7-8 hours.

15

E101. The system, or LNP composition of embodiment 94 or 95, wherein the system or LNP composition results in sustained expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., as measured by an assay in Example 4.

20

E102. The system, or LNP composition of embodiment 101, wherein at least 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more of the expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is sustained for a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 24 or 36 hours.

25

E103. The system, or LNP composition of embodiment 102, wherein the sustained expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding
30 element of the first polynucleotide.

E104. The system, or LNP composition of embodiment 101, wherein the system, or LNP composition results in a decreased loss, e.g., about a 1.2-fold, 2-fold, 3-fold, 4-fold or 5-fold decrease in loss, of mRNA encoding the therapeutic payload or prophylactic payload.

5

E105. The system, of LNP composition of embodiments 104, wherein the decrease in loss of mRNA encoding the therapeutic payload or prophylactic payload occurs over a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, or 24 hours.

10 E106. The system, or LNP composition of embodiment 101, wherein the system, or LNP composition results in a decreased loss, e.g., about a 1.2-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold decrease in loss, of translating mRNA.

15 E107. The system, or LNP composition of embodiment 106, wherein the decrease in loss of mRNA encoding the therapeutic payload or prophylactic payload occurs over a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, or 24 hours.

20 E108. The system, or LNP composition of embodiment 101, wherein the system, or LNP composition results in a sustained, e.g., maintained, level of translation of an mRNA encoding the therapeutic payload or prophylactic payload.

E109. The system, or LNP composition of embodiment 94 or 95, wherein the system results in increased expression and/or level of the therapeutic payload or prophylactic payload, e.g., increased protein level, translation, or half-life, e.g., as measured by an
25 assay of Example 4.

5 E110. The system, or LNP composition of embodiment 109, wherein the increased expression and/or level of the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding element of the first polynucleotide.

10 E111. The system, or LNP composition of embodiment 109 or 110, wherein the increase in expression and/or level of the therapeutic payload or prophylactic payload is about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 30, 35, 40 or 50 fold.

15 E112. The system, or LNP composition of embodiment 94 or 95, wherein the system results in sustained expression and/or level of the therapeutic payload or prophylactic payload.

20 E113. The system, or LNP composition of any one of the preceding embodiments, wherein:

(a) the first polynucleotide comprises:

(1) a sequence encoding a therapeutic payload or prophylactic payload,
and

25 (2) a binding element comprising an MS2 sequence, e.g., 6 MS2 sequences of 19 nucleotides separated by spacers of 20 nucleotides in length;

(b) the second polynucleotide comprises a sequence encoding:

(1) an effector molecule comprising eIF4G, e.g., wildtype eIF4G, a variant or a fragment thereof; and

30 (2) a tether molecule comprising MBP, e.g., wildtype MBP, a variant or fragment thereof.

E114. The system, of LNP composition of embodiment 113, wherein the first polynucleotide and the second polynucleotide each comprises an mRNA.

35

E115. The system, of LNP composition of embodiment 113, wherein the tether molecule of (b)(2) binds to, e.g., recognizes, the binding element of (a)(2).

5 E116. The system of embodiment 113, wherein the first polynucleotide is formulated as an LNP or the second polynucleotide is formulated as an LNP.

E117. The system of embodiment 113, wherein the first polynucleotide is formulated as an LNP and the second polynucleotide is formulated as an LNP.

10 E118. The system, or LNP composition of embodiment 117, wherein the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are the same.

15 E119. The system, or LNP composition of embodiment 117, wherein the first polynucleotide and the second polynucleotide are each formulated as separate LNPs, optionally wherein:

(i) the LNP comprising the first polynucleotide is in a first composition and the LNP comprising the second polynucleotide is in a separate composition; or

20 (ii) the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are in the same composition.

E120. The system, or LNP composition of any one of embodiments 113-119, wherein the binding element comprises a sequence comprising 19 nucleotides, e.g., a binding element nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%,
25 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

E121. The system, or LNP composition of any one of embodiments 113-120, wherein the spacer sequence comprises 20 nucleotides, e.g., a spacer sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100%
30 identity thereof.

E122. The system, or LNP composition of any one of embodiments 113-121, wherein the effector molecule comprises an amino acid sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

5

E123. The system, or LNP composition of any one of embodiments 113-122, wherein the effector molecule is encoded by a nucleotide sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

10

E124. The system, or LNP composition of any one of embodiments 113-123, wherein the tether molecule comprises an amino acid sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

15

E125. The system, or LNP composition of any one of embodiments 113-124, wherein the tether molecule is encoded by a nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

20

E126. The system, or LNP composition of any one of the preceding embodiments, wherein the first polynucleotide comprises an mRNA comprising at least one chemical modification.

25

E127. The system, or LNP composition of any one of the preceding embodiments, wherein the second polynucleotide comprises an mRNA comprising at least one chemical modification.

- E128. The system, or LNP composition of embodiment 126 or 127, wherein the chemical modification is selected from the group consisting of pseudouridine, NI-methylpseudouridine, 2-thiouridine, 4'-thiouridine, 5-methylcytosine, 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-methyluridine, 5-methoxyuridine, and 2'-O-methyl uridine.
- 10 E129. The system, or LNP composition of embodiment 128, wherein the chemical modification is selected from the group consisting of pseudouridine, NI-methylpseudouridine, 5-methylcytosine, 5-methoxyuridine, and a combination thereof.
- E130. The system, or LNP composition of embodiment 128, wherein the chemical
15 modification is NI-methylpseudouridine.
- E131. The LNP composition of any one of the preceding embodiments, wherein the mRNA comprises fully modified NI-methylpseudouridine.
- 20 E132. The system of any one of embodiments 126-131, or the LNP composition of any one of the preceding embodiments, wherein the LNP composition comprises: (i) an ionizable lipid, e.g., an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-lipid.
- 25 E133. The system or LNP composition of embodiment 132, wherein the ionizable lipid comprises an amino lipid.
- E134. The system or LNP composition of embodiment 132 or 133, wherein the ionizable lipid comprises a compound of any of Formulae (I), (IA), (IB), (II), (IIa), (IIb), (IIc), (IId), (IIe), (IIf), (IIg), (III), (IIIa1), (IIIa2), (IIIa3), (IIIa4), (IIIa5), (IIIa6), (IIIa7), or (IIIa8).
- 30

E135. The system or LNP composition of any one of embodiments 132-134, wherein the ionizable lipid comprises a compound of Formula (I).

- 5 E136. The system or LNP composition of any one of embodiments 132-135, wherein the ionizable lipid comprises a compound of Formula (IIa).

E137. The system or LNP composition of any one of embodiments 132-135, wherein the ionizable lipid comprises a compound of Formula (IIe).

10

E138. The system or LNP composition of any one of embodiments 132-137, wherein the non-cationic helper lipid or phospholipid comprises a compound selected from the group consisting of DSPC, DPPC, or DOPC.

- 15 E139. The system or LNP composition of 138, wherein the phospholipid is DSPC, e.g., a variant of DSPC, e.g., a compound of Formula (IV).

E140. The system or LNP composition of any one of embodiments 132-141, wherein the structural lipid is chosen from alpha-tocopherol, β -sitosterol or cholesterol.

20

E141. The system or LNP composition of embodiment 140, wherein the structural lipid is alpha-tocopherol.

- E142. The system or LNP composition of embodiment 140, wherein the structural lipid
25 is β -sitosterol.

E143. The system or LNP composition of embodiment 140, wherein the structural lipid is cholesterol.

E144. The system or LNP composition of any one of embodiments 132-143, wherein the PEG lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof.

E145. The system or LNP composition of embodiments 144, wherein the PEG lipid is selected from the group consisting of PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC and PEG-DSPE lipid.

E146. The system or LNP composition of embodiment 145, wherein the PEG-lipid is PEG-DMG.

E147. The system or LNP composition of embodiments 144, wherein the PEG lipid is chosen from a compound of: Formula (V), Formula (VI-A), Formula (VI-B), Formula (VI-C) or Formula (VI-D).

E148. The system or LNP composition of embodiment 147, wherein the PEG lipid is a compound of Formula (VI-A).

E149. The system or LNP composition of embodiment 147, is a compound of Formula (VI-B).

E150. The system or LNP composition of any one of embodiments 132-149, wherein the LNP comprises a molar ratio of about 20-60% ionizable lipid: 5-25% phospholipid: 25-55% cholesterol; and 0.5-15% PEG lipid.

E151. The system or LNP composition of embodiment 149, wherein the LNP comprises a molar ratio of about 50% ionizable lipid: about 10% phospholipid: about 38.5% cholesterol; and about 1.5% PEG lipid.

E152. The system or LNP composition of embodiment 149 or 150, wherein the LNP comprises a molar ratio of about 49.83% ionizable lipid: about 9.83% phospholipid: about 30.33% cholesterol; and about 2.0% PEG lipid.

- 5 E153. The system or LNP composition of any one of the preceding embodiments, which is formulated for intravenous, subcutaneous, intramuscular, intranasal, intraocular, rectal, pulmonary or oral delivery.

- E154. The system or LNP composition of any one of the preceding embodiments,
10 which is formulated for intravenous delivery.

E155. The system or LNP composition of any one of the preceding embodiments, further comprising a pharmaceutically acceptable carrier or excipient.

- 15 E156. The system, or LNP composition of any one of the preceding embodiments, wherein the polynucleotide, e.g., the first and/or second polynucleotide comprises a cap, a 3' UTR, a 5' UTR, a Poly A tail and/or a micro RNA (miRNA) binding site.

- E157. The system, or LNP composition of embodiment 156, wherein the cap comprises
20 a cap disclosed herein.

E158 The system, or LNP composition of embodiment 156, wherein the polynucleotide, e.g., the first and/or second polynucleotide does not comprise a cap.

- 25 E159. The system, or LNP composition of any one of embodiments 156-158, wherein the 3' UTR comprises a 3' UTR disclosed herein, e.g., a v1.1 3' UTR or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

- E160. The system, or LNP composition of any one of embodiments 156-159, wherein
30 the 5' UTR comprises a 5' UTR disclosed herein.

E161. The system, or LNP composition of any one of embodiments 156-160, wherein the Poly A tail comprises a Poly A tail sequence disclosed herein or a fragment thereof.

5 E162. The system, or LNP composition of any one of embodiments 156-160, wherein the polynucleotide, e.g., the first and/or second polynucleotide does not comprise a Poly A tail.

E163. The system, or LNP composition of any one of embodiments 156-162, wherein the miRNA binding site comprises a miRNA binding site disclosed herein.

10

E164. The system, or LNP composition of any one of the preceding embodiments, wherein the polynucleotide, e.g., the first and/or second polynucleotide is a circular polynucleotide.

15 E165. The system, or LNP composition of embodiment E165, wherein the first polynucleotide is a circular polynucleotide.

E166. The system, or LNP composition of embodiment E165, wherein the second polynucleotide is a circular polynucleotide.

20

E167. A pharmaceutical composition comprising the system, or LNP composition of any one of the preceding embodiments.

25 E168. A cell comprising a system, or LNP composition of any one of the preceding embodiments.

E169. The cell of embodiment 168, which has been contacted with the system, or LNP composition.

E170. The cell of embodiment 168 or 169, which is maintained under conditions sufficient to allow for expression of one or both polynucleotides of the system, or LNP composition.

- 5 E170. A method of increasing expression of a therapeutic payload or prophylactic payload in a cell, comprising administering to the cell a system, or LNP composition of any one of embodiments 1-166.

- 10 E171. The LNP composition or system of any one of embodiments 1-166, for use in a method of increasing expression of a therapeutic payload or prophylactic payload in a cell.

- 15 E172. A method of increasing expression of a therapeutic payload or prophylactic payload, in a subject, comprising administering to the subject an effective amount of a system or LNP composition of any one of embodiments 1-166.

- 20 E173. The LNP composition or system of any one of embodiments 1-166, for use in a method of increasing expression of a therapeutic payload or prophylactic payload in a subject.

- E174. A method of delivering a system, or LNP composition of any one of embodiments 1-166, to a cell.

- 25 E175. The LNP composition or system of any one of embodiments 1-166, for use in a method of delivering the system or LNP composition to a cell.

- 30 E176. The method of embodiment 174, or the LNP composition or system for use of embodiment 175, comprising contacting the cell in vitro, in vivo or ex vivo with the system or LNP composition.

E177. A method of delivering a system or LNP composition of any one of embodiments 1-166, to a subject having a disease or disorder, e.g., as described herein.

5 E178. The LNP composition or system of any one of embodiments 1-166, for use in a method of delivering the system or LNP composition to a subject having a disease or disorder, e.g., as described herein.

E179. A method of modulating an immune response in a subject, comprising administering to the subject in need thereof an effective amount of a system, or LNP
10 composition of any one of embodiments 1-166.

E180. The LNP composition or system of any one of embodiments 1-166, for use in a method of modulating an immune response in a subject, comprising administering to the subject an effective amount of the system, or LNP composition.

15

E181. A method of treating, preventing, or preventing a symptom of, a disease or disorder comprising administering to a subject in need thereof an effective amount of a system, or LNP composition of any one of embodiments 1-166.

20 E182. The LNP composition or system of any one of embodiments 1-166, for use in a method of treating, preventing, or preventing a symptom of, a disease or disorder in a subject, comprising administering to the subject in need thereof an effective amount of the system, or LNP composition.

25 E184. The method, or system for use of any one of embodiments 170-183, wherein the first polynucleotide and/or the second polynucleotide of the system is formulated as an LNP.

E185. The method of embodiment 184, wherein both the first and the second
30 polynucleotides of the system are each formulated as LNPs, e.g., the same or different LNPs.

E186. The method, or the LNP composition or system for use of any one of embodiments 170-185, wherein the LNP comprising (a) and the LNP comprising (b) are administered simultaneously, e.g., substantially simultaneously.

5

E187. The method, or the LNP composition or system for use of any one of embodiments 170-185, wherein the LNP comprising (a) and the LNP comprising (b) are administered sequentially.

10 E188. The method, or the LNP composition or system for use of embodiment 187, wherein the LNP comprising (a) is administered first.

E189. The method, or the LNP composition or system for use of embodiment 187 or 188, wherein the LNP comprising (a) is administered first followed by administration of
15 the LNP comprising (b).

E190. The method, or the LNP composition or system for use of embodiment 187, wherein the LNP comprising (b) is administered first.

20 E191. The method, or the LNP composition or system for use of embodiment 187 or 190, wherein the LNP comprising (b) is administered first followed by administration of the LNP comprising (a).

E192. The method, or the LNP composition or system for use of any one of
25 embodiments 170-191, wherein the LNP comprising the first polynucleotide is administered at a lower dose compared to a reference LNP.

E193. The method, or the LNP composition or system for use of embodiment 192, wherein the LNP comprising the first polynucleotide is administered at a dose that is at
30 least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% lower compared to the dose of a reference LNP.

E194. The method, or the LNP composition or system for use of embodiment 192 or 193, wherein the reference LNP is chosen from: an otherwise similar LNP comprising a polynucleotide which does not have the binding element of the first polynucleotide; or
5 an LNP that does not comprise the second polynucleotide.

E195. The method, or the LNP composition or system for use of any one of embodiments 170-191, wherein the LNP comprising the first polynucleotide is administered at a higher dose compared to the LNP comprising the second
10 polynucleotide.

E196. The method or the LNP composition or system for use of embodiment 195, wherein the LNP comprising the first polynucleotide is administered at a dose that is at least 10%, 20%, 30%,
15 40%, 50%, 60%, 70%, 80%, or 90% higher compared to the dose of the LNP comprising the second polynucleotide.

E197. The method, or the LNP composition or system for use of any one of
20 embodiments 170-191, wherein the LNP comprising the first polynucleotide is in molar excess compared to the LNP comprising the second polynucleotide.

E198. The method or the LNP composition or system for use of embodiment 197, wherein the LNP comprising the first polynucleotide is in about 1-800X molar excess,
25 compared to the LNP comprising the second polynucleotide.

E199. The method or the LNP composition or system for use of embodiment 197 or 198, wherein the LNP comprising the first polynucleotide is in about 1-750x, about 2-700x, about 3-650 x, about 4-600 x, about 5-550 x, about 6-500 x, about 7-450 x, about 8-400 x, about 10-350 x, about 15-300 x, about 20-250 x, about 25-200 x, about 30-150
5 x, about 35-100 x, about 40-90 x, about 45-80 x, about 50-75 x, about 60-70x molar excess compared to the LNP comprising the second polynucleotide.

E200. The method or the LNP composition or system for use of embodiment 197 or 198, wherein the LNP comprising the first polynucleotide is in about 2x, about 3x,
10 about 4x, about 5 x, about 6x, about 7x, about 8x, about 9 x, about 10x, about 11 x, about 12 x, about 13 x, about 14 x, about 15 x, about 20 x, about 25 x, about 30 x, about 35 x, about 40 x, about 50 x, about 60 x, about 70 x, about 80 x, about 90 x, about 100 x, about 150 x, about 200 x, about 250 x, about 300 x, about 350 x, about 400 x, about 450 x, about 500 x, about 600 x, about 650 x, about 700 x, about 750 x, or about 800x
15 molar excess compared to the LNP comprising the second polynucleotide.

E201. The method or the LNP composition or system for use of any one of embodiments 197-200, wherein the LNP comprising the first polynucleotide is in about 9x molar excess compared to the LNP comprising the second polynucleotide.
20

E202. The method or the LNP composition or system for use of any one of embodiments 197-200, wherein the LNP comprising the first polynucleotide is in about 10x molar excess compared to the LNP comprising the second polynucleotide.

25 E203. The method or the LNP composition or system for use of any one of embodiments 197-200, wherein the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are at the same molar amounts.

E204. The method, or the LNP composition or system for use of any one of embodiments 171-203, which results in one, two, three, four, five, six or all, or any combination thereof, of the following in a cell (e.g., in a cell contacted with the system or LNP composition):

5

E205. The method, or the LNP composition or system for use of embodiment 204, wherein any one, or all of (i)-(vii) is compared to a cell which:

- (a) has not been contacted with the system of embodiment 7;
- (b) has not been contacted with the LNP composition of embodiment 1;
- 10 (c) has not been contacted with an LNP comprising the first polynucleotide; or
- (d) has been contacted with an LNP comprising the first polynucleotide but has not been contacted with the second polynucleotide, e.g., an LNP comprising the second polynucleotide.

- 15 E206. The method, or the LNP composition or system for use of embodiment 204 or 205, wherein the system, or LNP composition results in increased expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., as measured by an assay in Example 2 or 3.

- 20 E207. The method, or the LNP composition or system for use of embodiment 206, wherein the increased expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding element of the first polynucleotide.

25

E208. The method, or the LNP composition or system for use of embodiment 206 or 207, wherein the increase in expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 30, 35, 40 or 50 fold.

30

E209. The method, or the LNP composition or system for use embodiment 206 or 207, wherein the increase in expression and/or level of the mRNA comprises an increase in stability (e.g., half-life) of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 20 fold increase in stability of the mRNA encoding the therapeutic payload or prophylactic payload.

E210. The method, or the LNP composition or system for use of embodiment 209, wherein the mRNA encoding the therapeutic payload or prophylactic payload has a half-life of about 3-25 hours, about 4-20 hours, about 4-15 hours, about 5-10 hours, about 6-9 hours or about 7-8 hours.

E211. The method, or the LNP composition or system for use of embodiment 204 or 205, wherein the system or LNP composition results in sustained expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload, e.g., as measured by an assay in Example 4.

E212. The method, or the LNP composition or system for use of embodiment 211, wherein at least 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more of the expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is sustained for a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 24 or 36 hours.

E213. The method, or the LNP composition or system for use of embodiment 212, wherein the sustained expression and/or level of the mRNA encoding the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding element of the first polynucleotide.

E214. The method, or the LNP composition or system for use of embodiment 211, wherein the system, or LNP composition results in a decreased loss, e.g., about a 1.2-fold, 2-fold, 3-fold, 4-fold or 5-fold decrease in loss, of mRNA encoding the therapeutic payload or prophylactic payload.

5

E215. The method, or the LNP composition or system for use of embodiments 214, wherein the decrease in loss of mRNA encoding the therapeutic payload or prophylactic payload occurs over a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, or 24 hours.

10

E216. The method, or the LNP composition or system for use of embodiment 211, wherein the system, or LNP composition results in a decreased loss, e.g., about a 1.2-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold decrease in loss, of translating mRNA.

15

E217. The method, or the LNP composition or system for use of embodiment 216, wherein the decrease in loss of mRNA encoding the therapeutic payload or prophylactic payload occurs over a period of time, e.g., about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, or 24 hours.

20

E218. The method, or the LNP composition or system for use of embodiment 211, wherein the system, or LNP composition results in a sustained, e.g., maintained, level of translation of an mRNA encoding the therapeutic payload or prophylactic payload.

25

E219. The method, or the LNP composition or system for use of embodiment 204 or 205, wherein the system results in increased expression and/or level of the therapeutic payload or prophylactic payload, e.g., increased protein level, translation, or half-life, e.g., as measured by an assay of Example 4.

E220. The method, or the LNP composition or system for use of embodiment 219, wherein the increased expression and/or level of the therapeutic payload or prophylactic payload is compared to an otherwise similar cell that has been contacted with an mRNA encoding the therapeutic payload or prophylactic payload which mRNA lacks a binding
5 element of the first polynucleotide.

E221. The method, or the LNP composition or system for use of embodiment 219 or 220, wherein the increase in expression and/or level of the therapeutic payload or prophylactic payload is about 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 30, 35, 40 or 50 fold.
10

E222. The method, or the LNP composition or system for use of embodiment 204 or 205, wherein the system results in sustained expression and/or level of the therapeutic payload or prophylactic payload.

15 E223. The method, or the LNP composition or system for use of any one of embodiments 171-222, wherein the first polynucleotide comprises an mRNA comprising at least one chemical modification.

E224. The method, or the LNP composition or system for use of any one of
20 embodiments 171-222, wherein the second polynucleotide comprises an mRNA comprising at least one chemical modification.

E225. The method, or the LNP composition or system for use of embodiment 223 or 224 , wherein the chemical modification is selected from the group consisting of
25 pseudouridine, N1-methylpseudouridine, 2-thiouridine, 4'-thiouridine, 5-methylcytosine, 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,
30 dihydropseudouridine, 5-methyluridine, 5-methyluridine, 5-methoxyuridine, and 2'-O-methyl uridine.

E226. The method, or the LNP composition or system for use of embodiment 225, wherein the chemical modification is selected from the group consisting of pseudouridine, N1-methylpseudouridine, 5-methylcytosine, 5-methoxyuridine, and a
5 combination thereof.

E227. The method, or the LNP composition or system for use of embodiment 225, wherein the chemical modification is N1-methylpseudouridine.

10 E228. The method, or the LNP composition or system for use of any one of embodiments 171-227, wherein the mRNA comprises fully modified N1-methylpseudouridine.

E229. The method, or the LNP composition or system for use of any one of
15 embodiments 171-228, wherein the LNP composition comprises: (i) an ionizable lipid, e.g., an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-lipid.

E230. The method, or the LNP composition or system for use of embodiment 229,
20 wherein the ionizable lipid comprises an amino lipid.

E231. The method, or the LNP composition or system for use of embodiment 229 or 230, wherein the ionizable lipid comprises a compound of any of Formulae (I), (IA), (IB), (II), (IIa), (IIb), (IIc), (IId), (IIe), (IIf), (IIg), (III), (IIIa1), (IIIa2), (IIIa3), (IIIa4),
25 (IIIa5), (IIIa6), (IIIa7), or (IIIa8).

E232. The method, or the LNP composition or system for use of any one of embodiments 229-231, wherein the ionizable lipid comprises a compound of Formula (I).

30

E233. The method, or the LNP composition or system for use of any one of embodiments 229-232, wherein the ionizable lipid comprises a compound of Formula (IIa).

- 5 E234. The method, or the LNP composition or system for use of any one of embodiments 229-232, wherein the ionizable lipid comprises a compound of Formula (IIe).

- 10 E235. The method, or the LNP composition or system for use of any one of embodiments 229-234, wherein the non-cationic helper lipid or phospholipid comprises a compound selected from the group consisting of DSPC, DPPC, or DOPC.

- 15 E236. The method, or the LNP composition or system for use of embodiment 235, wherein the phospholipid is DSPC, e.g., a variant of DSPC, e.g., a compound of Formula (IV).

- 20 E237. The method, or the LNP composition or system for use of any one of embodiments 229-238, wherein the structural lipid is chosen from alpha-tocopherol, β -sitosterol or cholesterol.

E238. The method, or the LNP composition or system for use of embodiment 237, wherein the structural lipid is alpha-tocopherol.

- 25 E239. The method, or the LNP composition or system for use of embodiment 237, wherein the structural lipid is β -sitosterol.

E240. The method, or the LNP composition or system for use of embodiment 237, wherein the structural lipid is cholesterol.

E241. The method, or the LNP composition or system for use of any one of
embodiments 229-240, wherein the PEG lipid is selected from the group consisting of a
PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-
modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a
5 PEG-modified dialkylglycerol, and mixtures thereof.

E242. The method, or the LNP composition or system for use of embodiment 241,
wherein the PEG lipid is selected from the group consisting of PEG-c-DOMG, PEG-
DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC and PEG-DSPE lipid.

10

E243. The method, or the LNP composition or system for use of embodiment 241 or
242, wherein the PEG-lipid is PEG-DMG.

E244. The method, or the LNP composition or system for use of any one of
15 embodiments 229-240, wherein the PEG lipid is a compound chosen from: Formula
(V), Formula (VI-A), Formula (VI-B), Formula (VI-C) or Formula (VI-D).

E245. The method, or the LNP composition or system for use of embodiment 244,
wherein the PEG-lipid is a compound of Formula (VI-A).

20

E246. The method, or the LNP composition or system for use of embodiment 244,
wherein the PEG-lipid is a compound of Formula (VI-B).

E247. The method, or the LNP composition or system for use of any one of
25 embodiments 229-246, wherein the LNP comprises a molar ratio of about 20-60%
ionizable lipid: 5-25% phospholipid: 25-55% cholesterol; and 0.5-15% PEG lipid.

E248. The method, or the LNP composition or system for use of embodiment 247,
wherein the LNP comprises a molar ratio of about 50% ionizable lipid: about 10%
30 phospholipid: about 38.5% cholesterol; and about 1.5% PEG lipid.

E249. The method, or the LNP composition or system for use of embodiment 247 or 248, wherein the LNP comprises a molar ratio of about 49.83% ionizable lipid: about 9.83% phospholipid: about 30.33% cholesterol; and about 2.0% PEG lipid.

- 5 E250. The method, or the LNP composition or system for use of any one of embodiments 171-249, wherein the LNP or system is formulated for intravenous, subcutaneous, intramuscular, intranasal, intraocular, rectal, pulmonary or oral delivery.

- 10 E251. The method, or the LNP composition or system for use of any one of embodiments 171-250, wherein the subject is a mammal, e.g., a human.

E252. The method, or the LNP composition or system for use of any one of embodiments 171-251, wherein the subject has a disease or disorder disclosed herein.

- 15 E253. The LNP composition of embodiment 2, wherein the effector molecule recognizes the binding element.

E254. The LNP composition of any one of embodiments 1-4, wherein the second polynucleotide is DNA.

20

E255. The LNP composition of embodiment 254, wherein the sequence encoding the effector molecule is under the control of a tissue-specific promoter.

- 25 E256. The LNP composition of any one of embodiments 1-4, wherein expression or recruitment of the effector molecule is under the control of a trigger in a specific microenvironment or specific cell-type.

E257. The LNP composition of embodiment 256, wherein the trigger is microRNA, receptor-mediated activation, and/or a change in pH and/or hypoxia.

30

E258. A system comprising:

(a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or a prophylactic payload, and (2) a binding element; and/or

(b) a second polynucleotide comprising a sequence encoding an effector molecule,

5 optionally wherein, (a) and (b) each comprise an mRNA.

E259. The system of embodiment E258, wherein the effector molecule further comprises a polypeptide that recognizes the binding element (a tether molecule).

10 E260. The system of embodiment E258, wherein the effector molecule recognizes the binding element.

E261. The system of any embodiment 7, wherein the second polynucleotide is DNA.

15 E262. The system of embodiment 7, wherein the effector molecule is under the control of a tissue-specific promoter.

E263. The system of embodiment 262, wherein expression of the effector molecule or recruitment of the effector molecule is under the control of a trigger in a specific
20 microenvironment or specific cell-type.

E264. The system of embodiment 263, wherein the trigger is microRNA, receptor-mediated activation, and/or a change in pH and/or hypoxia.

25 E265. The system, or LNP composition of any one of embodiments 1-125 and 128-264, wherein the first polynucleotide comprises an mRNA which does not have any chemical modification.

E266. The system, or LNP composition of any one of embodiments 1-125 and 128-264,
30 wherein the second polynucleotide comprises an mRNA which does not have any chemical modification.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 provides a schematic of the 2 RNA system. Target mRNA with MS2 loops in the 3'UTR and another mRNA encoding MS2-binding protein (MBP) fused to truncated eIF4G (eIF4GAN) are co-delivered by transfection or electroporation. The MS2 loops either replace the 3'UTR completely or are added before or after v1.1 UTR sequence.

FIGs. 2A-2D depict increased protein output for target deg-GFP RNA in HeLa cells with tethered eIF4G. Total green intensity vs time is plotted. Target mRNA with MS2 loops in the 3'UTR was co-delivered with another mRNA encoding a control protein (EPO; SEQ ID NO: 81) or a tethered control (MBP-LacZ; SEQ ID NO: 39) or a tethered effector (MBP-eIF4GAN; SEQ ID NO: 11). Experiments were done in 10x molar excess of target. **FIG. 2A** shows the results of a control target mRNA with no binding sites in the 3'UTR (the '3 UTR comprises a v1.1 sequence; SEQ ID NO: 4). **FIG. 2B** shows the results of a target mRNA with MS2 binding sites upstream of the v1.1 sequence in the 3' UTR. **FIG. 2C** shows the results of a target mRNA with MS2 binding sites in the 3' UTR. **FIG. 2D** shows the results of a target mRNA with MS2 binding sites downstream of the v1.1 sequence in the 3' UTR. In this construct, the MS2 binding sites were adjacent to a PolyA sequence.

FIGs. 3A-3C depict increased protein output for target deg-GFP RNA in HeLa, HEK293 and Hep3b cells, with tethered eIF4G. Total AUC is plotted. 3'UTR status of the target mRNA is indicated on the x axis. Experiments were done in 3x molar excess of target. **FIG. 3A** shows deg-GFP RNA protein output in HeLa cells. **FIG. 3B** shows deg-GFP RNA protein output in HeLa cells. **FIG. 3C** shows deg-GFP RNA protein output in HeLa cells.

FIGs. 4A-4B shows that increasing the ratio of tethered effector increases the total protein output in HeLa cells, and half-life predicted by the model. 3'UTR status of the target deg-GFP mRNA is indicated in the schematic. **FIG. 4A** shows total green intensity over time. **FIG. 4B** shows predicted half-lives (hours) using the 4-parameter model. Values of half-life obtained under each condition are added as data labels. (Target mRNA was added at the same amount in all conditions, amounts of effector RNA are changing as shown).

FIGs. 5A-5B show Total green intensity vs time for target RNA co-delivered with different RNAs as depicted (nature of UTR is depicted on top of the panels, and **FIGs. 5C** and **5D** show the % of the starting quantity of the target mRNA remaining in HEK293 cells with tethered eIF4G. HEK293 cells were electroporated with target reporter RNA and an mRNA encoding control protein (EPO; SEQ ID NO: 81); or an mRNA encoding tethered control protein (MBP-LacZ; SEQ ID NO: 39); or an mRNA encoding tethered effector protein (MBP-eIF4GΔN; SEQ ID NO: 11). The experiments were done in 3x molar excess of target RNA. **FIGs. 5A-5B** show that both the amount of protein and duration of expression was increased in the presence of the tethered effector for the target RNA with MS2 sites in the UTR. **FIGs. 5C-5D** show that the % of the starting quantity of the target mRNA with MS2 sites reduced much more slowly with time for tethered effector condition. This suggests an increase in half-life of the target RNA.

FIGs. 6A-6E depict maintenance of robust translation for target mRNAs at later time points in Hep3b with a tethered effector. Hep3b cells were imaged at the indicated time points after electroporation with target mRNA (MS2 loops containing UTR) with non-tethered control (EPO; SEQ ID NO: 81), tethered control (MBP-LacZ; SEQ ID NO: 39) or tethered effector (MBP-eIF4GΔN; SEQ ID NO: 11). All experiments were done with 1.5x molar target excess. Control cells were electroporated with Luciferase (FLuc) or no RNAs. Images were analyzed in the red and green channel for smFISH and V5 signal respectively. **FIG. 6A** shows tethered effector reduces the rate of target mRNA decay. **FIG. 6B** shows tethered effector reduces the loss of translating mRNAs with time. **FIG. 6C** shows a higher percentage of cells maintained robust translation at later time points with tethered effector. **FIG. 6D** shows tethered effector maintained translation per target mRNA over time. **FIG. 6E** shows similar data as **FIG. 6D**, average NPI intensity per cell is plotted. N in the figures represents the number of cells assessed. Colors represent target RNA transfection with RNA encoding non-tethered unrelated protein (black), tethered control protein (orange) and tethered effector protein (green).

FIGs. 7A-7D depict maintenance of robust translation for target mRNAs at later time points in HeLa with tethered effector. HeLa cells were imaged at the indicated time points after electroporation with target mRNA (MS2 loops containing UTR) with non-tethered control (EPO; SEQ ID NO: 81), tethered control (MBP-LacZ; SEQ ID NO: 39) or tethered effector (MBP- eIF4GΔN; SEQ ID NO: 11). All experiments were done with 1.5x molar target excess. Control cells were electroporated with Luciferase (Fluc) or no RNAs. Images were analyzed in the red and green channel for smFISH and V5 signal respectively. **FIG. 7A** shows no appreciable change in cytosolic mRNAs over time in any condition. **FIG. 7B** shows no appreciable change in translation per mRNA with time in any condition. **FIG. 7C** shows tethered effector maintained higher percentage of translating mRNAs over time. **FIG. 7E** shows a higher percentage of cells maintained robust translation at later time points with tethered eIF4G.

FIGs. 8A-8E depict the results of the experiments to identify the domain of eIF4G required for effector function. **FIG. 8A** provides a schematic of the constructs that were used. Target deg-GFP RNA was co-delivered with an mRNA encoding control protein, EPO (SEQ ID NO: 81); or an mRNA encoding tethered control protein, MBP-LacZ (SEQ ID NO: 39); or an mRNA encoding tethered effector protein, MBP-eIF4GΔN (SEQ ID NO: 11), or mRNAs encoding MBP-fused to eIF4G truncations and mutations as illustrated in **FIG. 8A**. The experiment was done in 10x molar excess of target RNA. **FIGs. 8B** and **8D** show the results in HeLa cells and **FIGs. 8C** and **8E** show the results in HEK293 cells.

FIG. 9 is a table depicting summary data of the binding site experiments of **FIGs. 8A-8E**.

FIGs. 10A-10C depict the tethering system and results of experiments done when the target RNA had 24MS2 sites. **FIG. 10A** is a schematic of the tethering system where RNA encoding target protein shows no detectable protein expression. This RNA can be activated upon interaction with a specific effector protein. The
 5 interaction is mediated by known RNA-binding protein-RNA interaction tether (MBP protein-MS2 stem loop structure in RNA). **FIG. 10B** depicts real time fluorescence curves for HeLa cells transfected with 2 RNAs. Each sample has target deg-GFP encoding RNA (3' UTR_v1.1; SEQ ID NO: 4 or 3'UTR_24 MS2; SEQ ID NO: 154) co-transfected with a control (EPO; SEQ ID NO: 81) or MBP-effector protein (MBP-eIF4GdN; SEQ ID NO: 11) encoding RNA. **FIG. 10C** shows the AUC for the same data in **FIG. 10A**.

FIG. 11 provides a schematic of a system to recruit potential effectors to target RNA with an MS2 binding protein (MBP)-MS2 tether. The target mRNA has MS2 loops in the 3'UTR and lacks a polyA tail (A0). The second mRNA encodes MS2-
 15 binding protein (MBP) fused to an effector/ eIF4G-mid (a truncated fragment of the eIF4G protein). The RNA-binding protein, MBP tethers the effector, eIF4G to target RNA via the recognition of the MS2 hairpins. This system can be coupled to a miRNA-dependent switch gate to permit tethering in specific cells, thereby turning ON expression of the target protein. nt:
 20 nucleotide; aa: amino acids.

FIG. 12A-12D depict that co-delivery with tethered effector rescues expression for tailless mRNAs in Hep3b or HeLa cells. Control (A100) target RNA (degGFP with v1.1_A100 tails) or test (A0) target RNA (degGFP with 6xMS2_A0 tails; SEQ ID NO: 3) were co-delivered
 25 with another mRNA encoding a control protein (t-ctrl; MBP-LacZ; SEQ ID NO: 39) or a tethered effector (t-eff; MBP-eIF4GMID2; SEQ ID NO: 23). Experiments were done in 10x molar excess of target. **FIG. 12A** shows the total green intensity over time in Hep3b cells co-transfected with the indicated RNA constructs. **FIG. 12B** is the AUC for the data in **FIG. 12A**, normalized to AUC obtained with standard A100.

RNAs. **FIG. 12C** shows the total green intensity over time in HeLa cells co-transfected with the indicated RNA constructs. **FIG. 12D** is the AUC for the data in **FIG. 12C**, normalized to AUC obtained with standard A100 RNAs.

FIG. 13 shows that tailless target RNA shows detectable expression only when co-delivered with effector, and overall expression is further improved when tailless RNA is ligated to idT. The figure depicts the level of green fluorescent protein (GFP fluorescence, y-axis) measured at the indicated times (x-axis) after transfection of Hep3b cells with 3'v1.1_MS2_A0 or 3'v1.1_MS2-A0-idT deg-GFP constructs with Effector (t-eff; MBP-mid2; SEQ ID NO: 23) or Control (t-ctrl; MBP-LacZ; SEQ ID NO: 39) RNA.

FIG. 14 shows that capless target RNA shows detectable expression only when co-delivered with effector. The figure depicts the level of Relative Light Units (RLU, y-axis) measured at the indicated times (x-axis) after transfection of HeLa cells with 5'PPP-end NpiLUC constructs with Effector (t-eff; MBP-eIF4G-mid2; SEQ ID NO: 23) or Control (t-ctrl; MBP-eIF4G-mid2mut; SEQ ID NO: 69) RNA.

FIG. 15 shows that capless-tailless target RNA shows detectable expression only when co-delivered with effector. The figure depicts the level of Relative Light Units (RLU, y-axis) measured at the indicated times (x-axis) after transfection of HeLa cells with 5'PPP-A0 ends NpiLUC constructs with Effector (t-eff; MBP-eIF4G-mid2; SEQ ID NO: 23) or Control (t-ctrl; MBP-eIF4G-mid2mut; SEQ ID NO: 69) RNA.

FIG. 16 shows that tethered effector decreases loss of translating mRNAs over time. Hep3b cells were imaged at the indicated time points after electroporation with target mRNA (MS2 loops containing UTR) in combination with non-tethered control (nt-ctrl; EPO), tethered control (t-ctrl; MBP-LacZ; SEQ ID NO: 39) or tethered effector (t-eff; MBP-eIF4GAN; SEQ ID NO: 11). The graph depicts the nascent peptide imaging (NPI) + smFISH + Spots count per cell. All experiments were done with 1.5x molar target excess. Control cells were electroporated with Luciferase (FLuc) or no RNAs. Images were analyzed in the red and green channel for smFISH and V5 signal respectively. The average number of spots positive for both Npi and smFISH per cell is plotted. Data represented as average +/- SEM.

FIGs. 17A-17C shows that the tethered effector maintains more translating mRNAs in more cells over time. Hep3b cells were electroporated with the target RNA (3'v1.1_MS2_A100 or 3'v1.1_A100) in combination with non-tethered control (nt-ctrl; LacZ; SEQ ID NO: 98), tethered control (t-ctrl; MBP-LacZ; SEQ ID NO: 39) or t-effector (MBP-eIF4GΔN; SEQ ID NO: 11). The graphs are cumulative frequency distribution plots showing the percentage of cells (y-axis) against the percentage of translating mRNAs (x-axis) at 4 hours (**FIG. 17A**), 8 hours (**FIG. 17B**) or 12 hours (**FIG. 17C**) post transfection. All experiments were done with 1.5x molar target excess. Control cells were electroporated with Luciferase (FLuc) or no RNAs. The average NPI intensity per cell is plotted at the various time points indicated (4h, 8h, 12h).

FIG. 18 shows that tethering increases secreted protein expression and translation. v1.1 target RNA constructs (SEQ ID NO: 4) with optimized reading frames for the light chain and heavy chain pairs of two secreted antibodies, (Ab1 and Ab2) were co-transfected into Hek293 cells with t-ctrl (MBP-LacZ; SEQ ID NO: 39) or t-eff (MBP-eIF4GΔN; SEQ ID NO: 11). Experiments were done with 5x molar target excess (Ab1) or 1x molar target excess (Ab2). In each of the antibody experiments, the tethered effector is tethered to two separate light and heavy chain encoding RNAs. The graphs show the concentration (μg/ml) of Ab1 or Ab2 over time.

FIG. 19 provides a schematic of the single RNA tethering system. The mRNA molecule from 5' to 3' includes a CAP, a 5'UTR, target ORF, 3 protease cleavage sites in tandem: T2A-P2A-E2A/ TPE (red), another ORF encoding for the RNA binding protein fused to an effector (orange-RBP, -green-Effector), and MS2 loops (orange stripes) in the 3'UTR. The MS2 loops after v1.1 UTR sequence. The TPE protease cleavage site leads to ribosome skipping during translation in a cell. The end product of this translation is that the target ORF encoded protein gets a C-terminal tag, and the RNA binding protein-effector fusion starts with a residual Proline (C-terminal amino acid) from TPE.

FIGs. 20A-20C show that the tethered single RNA system increases protein expression (Area under the curve) and overall duration of expression from target RNA. In this system, Balb/c mice are injected with an RNA construct containing a target (Luc) and effector/ control

separated by a TPE element (SEQ ID NOs: 94 and 96 respectively). **FIG. 20A** shows the level of luminescence over the various timepoints indicated for mice that were injected with a t-eff construct (MBP_eIF4G-mid2 (653-1130); Effector; SEQ ID NO: 23) or a t-ctrl construct (MBP-LacZ; Control; SEQ ID NO: 39). **FIG. 20B** shows the cumulative luminiscence plotted as total Area Under Curve; AUC) corresponding to **FIG. 20A**. **FIG. 20C** shows the luminescence over the time points indicated for mice that were injected with a t-eff construct (Effector) or a t-ctrl construct (Control). Data is shown as mean +/- SEM.

FIGs. 21A-21B depict increased protein output for target deg-GFP RNA in Hep3b cells with tethered effectors that are full length or truncations in eIF4G1 and eIF4G3. Total green intensity vs time is plotted. Target mRNA with v1.1_MS2 loops in the 3'UTR (SEQ ID NO: 1) was co-delivered with another mRNA encoding a control protein (LacZ) (SEQ ID NO: 98) or a tethered control (MBP-LacZ) (SEQ ID NO: 39) or different tethered effector proteins as depicted (SEQ ID NOs: 51, 55, 59, or 79). Experiments were done in 10x molar excess of target. **FIG. 21A** shows the results of a target mRNA with MS2 binding sites downstream of the v1.1 sequence in the 3' UTR. In this construct, the MS2 binding sites were adjacent to a PolyA sequence. **FIG. 21B** shows the total integrated area under the curve for the data in **FIG. 22A**.

FIGs. 22A-22B depict increased protein output for target deg-GFP RNA in Hep3b cells with tethered effectors that are different proteins that bind to (PABP) or can modulate the polyA tail (Gld2, TENT4A, TENT4B) of an RNA. Total green intensity vs time is plotted. Target mRNA with v1.1_MS2 loops in the 3'UTR (SEQ ID NO: 1) was co-delivered with another mRNA encoding a control protein (LacZ) or a tethered control (MBP-LacZ) (SEQ ID NO: 39) or different tethered effector proteins as depicted (SEQ ID NOs: 51, 55, 59, or 79). Experiments were done in 10x molar excess of target. **FIG. 22A** shows the results of a target mRNA with MS2 binding sites downstream of the v1.1 sequence in the 3' UTR. In this construct, the MS2 binding sites were adjacent to a PolyA sequence. **FIG. 22B** shows the total integrated area under the curve for the data in **FIG. 22A**.

DETAILED DESCRIPTION

Efforts to increase mRNA potency have focused on generating canonical linear mRNAs with optimal sequence design for the untranslated regions (UTRs) and open reading frame (ORFs). Disclosed herein, *inter alia*, is the discovery that RNA-binding-protein (RBP)- RNA interactions can modulate protein expression from exogenous mRNAs and/or mRNA sub-cellular localization. In some embodiments, the disclosure provides compositions and methods that can increase the efficacy, e.g., level and/or activity, of an mRNA, e.g., by making the mRNA independent of certain endogenous factors (e.g., by recruiting translation initiation factors). Additionally, disclosed herein is the discovery that effector molecules, e.g., effector proteins, can be recruited to an mRNA encoding a therapeutic payload or prophylactic payload to, e.g., promote desirable interactions between the mRNA and the effector molecule. In some embodiments, the effector molecule is recruited to the mRNA encoding the therapeutic payload or prophylactic payload by an RNA-binding protein, e.g., a tether molecule. Schematics of exemplary systems comprising these components are provided in **FIG. 1** and **FIG. 19**.

Without wishing to be bound by theory it is believed that, in some embodiments, a system or LNP composition comprising: (1) an mRNA encoding a therapeutic payload or prophylactic payload; and (2) an mRNA encoding: an RBP, e.g., a tether molecule, and/or an effector molecule, utilizes the RBP-RNA interaction and allows for increased protein expression from the mRNA encoding the therapeutic payload or prophylactic payload.

Exemplary effects on mRNA expression with systems disclosed herein are provided in Examples 2-15. Example 2 shows increased potency of a target mRNA (e.g., increased protein expression and/or duration of protein expression) when co-delivered with an RNA encoding a tethered effector protein, e.g., tethered eIF4G. Examples 3 and 4 describe the increased half-life of a target mRNA when co-delivered with a tethered effector and the effects of a tethered effector on the translation of target mRNA analysis of the domains of eIF4G which are required for effector function is provided in Example 5. Example 6 shows that a tethered effector can rescue protein expression and increase mRNA stability. Examples 7 and 8 shows that tethered effector

increases expression (and inferred mRNA stability) of a tailless RNA. Tailless target RNA with minimal-no baseline expression can be rescued (translation is restored) by effectors recruited using RBP-RNA (MBP-MS2) tethers. Examples 9-11 show that modified tailless and/or capless RNAs can be rescued by tethered effectors. Example 5 12 shows that tethered effector maintains more translating mRNAs over time. Example 13 shows that tethering increases secreted protein expression and translation. Example 14 shows that a single RNA tethering system where the target (i.e., therapeutic payload or prophylactic payload), and the tethered effector within the same RNA molecule, enhances target expression in vivo. Example 14 identifies other useful effector proteins 10 or domains thereof for use in the systems disclosed herein.

Accordingly, disclosed herein are lipid nanoparticle (LNP) compositions or systems comprising a therapeutic payload or prophylactic payload, a binding element, a tether molecule and/or an effector molecule and uses thereof. The LNP compositions or systems of the present disclosure comprise: (a) a first polynucleotide (e.g., mRNA) 15 comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide (e.g., mRNA) comprising a sequence encoding: (1) an effector molecule, and/or (2) a polypeptide that recognizes the binding element (a tether molecule). In an embodiment, the effector molecule further comprises a tether molecule. In an embodiment, the In an embodiment the 20 effector molecule polypeptide comprising a tether molecule comprises a first domain which modulates a parameter of, e.g., level and/or activity of: an RNA (e.g., an mRNA); or a protein encoded by the RNA. In an embodiment, the parameter comprises one, two, three or all of: (1) mRNA level and/or activity and/or subcellular localization (e.g., half-life and/or expression); (2) protein level and/or activity (e.g., half-life and/or 25 expression); (3) protein translation rate or (4) protein localization, e.g., location. In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a second domain which binds to, e.g., recognizes, the binding element (a tether molecule). In an embodiment the effector molecule comprising the tether molecule comprises a polypeptide comprising the first domain and the second domain. In an 30 embodiment, the first and second domains are operatively linked.

In an embodiment, the first polynucleotide and the second polynucleotide are disposed in the same or different polynucleotides. In an embodiment, a system disclosed herein is formulated as an LNP. In an embodiment, the LNP comprising the first polynucleotide is formulated as an LNP. In some embodiments, the LNP comprising the second polynucleotide is formulated as an LNP. In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are the same LNP. In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are different LNPs. In an aspect, the LNP compositions or systems of the present disclosure can: increase the level and/or activity of the therapeutic payload or prophylactic payload, e.g., increase the level and/or activity of the mRNA encoding the therapeutic payload or prophylactic payload, or increase the level and/or activity of the therapeutic payload or prophylactic payload protein. In an aspect, the LNP compositions or systems can be used in a method of treating a disease or disorder, or for modulating an immune response in a subject.

Definitions

Administering: As used herein, “administering” refers to a method of delivering a composition to a subject or patient. A method of administration may be selected to target delivery (e.g., to specifically deliver) to a specific region or system of a body. For example, an administration may be parenteral (e.g., subcutaneous, intracutaneous, intravenous, intraperitoneal, intramuscular, intraarticular, intraarterial, intrasynovial, intrasternal, intrathecal, intralesional, or intracranial injection, as well as any suitable infusion technique), oral, trans- or intra-dermal, interdermal, rectal, intravaginal, topical (e.g., by powders, ointments, creams, gels, lotions, and/or drops), mucosal, nasal, buccal, enteral, vitreal, intratumoral, sublingual, intranasal; by intratracheal instillation, bronchial instillation, and/or inhalation; as an oral spray and/or powder, nasal spray, and/or aerosol, and/or through a portal vein catheter. Preferred means of administration are intravenous or subcutaneous.

Antibody molecule: In one embodiment, antibody molecules can be used for targeting to desired cell types. As used herein, “antibody molecule” refers to a naturally occurring antibody, an engineered antibody, or a fragment thereof, e.g., an antigen

binding portion of a naturally occurring antibody or an engineered antibody. An antibody molecule can include, e.g., an antibody or an antigen-binding fragment thereof (e.g., Fab, Fab', F(ab')₂, Fv fragments, scFv antibody fragments, disulfide-linked Fvs (sdFv), a Fd fragment consisting of the VH and CH1 domains, linear antibodies, single domain antibodies such as sdAb (either VL or VH), nanobodies, or camelid VHH domains), an antigen-binding fibronectin type III (Fn3) scaffold such as a fibronectin polypeptide minibody, a ligand, a cytokine, a chemokine, or a T cell receptor (TCRs). Exemplary antibody molecules include, but are not limited to, humanized antibody molecule, intact IgA, IgG, IgE or IgM antibody; bi- or multi- specific antibody (e.g., Zybodyes®, etc); antibody fragments such as Fab fragments, Fab' fragments, F(ab')₂ fragments, Fd' fragments, Fd fragments, and isolated CDRs or sets thereof; single chain Fvs; polypeptide-Fc fusions; single domain antibodies (e.g., shark single domain antibodies such as IgNAR or fragments thereof); cameloid antibodies; masked antibodies (e.g., Probodyes®); Small Modular ImmunoPharmaceuticals ("SMIPs™"); single chain or Tandem diabodies (TandAb®); VHHs; Anticalins®; Nanobodies®; minibodies; BiTE®; ankyrin repeat proteins or DARPINs®; Avimers®; DARTs; TCR-like antibodies; Adnectins®; Affilins®; Trans-bodyes®; Affibodies®; TrimerX®; MicroProteins; Fynomers®, Centyrins®, and KALBITOR®s.

Approximately, about: As used herein, the terms "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term "approximately" or "about" refers to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value). For example, when used in the context of an amount of a given compound in a lipid component of an LNP, "about" may mean +/- 5% of the recited value. For instance, an LNP including a lipid component having about 40% of a given compound may include 30-50% of the compound.

Conjugated: As used herein, the term "conjugated," when used with respect to two or more moieties, means that the moieties are physically associated or connected

with one another, either directly or via one or more additional moieties that serves as a linking agent, to form a structure that is sufficiently stable so that the moieties remain physically associated under the conditions in which the structure is used, e.g., physiological conditions. In some embodiments, two or more moieties may be
5 conjugated by direct covalent chemical bonding. In other embodiments, two or more moieties may be conjugated by ionic bonding or hydrogen bonding.

Contacting: As used herein, the term “contacting” means establishing a physical connection between two or more entities. For example, contacting a cell with an mRNA or a lipid nanoparticle composition means that the cell and mRNA or lipid
10 nanoparticle are made to share a physical connection. Methods of contacting cells with external entities both *in vivo*, *in vitro*, and *ex vivo* are well known in the biological arts. In exemplary embodiments of the disclosure, the step of contacting a mammalian cell with a composition (e.g., a nanoparticle, or pharmaceutical composition of the disclosure) is performed *in vivo*. For example, contacting a lipid nanoparticle
15 composition and a cell (for example, a mammalian cell) which may be disposed within an organism (e.g., a mammal) may be performed by any suitable administration route (e.g., parenteral administration to the organism, including intravenous, intramuscular, intradermal, and subcutaneous administration). For a cell present *in vitro*, a composition (e.g., a lipid nanoparticle) and a cell may be contacted, for example, by
20 adding the composition to the culture medium of the cell and may involve or result in transfection. Moreover, more than one cell may be contacted by a nanoparticle composition.

Delivering: As used herein, the term “delivering” means providing an entity to a destination. For example, delivering a therapeutic and/or prophylactic to a subject may
25 involve administering a LNP including the therapeutic and/or prophylactic to the subject (e.g., by an intravenous, intramuscular, intradermal, pulmonary or subcutaneous route). Administration of a LNP to a mammal or mammalian cell may involve contacting one or more cells with the lipid nanoparticle.

Encapsulate: As used herein, the term “encapsulate” means to enclose,
30 surround, or encase. In some embodiments, a compound, polynucleotide (e.g., an mRNA), or other composition may be fully encapsulated, partially encapsulated, or

substantially encapsulated. For example, in some embodiments, an mRNA of the disclosure may be encapsulated in a lipid nanoparticle, e.g., a liposome.

Encapsulation efficiency: As used herein, “encapsulation efficiency” refers to the amount of a therapeutic and/or prophylactic that becomes part of a LNP, relative to the initial total amount of therapeutic and/or prophylactic used in the preparation of a LNP. For example, if 97 mg of therapeutic and/or prophylactic are encapsulated in a LNP out of a total 100 mg of therapeutic and/or prophylactic initially provided to the composition, the encapsulation efficiency may be given as 97%. As used herein, “encapsulation” may refer to complete, substantial, or partial enclosure, confinement, surrounding, or encasement.

Effective amount: As used herein, the term “effective amount” of an agent is that amount sufficient to effect beneficial or desired results, for example, clinical results, and, as such, an “effective amount” depends upon the context in which it is being applied. For example, in the context of the amount of a target cell delivery potentiating lipid in a lipid composition (e.g., LNP) of the disclosure, an effective amount of a target cell delivery potentiating lipid is an amount sufficient to effect a beneficial or desired result as compared to a lipid composition (e.g., LNP) lacking the target cell delivery potentiating lipid. Non-limiting examples of beneficial or desired results effected by the lipid composition (e.g., LNP) include increasing the percentage of cells transfected and/or increasing the level of expression of a protein encoded by a nucleic acid associated with/encapsulated by the lipid composition (e.g., LNP). In the context of administering a target cell delivery potentiating lipid-containing lipid nanoparticle such that an effective amount of lipid nanoparticles are taken up by target cells in a subject, an effective amount of target cell delivery potentiating lipid-containing LNP is an amount sufficient to effect a beneficial or desired result as compared to an LNP lacking the target cell delivery potentiating lipid. Non-limiting examples of beneficial or desired results in the subject include increasing the percentage of cells transfected, increasing the level of expression of a protein encoded by a nucleic acid associated with/encapsulated by the target cell delivery potentiating lipid-containing LNP and/or increasing a prophylactic or therapeutic effect *in vivo* of a nucleic acid, or its encoded protein, associated with/encapsulated by the target cell delivery potentiating lipid-

containing LNP, as compared to an LNP lacking the target cell delivery potentiating lipid. In some embodiments, a therapeutically effective amount of target cell delivery potentiating lipid-containing LNP is sufficient, when administered to a subject suffering from or susceptible to an infection, disease, disorder, and/or condition, to treat, improve
5 symptoms of, diagnose, prevent, and/or delay the onset of the infection, disease, disorder, and/or condition. In another embodiment, an effective amount of a lipid nanoparticle is sufficient to result in expression of a desired protein in at least about 5%, 10%, 15%, 20%, 25% or more of target cells. For example, an effective amount of target cell delivery potentiating lipid-containing LNP can be an amount that results in
10 transfection of at least 5%, 10%, 15%, 20%, 25%, 30%, or 35% of target cells after a single intravenous injection.

Expression: As used herein, “expression” of a nucleic acid sequence refers to one or more of the following events: (1) production of an RNA template from a DNA sequence (e.g., by transcription); (2) processing of an RNA transcript (e.g., by splicing,
15 editing, 5' cap formation, and/or 3' end processing); (3) translation of an RNA into a polypeptide or protein; and (4) post-translational modification of a polypeptide or protein.

Ex vivo: As used herein, the term “ex vivo” refers to events that occur outside of an organism (e.g., animal, plant, or microbe or cell or tissue thereof). Ex vivo events
20 may take place in an environment minimally altered from a natural (e.g., *in vivo*) environment.

Fragment: A “fragment,” as used herein, refers to a portion. For example, fragments of proteins may include polypeptides obtained by digesting full-length protein isolated from cultured cells or obtained through recombinant DNA techniques.
25 A fragment of a protein can be, for example, a portion of a protein that includes one or more functional domains such that the fragment of the protein retains the functional activity of the protein.

GC-rich: As used herein, the term “GC-rich” refers to the nucleobase composition of a polynucleotide (e.g., mRNA), or any portion thereof (e.g., an RNA
30 element), comprising guanine (G) and/or cytosine (C) nucleobases, or derivatives or analogs thereof, wherein the GC-content is greater than about 50%. The term “GC-rich”

refers to all, or to a portion, of a polynucleotide, including, but not limited to, a gene, a non-coding region, a 5' UTR, a 3' UTR, an open reading frame, an RNA element, a sequence motif, or any discrete sequence, fragment, or segment thereof which comprises about 50% GC-content. In some embodiments of the disclosure, GC-rich
5 polynucleotides, or any portions thereof, are exclusively comprised of guanine (G) and/or cytosine (C) nucleobases.

GC-content: As used herein, the term "GC-content" refers to the percentage of nucleobases in a polynucleotide (e.g., mRNA), or a portion thereof (e.g., an RNA element), that are either guanine (G) and cytosine (C) nucleobases, or derivatives or
10 analogs thereof, (from a total number of possible nucleobases, including adenine (A) and thymine (T) or uracil (U), and derivatives or analogs thereof, in DNA and in RNA). The term "GC-content" refers to all, or to a portion, of a polynucleotide, including, but not limited to, a gene, a non-coding region, a 5' or 3' UTR, an open reading frame, an RNA element, a sequence motif, or any discrete sequence, fragment, or segment
15 thereof.

Heterologous: As used herein, "heterologous" indicates that a sequence (e.g., an amino acid sequence or the polynucleotide that encodes an amino acid sequence) is not normally present in a given polypeptide or polynucleotide. For example, an amino acid sequence that corresponds to a domain or motif of one protein may be heterologous to a
20 second protein.

Isolated: As used herein, the term "isolated" refers to a substance or entity that has been separated from at least some of the components with which it was associated (whether in nature or in an experimental setting). Isolated substances may have varying levels of purity in reference to the substances from which they have been associated.
25 Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, or more of the other components with which they were initially associated. In some embodiments, isolated agents are more than about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%,
30 about 98%, about 99%, or more than about 99% pure. As used herein, a substance is "pure" if it is substantially free of other components.

Kozak Sequence: The term “Kozak sequence” (also referred to as “Kozak consensus sequence”) refers to a translation initiation enhancer element to enhance expression of a gene or open reading frame, and which in eukaryotes, is located in the 5’ UTR. The Kozak consensus sequence was originally defined as the sequence

5 GCCRCC, where R = a purine, following an analysis of the effects of single mutations surrounding the initiation codon (AUG) on translation of the preproinsulin gene (Kozak (1986) Cell 44:283-292). Polynucleotides disclosed herein comprise a Kozak consensus sequence, or a derivative or modification thereof. (Examples of translational enhancer compositions and methods of use thereof, see U.S. Pat. No. 5,807,707 to

10 Andrews et al., incorporated herein by reference in its entirety; U.S. Pat. No. 5,723,332 to Chernajovsky, incorporated herein by reference in its entirety; U.S. Pat. No. 5,891,665 to Wilson, incorporated herein by reference in its entirety.)

Leaky scanning: A phenomenon known as “leaky scanning” can occur whereby the PIC bypasses the initiation codon and instead continues scanning downstream until

15 an alternate or alternative initiation codon is recognized. Depending on the frequency of occurrence, the bypass of the initiation codon by the PIC can result in a decrease in translation efficiency. Furthermore, translation from this downstream AUG codon can occur, which will result in the production of an undesired, aberrant translation product that may not be capable of eliciting the desired therapeutic response. In some cases, the

20 aberrant translation product may in fact cause a deleterious response (Kracht et al., (2017) Nat Med 23(4):501-507).

Liposome: As used herein, by “liposome” is meant a structure including a lipid-containing membrane enclosing an aqueous interior. Liposomes may have one or more lipid membranes. Liposomes include single-layered liposomes (also known in the art as

25 unilamellar liposomes) and multi-layered liposomes (also known in the art as multilamellar liposomes).

Modified: As used herein “modified” refers to a changed state or structure of a molecule of the disclosure, e.g., a change in a composition or structure of a polynucleotide (e.g., mRNA). Molecules, e.g., polynucleotides, may be modified in

30 various ways including chemically, structurally, and/or functionally. For example, molecules, e.g., polynucleotides, may be structurally modified by the incorporation of

one or more RNA elements, wherein the RNA element comprises a sequence and/or an RNA secondary structure(s) that provides one or more functions (e.g., translational regulatory activity). Accordingly, molecules, e.g., polynucleotides, of the disclosure may be comprised of one or more modifications (e.g., may include one or more
5 chemical, structural, or functional modifications, including any combination thereof). In one embodiment, polynucleotides, e.g., mRNA molecules, of the present disclosure are modified by the introduction of non-natural nucleosides and/or nucleotides, e.g., as it relates to the natural ribonucleotides A, U, G, and C. Noncanonical nucleotides such as the cap structures are not considered “modified” although they differ from the chemical
10 structure of the A, C, G, U ribonucleotides.

mRNA: As used herein, an “mRNA” refers to a messenger ribonucleic acid. An mRNA may be naturally or non-naturally occurring. For example, an mRNA may include modified and/or non-naturally occurring components such as one or more nucleobases, nucleosides, nucleotides, or linkers. An mRNA may include a cap
15 structure, a chain terminating nucleoside, a stem loop, a polyA sequence, and/or a polyadenylation signal. An mRNA may have a nucleotide sequence encoding a polypeptide. Translation of an mRNA, for example, *in vivo* translation of an mRNA inside a mammalian cell, may produce a polypeptide. Traditionally, the basic components of an mRNA molecule include at least a coding region, a 5'-untranslated
20 region (5'-UTR), a 3'UTR, a 5' cap and a polyA sequence. In an embodiment, the mRNA is a circular mRNA.

Nanoparticle: As used herein, “nanoparticle” refers to a particle having any one structural feature on a scale of less than about 1000nm that exhibits novel properties as compared to a bulk sample of the same material. Routinely, nanoparticles have any one
25 structural feature on a scale of less than about 500 nm, less than about 200 nm, or about 100 nm. Also routinely, nanoparticles have any one structural feature on a scale of from about 50 nm to about 500 nm, from about 50 nm to about 200 nm or from about 70 to about 120 nm. In exemplary embodiments, a nanoparticle is a particle having one or more dimensions of the order of about 1 - 1000nm. In other exemplary embodiments, a
30 nanoparticle is a particle having one or more dimensions of the order of about 10- 500 nm. In other exemplary embodiments, a nanoparticle is a particle having one or more

dimensions of the order of about 50- 200 nm. A spherical nanoparticle would have a diameter, for example, of between about 50-100 or 70-120 nanometers. A nanoparticle most often behaves as a unit in terms of its transport and properties. It is noted that novel properties that differentiate nanoparticles from the corresponding bulk material typically develop at a size scale of under 1000nm, or at a size of about 100nm, but nanoparticles can be of a larger size, for example, for particles that are oblong, tubular, and the like. Although the size of most molecules would fit into the above outline, individual molecules are usually not referred to as nanoparticles.

Nucleic acid: As used herein, the term “nucleic acid” is used in its broadest sense and encompasses any compound and/or substance that includes a polymer of nucleotides. These polymers are often referred to as polynucleotides. Exemplary nucleic acids or polynucleotides of the disclosure include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), DNA-RNA hybrids, RNAi-inducing agents, RNAi agents, siRNAs, shRNAs, miRNAs, antisense RNAs, ribozymes, catalytic DNA, RNAs that induce triple helix formation, threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs, including LNA having a β -D-ribo configuration, α -LNA having an α -L-ribo configuration (a diastereomer of LNA), 2'-amino-LNA having a 2'-amino functionalization, and 2'-amino- α -LNA having a 2'-amino functionalization) or hybrids thereof.

Nucleic Acid Structure: As used herein, the term “nucleic acid structure” (used interchangeably with “polynucleotide structure”) refers to the arrangement or organization of atoms, chemical constituents, elements, motifs, and/or sequence of linked nucleotides, or derivatives or analogs thereof, that comprise a nucleic acid (e.g., an mRNA). The term also refers to the two-dimensional or three-dimensional state of a nucleic acid. Accordingly, the term “RNA structure” refers to the arrangement or organization of atoms, chemical constituents, elements, motifs, and/or sequence of linked nucleotides, or derivatives or analogs thereof, comprising an RNA molecule (e.g., an mRNA) and/or refers to a two-dimensional and/or three dimensional state of an RNA molecule. Nucleic acid structure can be further demarcated into four organizational categories referred to herein as “molecular structure”, “primary

structure”, “secondary structure”, and “tertiary structure” based on increasing organizational complexity.

Nucleobase: As used herein, the term “nucleobase” (alternatively “nucleotide base” or “nitrogenous base”) refers to a purine or pyrimidine heterocyclic compound found in nucleic acids, including any derivatives or analogs of the naturally occurring purines and pyrimidines that confer improved properties (e.g., binding affinity, nuclease resistance, chemical stability) to a nucleic acid or a portion or segment thereof.

Adenine, cytosine, guanine, thymine, and uracil are the nucleobases predominately found in natural nucleic acids. Other natural, non-natural, and/or synthetic nucleobases, as known in the art and/or described herein, can be incorporated into nucleic acids.

Nucleoside/Nucleotide: As used herein, the term “nucleoside” refers to a compound containing a sugar molecule (e.g., a ribose in RNA or a deoxyribose in DNA), or derivative or analog thereof, covalently linked to a nucleobase (e.g., a purine or pyrimidine), or a derivative or analog thereof (also referred to herein as

“nucleobase”), but lacking an internucleoside linking group (e.g., a phosphate group).

As used herein, the term “nucleotide” refers to a nucleoside covalently bonded to an internucleoside linking group (e.g., a phosphate group), or any derivative, analog, or modification thereof that confers improved chemical and/or functional properties (e.g., binding affinity, nuclease resistance, chemical stability) to a nucleic acid or a portion or segment thereof.

Open Reading Frame: As used herein, the term “open reading frame”, abbreviated as “ORF”, refers to a segment or region of an mRNA molecule that encodes a polypeptide. The ORF comprises a continuous stretch of non-overlapping, in-frame codons, beginning with the initiation codon and ending with a stop codon, and is translated by the ribosome.

Patient: As used herein, “patient” refers to a subject who may seek or be in need of treatment, requires treatment, is receiving treatment, will receive treatment, or a subject who is under care by a trained professional for a particular disease or condition. In particular embodiments, a patient is a human patient. In some embodiments, a patient is a patient suffering from an autoimmune disease, e.g., as described herein.

Pharmaceutically acceptable: The phrase “pharmaceutically acceptable” is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

Pharmaceutically acceptable excipient: The phrase “pharmaceutically acceptable excipient,” as used herein, refers any ingredient other than the compounds described herein (for example, a vehicle capable of suspending or dissolving the active compound) and having the properties of being substantially nontoxic and non-inflammatory in a patient. Excipients may include, for example: antiadherents, antioxidants, binders, coatings, compression aids, disintegrants, dyes (colors), emollients, emulsifiers, fillers (diluents), film formers or coatings, flavors, fragrances, glidants (flow enhancers), lubricants, preservatives, printing inks, sorbents, suspending or dispersing agents, sweeteners, and waters of hydration. Exemplary excipients include, but are not limited to: butylated hydroxytoluene (BHT), calcium carbonate, calcium phosphate (dibasic), calcium stearate, croscarmellose, crosslinked polyvinyl pyrrolidone, citric acid, crospovidone, cysteine, ethylcellulose, gelatin, hydroxypropyl cellulose, hydroxypropyl methylcellulose, lactose, magnesium stearate, maltitol, mannitol, methionine, methylcellulose, methyl paraben, microcrystalline cellulose, polyethylene glycol, polyvinyl pyrrolidone, povidone, pregelatinized starch, propyl paraben, retinyl palmitate, shellac, silicon dioxide, sodium carboxymethyl cellulose, sodium citrate, sodium starch glycolate, sorbitol, starch (corn), stearic acid, sucrose, talc, titanium dioxide, vitamin A, vitamin E, vitamin C, and xylitol.

Pharmaceutically acceptable salts: As used herein, “pharmaceutically acceptable salts” refers to derivatives of the disclosed compounds wherein the parent compound is modified by converting an existing acid or base moiety to its salt form (e.g., by reacting the free base group with a suitable organic acid). Examples of pharmaceutically acceptable salts include, but are not limited to, mineral or organic acid salts of basic residues such as amines; alkali or organic salts of acidic residues such as carboxylic acids; and the like. Representative acid addition salts include acetate, acetic

acid, adipate, alginate, ascorbate, aspartate, benzenesulfonate, benzene sulfonic acid, benzoate, bisulfate, borate, butyrate, camphorate, camphorsulfonate, citrate, cyclopentanepropionate, digluconate, dodecylsulfate, ethanesulfonate, fumarate, glucoheptonate, glycerophosphate, hemisulfate, heptonate, hexanoate, hydrobromide, hydrochloride, hydroiodide, 2-hydroxy-ethanesulfonate, lactobionate, lactate, laurate, lauryl sulfate, malate, maleate, malonate, methanesulfonate, 2-naphthalenesulfonate, nicotinate, nitrate, oleate, oxalate, palmitate, pamoate, pectinate, persulfate, 3-phenylpropionate, phosphate, picrate, pivalate, propionate, stearate, succinate, sulfate, tartrate, thiocyanate, toluenesulfonate, undecanoate, valerate salts, and the like.

Representative alkali or alkaline earth metal salts include sodium, lithium, potassium, calcium, magnesium, and the like, as well as nontoxic ammonium, quaternary ammonium, and amine cations, including, but not limited to ammonium, tetramethylammonium, tetraethylammonium, methylamine, dimethylamine, trimethylamine, triethylamine, ethylamine, and the like. The pharmaceutically acceptable salts of the present disclosure include the conventional non-toxic salts of the parent compound formed, for example, from non-toxic inorganic or organic acids. The pharmaceutically acceptable salts of the present disclosure can be synthesized from the parent compound which contains a basic or acidic moiety by conventional chemical methods. Generally, such salts can be prepared by reacting the free acid or base forms of these compounds with a stoichiometric amount of the appropriate base or acid in water or in an organic solvent, or in a mixture of the two; generally, nonaqueous media like ether, ethyl acetate, ethanol, isopropanol, or acetonitrile are preferred. Lists of suitable salts are found in *Remington's Pharmaceutical Sciences*, 17th ed., Mack Publishing Company, Easton, Pa., 1985, p. 1418, *Pharmaceutical Salts: Properties, Selection, and Use*, P.H. Stahl and C.G. Wermuth (eds.), Wiley-VCH, 2008, and Berge et al., *Journal of Pharmaceutical Science*, 66, 1-19 (1977), each of which is incorporated herein by reference in its entirety.

Polypeptide: As used herein, the term "polypeptide" or "polypeptide of interest" refers to a polymer of amino acid residues typically joined by peptide bonds that can be produced naturally (e.g., isolated or purified) or synthetically.

RNA: As used herein, an “RNA” refers to a ribonucleic acid that may be naturally or non-naturally occurring. For example, an RNA may include modified and/or non-naturally occurring components such as one or more nucleobases, nucleosides, nucleotides, or linkers. An RNA may include a cap structure, a chain
5 terminating nucleoside, a stem loop, a polyA sequence, and/or a polyadenylation signal. An RNA may have a nucleotide sequence encoding a polypeptide of interest. For example, an RNA may be a messenger RNA (mRNA). Translation of an mRNA encoding a particular polypeptide, for example, *in vivo* translation of an mRNA inside a mammalian cell, may produce the encoded polypeptide. RNAs may be selected from
10 the non-limiting group consisting of small interfering RNA (siRNA), asymmetrical interfering RNA (aiRNA), microRNA (miRNA), Dicer-substrate RNA (dsRNA), small hairpin RNA (shRNA), mRNA, long non-coding RNA (lncRNA) and mixtures thereof.

RNA element: As used herein, the term “RNA element” refers to a portion, fragment, or segment of an RNA molecule that provides a biological function and/or
15 has biological activity (e.g., translational regulatory activity). Modification of a polynucleotide by the incorporation of one or more RNA elements, such as those described herein, provides one or more desirable functional properties to the modified polynucleotide. RNA elements, as described herein, can be naturally-occurring, non-naturally occurring, synthetic, engineered, or any combination thereof. For example,
20 naturally-occurring RNA elements that provide a regulatory activity include elements found throughout the transcriptomes of viruses, prokaryotic and eukaryotic organisms (e.g., humans). RNA elements in particular eukaryotic mRNAs and translated viral RNAs have been shown to be involved in mediating many functions in cells. Exemplary natural RNA elements include, but are not limited to, translation initiation
25 elements (e.g., internal ribosome entry site (IRES), see Kieft et al., (2001) RNA 7(2):194-206), translation enhancer elements (e.g., the APP mRNA translation enhancer element, see Rogers et al., (1999) J Biol Chem 274(10):6421-6431), mRNA stability elements (e.g., AU-rich elements (AREs), see Garneau et al., (2007) Nat Rev Mol Cell Biol 8(2):113-126), translational repression element (see e.g., Blumer et al., (2002)
30 Mech Dev 110(1-2):97-112), protein-binding RNA elements (e.g., iron-responsive element, see Selezneva et al., (2013) J Mol Biol 425(18):3301-3310), cytoplasmic

polyadenylation elements (Villalba et al., (2011) Curr Opin Genet Dev 21(4):452-457), and catalytic RNA elements (e.g., ribozymes, see Scott et al., (2009) Biochim Biophys Acta 1789(9-10):634-641).

Specific delivery: As used herein, the term “specific delivery,” “specifically deliver,” or “specifically delivering” means delivery of more (e.g., at least 10% more, at least 20% more, at least 30% more, at least 40% more, at least 50% more, at least 1.5 fold more, at least 2-fold more, at least 3-fold more, at least 4-fold more, at least 5-fold more, at least 6-fold more, at least 7-fold more, at least 8-fold more, at least 9-fold more, at least 10-fold more) of a therapeutic and/or prophylactic by a nanoparticle to a target cell of interest (e.g., mammalian target cell) compared to an off-target cell (e.g., non-target cells). The level of delivery of a nanoparticle to a particular cell may be measured by comparing the amount of protein produced in target cells versus non-target cells (e.g., by mean fluorescence intensity using flow cytometry, comparing the % of target cells versus non-target cells expressing the protein (e.g., by quantitative flow cytometry), comparing the amount of protein produced in a target cell versus non-target cell to the amount of total protein in said target cells versus non-target cell, or comparing the amount of therapeutic and/or prophylactic in a target cell versus non-target cell to the amount of total therapeutic and/or prophylactic in said target cell versus non-target cell. It will be understood that the ability of a nanoparticle to specifically deliver to a target cell need not be determined in a subject being treated, it may be determined in a surrogate such as an animal model (e.g., a mouse or NHP model).

Substantially: As used herein, the term “substantially” refers to the qualitative condition of exhibiting total or near-total extent or degree of a characteristic or property of interest. One of ordinary skill in the biological arts will understand that biological and chemical phenomena rarely, if ever, go to completion and/or proceed to completeness or achieve or avoid an absolute result. The term “substantially” is therefore used herein to capture the potential lack of completeness inherent in many biological and chemical phenomena.

Suffering from: An individual who is “suffering from” a disease, disorder, and/or condition has been diagnosed with or displays one or more symptoms of a disease, disorder, and/or condition.

5 *Targeting moiety:* As used herein, a “targeting moiety” is a compound or agent that may target a nanoparticle to a particular cell, tissue, and/or organ type.

10 *Effector Molecule:* As used herein, the term “effector molecule” refers to a molecule that can modulate a parameter of, e.g., level and/or activity of: an RNA (e.g., an mRNA); or a protein encoded by the RNA. In an embodiment, the parameter comprises one, two or all of: (1) mRNA level and/or activity and/or subcellular
 15 localization (e.g., half-life and/or expression); (2) protein level and/or activity (e.g., half-life and/or expression); (3) protein translation rate or (4) protein localization, e.g., location. In an embodiment, an effector molecule comprises: a translation factor, a splicing factor, an RNA stabilizing factor, an RNA editing factor, an RNA-binding factor, an RNA localizing factor, or any combination thereof, e.g., as provided in Table
 20 4. An effector molecule comprises wildtype (e.g., naturally occurring, e.g., human), full length, a fragment (e.g., biologically active or functional fragment), or a variant of any of the aforementioned classes of effector molecules. In an embodiment, the effector molecule further comprises a tether molecule. In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a first domain which modulates a
 25 parameter of, e.g., level and/or activity of: an RNA (e.g., an mRNA); or a protein encoded by the RNA, e.g., as described herein. In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a second domain which binds to, e.g., recognizes, the binding element (a tether molecule). In an embodiment, the effector molecule comprises an RNA-binding protein or a fragment thereof. In an embodiment,
 30 an effector molecule comprises a translation factor, e.g., eIF4G. In an embodiment, an effector molecule comprises wildtype (e.g., naturally occurring, e.g., human), full length, a fragment (e.g., biologically active or functional fragment), or a variant of eIF4G. Exemplary eIF4G constructs are provided herein, e.g., in Table 2.

30 *Binding element:* As used herein, the term “binding element” refers to a nucleic acid sequence, e.g., a DNA or RNA sequence, which is recognized by a tether molecule. In an embodiment, the binding element forms a structure, e.g., a three-dimensional

structure, e.g., a kink-turn, a loop, a stem or other known structure. Exemplary binding elements include, but are not limited, to those provided in Table 1.

Tether Molecule: As used herein, the term “tether molecule” refers to a molecule which binds to, e.g., recognizes, a binding element or a fragment thereof. In an embodiment, the tether molecule binds to, e.g., recognizes, a sequence, e.g., a DNA or RNA sequence, comprising the binding element, or fragment thereof. In an embodiment, the tether molecule binds to, e.g., recognizes, a structure comprising a sequence, e.g., a DNA or RNA sequence, comprising the binding element, or fragment thereof. In an embodiment, the effector molecule comprises an RNA-binding protein or a fragment thereof. Exemplary tether molecules are provided in Table 1.

Therapeutic Agent: The term “therapeutic agent” refers to any agent that, when administered to a subject, has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect. In some embodiments, the therapeutic agent comprises or is a therapeutic payload. In some embodiments, the therapeutic agent comprises or is a small molecule or a biologic (e.g., an antibody molecule).

Therapeutic payload or prophylactic payload: As used herein, the term “therapeutic payload or prophylactic payload” refers to an agent which elicits a desired biological and/or pharmacological effect. In an embodiment, the therapeutic payload or prophylactic payload has a therapeutic and/or prophylactic effect. In an embodiment, the therapeutic payload or prophylactic payload comprises a protein, a polypeptide, a peptide or a fragment (e.g., a biologically active fragment) thereof. In an embodiment, the therapeutic payload or prophylactic payload includes a sequence encoding a protein, e.g., a therapeutic protein. Some examples of therapeutic payload or prophylactic payloads may include, but are not limited to a secreted protein, a membrane-bound protein, or an intracellular protein. In an embodiment, the therapeutic payload or prophylactic payload includes a cytokine, an antibody, a vaccine (e.g., an antigen, or an immunogenic epitope), a receptor, an enzyme, a hormone, a transcription factor, a ligand, a membrane transporter, a structural protein, a nuclease, or a component, a variant or a fragment (e.g., a biologically active fragment) thereof. The terms protein, polypeptide and peptide are used interchangeably herein.

Transfection: As used herein, the term “transfection” refers to methods to introduce a species (e.g., a polynucleotide, such as a mRNA) into a cell.

Translational Regulatory Activity: As used herein, the term “translational regulatory activity” (used interchangeably with “translational regulatory function”) refers to a biological function, mechanism, or process that modulates (e.g., regulates, influences, controls, varies) the activity of the translational apparatus, including the activity of the PIC and/or ribosome. In some aspects, the desired translation regulatory activity promotes and/or enhances the translational fidelity of mRNA translation. In some aspects, the desired translational regulatory activity reduces and/or inhibits leaky scanning.

Subject: As used herein, the term “subject” refers to any organism to which a composition in accordance with the disclosure may be administered, e.g., for experimental, diagnostic, prophylactic, and/or therapeutic purposes. Typical subjects include animals (e.g., mammals such as mice, rats, rabbits, non-human primates, and humans) and/or plants. In some embodiments, a subject may be a patient.

Treating: As used herein, the term “treating” refers to partially or completely alleviating, ameliorating, improving, relieving, delaying onset of, inhibiting progression of, reducing severity of, and/or reducing incidence of one or more symptoms or features of a particular infection, disease, disorder, and/or condition. For example, “treating” cancer may refer to inhibiting survival, growth, and/or spread of a tumor. Treatment may be administered to a subject who does not exhibit signs of a disease, disorder, and/or condition and/or to a subject who exhibits only early signs of a disease, disorder, and/or condition for the purpose of decreasing the risk of developing pathology associated with the disease, disorder, and/or condition.

Preventing: As used herein, the term “preventing” refers to partially or completely inhibiting the onset of one or more symptoms or features of a particular infection, disease, disorder, and/or condition.

Unmodified: As used herein, “unmodified” refers to any substance, compound or molecule prior to being changed in any way. Unmodified may, but does not always, refer to the wild type or native form of a biomolecule. Molecules may undergo a series

of modifications whereby each modified molecule may serve as the “unmodified” starting molecule for a subsequent modification.

Uridine Content: The terms "uridine content" or "uracil content" are interchangeable and refer to the amount of uracil or uridine present in a certain nucleic acid sequence. Uridine content or uracil content can be expressed as an absolute value (total number of uridine or uracil in the sequence) or relative (uridine or uracil percentage respect to the total number of nucleobases in the nucleic acid sequence).

Uridine-Modified Sequence: The terms "uridine-modified sequence" refers to a sequence optimized nucleic acid (e.g., a synthetic mRNA sequence) with a different overall or local uridine content (higher or lower uridine content) or with different uridine patterns (e.g., gradient distribution or clustering) with respect to the uridine content and/or uridine patterns of a candidate nucleic acid sequence. In the content of the present disclosure, the terms "uridine-modified sequence" and "uracil-modified sequence" are considered equivalent and interchangeable.

A "high uridine codon" is defined as a codon comprising two or three uridines, a "low uridine codon" is defined as a codon comprising one uridine, and a "no uridine codon" is a codon without any uridines. In some embodiments, a uridine-modified sequence comprises substitutions of high uridine codons with low uridine codons, substitutions of high uridine codons with no uridine codons, substitutions of low uridine codons with high uridine codons, substitutions of low uridine codons with no uridine codons, substitution of no uridine codons with low uridine codons, substitutions of no uridine codons with high uridine codons, and combinations thereof. In some embodiments, a high uridine codon can be replaced with another high uridine codon. In some embodiments, a low uridine codon can be replaced with another low uridine codon. In some embodiments, a no uridine codon can be replaced with another no uridine codon. A uridine-modified sequence can be uridine enriched or uridine rarefied.

Uridine Enriched: As used herein, the terms "uridine enriched" and grammatical variants refer to the increase in uridine content (expressed in absolute value or as a percentage value) in a sequence optimized nucleic acid (e.g., a synthetic mRNA sequence) with respect to the uridine content of the corresponding candidate nucleic acid sequence. Uridine enrichment can be implemented by substituting codons in the

candidate nucleic acid sequence with synonymous codons containing less uridine nucleobases. Uridine enrichment can be global (i.e., relative to the entire length of a candidate nucleic acid sequence) or local (i.e., relative to a subsequence or region of a candidate nucleic acid sequence).

5 *Uridine Rarefied*: As used herein, the terms "uridine rarefied" and grammatical variants refer to a decrease in uridine content (expressed in absolute value or as a percentage value) in an sequence optimized nucleic acid (e.g., a synthetic mRNA sequence) with respect to the uridine content of the corresponding candidate nucleic acid sequence. Uridine rarefication can be implemented by substituting codons in the candidate nucleic acid sequence with synonymous codons containing less uridine nucleobases. Uridine rarefication can be global (i.e., relative to the entire length of a candidate nucleic acid sequence) or local (i.e., relative to a subsequence or region of a candidate nucleic acid sequence).

15 *Variant*: As used herein, the term "variant" refers to a molecule having at least 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% activity of the wild type molecule, e.g., as measured by an art-recognized assay.

Delivery of therapeutic payload or prophylactic payload with a tethered effector

Disclosed herein are, *inter alia*, LNP compositions or systems comprising a therapeutic payload or prophylactic payload, a binding element, a tether molecule and/or an effector molecule and uses thereof. In an embodiment, the LNP compositions of the present disclosure comprise: (a) a first polynucleotide (e.g., mRNA) comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide (e.g., mRNA) comprising a sequence encoding: (1) an effector molecule, and/or (2) a polypeptide that recognizes the binding element (a tether molecule). In an embodiment, the effector molecule further comprises a tether molecule.

In an aspect, the LNP compositions or systems of the present disclosure can: increase the level and/or activity of the therapeutic payload or prophylactic payload, e.g., increase the level and/or activity of the mRNA encoding the therapeutic payload or prophylactic payload, increase the stability of the mRNA encoding the therapeutic payload or prophylactic payload, or increase the level and/or activity of the therapeutic

payload or prophylactic payload protein. In an aspect, the LNP compositions of the present disclosure are contacted with cells, e.g., ex vivo or in vivo and can be used for treating a disease or disorder, for modulating an immune response in a subject, or to deliver a secreted polypeptide, an intracellular polypeptide or a transmembrane polypeptide to a subject.

In an embodiment, the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide. In an embodiment, the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

In an aspect, the system disclosed herein is formulated as an LNP. In an embodiment, a system disclosed herein comprises (1) a polynucleotide, e.g., a first polynucleotide, encoding a therapeutic payload or prophylactic payload, e.g., as described herein; and/or (2) a polynucleotide, e.g., a second polynucleotide, encoding a tether molecule and an effector molecule. In an embodiment, the system comprising the first and/or second polynucleotides is formulated as an LNP.

In an aspect, the system disclosed herein is formulated as an LNP. In an embodiment, a system disclosed herein comprises (1) a polynucleotide, e.g., a first polynucleotide, encoding a therapeutic payload or prophylactic payload, e.g., as described herein; and/or (2) a polynucleotide, e.g., a second polynucleotide, encoding an effector molecule. In an embodiment, the effector molecule further comprises a tether molecule.

In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are the same, e.g., the same LNP comprises the first polynucleotide and the second polynucleotide.

In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are different, e.g., a first LNP comprises the first polynucleotide and a second LNP comprises the second polynucleotide.

In an embodiment, the LNP comprising the first polynucleotide is in a composition. In an embodiment, the LNP comprising the second polynucleotide is in a separate composition.

In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are in the same composition.

In aspects where a single RNA system is used (e.g., the system of **FIG. 19**), the first and second polynucleotides are in the same RNA molecule and can be separated by a protease cleavage site (e.g., a P2A, or T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site. The TPE region encodes for self cleaving peptides: The cleavage of the single RNA is triggered by ribosomal skipping of the peptide bond between the Proline (P) and Glycine (G) in C-terminal of 2A peptide, resulting in the peptide located upstream of the 2A peptide (i.e., the target peptide) to have extra amino acids on its C-terminal end while the peptide located downstream the 2A peptide (i.e., the tethered effector) will have an extra Proline on its N-terminal end. The molecular mechanism of 2A-peptide-mediated cleavage is believed to involve ribosomal "skipping" of glycyl-prolyl peptide bond formation rather than true proteolytic cleavage. Liu et al, 2017 Scientific Reports 7: 2193.

The 2A self-cleaving peptides that can be used in the present disclosure include, but are not limited to the following:

T2A (GSG)EGRGSLTTCGDVEENPGP (SEQ ID NO: 89);
 P2A (GSG)ATNFSLKQAGDVEENPGP (SEQ ID NO: 90);
 E2A (GSG)QCTNYALLKLAGDVESNPGP (SEQ ID NO: 91); and
 F2A (GSG)VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 92); or a combination of the above, such as TPE.

In an aspect, an LNP composition comprising a polynucleotide, e.g., a first polynucleotide encoding a therapeutic payload or prophylactic payload, and/or a polynucleotide, e.g., a second polynucleotide encoding a tether molecule and/or an effector molecule, comprises: (i) an ionizable lipid, e.g., an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-lipid.

In another aspect, the LNP compositions of the disclosure are used in a method of treating a disease or disorder, or in a method of inhibiting an immune response in a subject.

In an aspect, an LNP composition or system comprising (1) a polynucleotide, e.g., a first polynucleotide, encoding a therapeutic payload or prophylactic payload, e.g., as described herein; and/or (2) a polynucleotide, e.g., a second polynucleotide, encoding a tether molecule and an effector molecule, can be administered with an additional agent, e.g., as described herein.

Therapeutic payload or prophylactic payload

Disclosed herein, *inter alia*, is a system or LNP comprising a polynucleotide encoding a therapeutic payload or a prophylactic payload. In some embodiments, the therapeutic payload or prophylactic payload comprises an mRNA encoding: a secreted protein; a membrane-bound protein; or an intercellular protein, or peptides, polypeptides or biologically active fragments thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises an mRNA encoding a secreted protein, or a peptide, a polypeptide or a biologically active fragment thereof. In some embodiments, the secreted protein comprises a cytokine, or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the secreted protein comprises an antibody or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the secreted protein comprises an enzyme or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the secreted protein comprises a hormone or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the secreted protein comprises a ligand, or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the secreted protein comprises a vaccine (e.g., an antigen, an immunogenic epitope), or a component, variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the vaccine is a prophylactic vaccine. In some embodiments, the vaccine is a therapeutic vaccine, e.g., a cancer vaccine. In some embodiments, the secreted protein comprises a growth factor or a component, variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the secreted protein comprises an immune modulator, e.g., an immune checkpoint agonist or antagonist.

In some embodiments, the therapeutic payload or prophylactic payload comprises an mRNA encoding a membrane-bound protein, or a peptide, a polypeptide or a biologically active fragment thereof. In some embodiments, the membrane-bound protein comprises a vaccine (e.g., an antigen, an immunogenic epitope), or a component, variant or fragment (e.g., a biologically active fragment) thereof. In some 5 embodiments, the vaccine is a prophylactic vaccine. In some embodiments, the vaccine is a therapeutic vaccine, e.g., a cancer vaccine. In some embodiments, the membrane-bound protein comprises a ligand, a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the membrane-bound protein comprises a membrane transporter, a variant or fragment (e.g., a biologically active fragment) 10 thereof. In some embodiments, the membrane-bound protein comprises a structural protein, a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the membrane-bound protein comprises an immune modulator, e.g., an immune checkpoint agonist or antagonist.

15 In some embodiments, the therapeutic payload or prophylactic payload comprises an mRNA encoding an intracellular protein, or a peptide, a polypeptide or a biologically active fragment thereof. In some embodiments, the intracellular protein comprises an enzyme, or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the intracellular protein comprises a hormone, or a 20 variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the intracellular protein comprises a cytokine, or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the intracellular protein comprises a transcription factor, or a variant or fragment (e.g., a biologically active fragment) thereof. In some 25 embodiments, the intracellular protein comprises a nuclease, or a variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the intracellular protein comprises a vaccine (e.g., an antigen, an immunogenic epitope), or a component, variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the vaccine is a prophylactic vaccine. In some embodiments, the vaccine is a therapeutic vaccine, e.g., a cancer vaccine. In 30 some embodiments, the intracellular protein comprises a structural protein, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload is chosen from a cytokine, an antibody, a vaccine (e.g., an antigen, an immunogenic epitope), a receptor, an enzyme, a hormone, a transcription factor, a ligand, a membrane transporter, a structural protein, a nuclease, a growth factor, an immune modulator, or a component, variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a cytokine, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises an antibody or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a vaccine (e.g., an antigen, an immunogenic epitope), or a component, variant or fragment (e.g., a biologically active fragment) thereof. In some embodiments, the vaccine is a prophylactic vaccine. In some embodiments, the vaccine is a therapeutic vaccine, e.g., a cancer vaccine.

In some embodiments, the therapeutic payload or prophylactic payload comprises a receptor, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises an enzyme, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a hormone, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a growth factor, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a nuclease, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a transcription factor, or a variant or fragment (e.g., a biologically active fragment) thereof.

5 In some embodiments, the therapeutic payload or prophylactic payload comprises a ligand, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises a membrane transporter, or a variant or fragment (e.g., a biologically active fragment) thereof.

10 In some embodiments, the therapeutic payload or prophylactic payload comprises a structural protein, or a variant or fragment (e.g., a biologically active fragment) thereof.

In some embodiments, the therapeutic payload or prophylactic payload comprises an immune modulator, or a variant or fragment (e.g., a biologically active
15 fragment) thereof. In some embodiments, the immune modulator comprises an immune checkpoint agonist or antagonist.

In some embodiments, the therapeutic payload or prophylactic payload comprises a protein or peptide.

In some embodiments, the first polynucleotide that comprises a sequence
20 encoding a therapeutic payload or a prophylactic payload and a binding element is inherently unstable, self-degrading and/or dormant due to the presence of an inactivating/destabilizing sequence or a degradation tag in the first polynucleotide. The first polynucleotide is subject to stabilization and/or protein expression when co-delivered with a second polynucleotide encoding a tethered effector that binds to the
25 binding element.

In some embodiments, the first polynucleotide that comprises a sequence encoding a therapeutic payload or a prophylactic payload and a binding element does not have a polyA tail and is therefore inherently unstable and/or unable to translate. The first polynucleotide is subject to stabilization when co-delivered with a second
30 polynucleotide encoding a tethered effector that binds to the binding element.

Binding element

Disclosed herein, *inter alia*, is a system or LNP comprising a polynucleotide, e.g., a first polynucleotide, comprising a binding element. A binding element comprises a sequence, e.g., a DNA or RNA sequence, which is bound, e.g., recognized by, an
5 RNA binding protein or a fragment thereof, e.g., a tether molecule, e.g., as disclosed herein. In some embodiments, the tether molecule binds to a sequence comprising the binding element, or a fragment thereof. In some embodiments, the tether molecule binds to a structure comprising the binding element, or a fragment thereof.

In some embodiments, the system or LNP comprises a second polynucleotide
10 encoding an RNA binding protein or a fragment thereof, e.g., a tether molecule, which binds to, e.g., recognizes, the binding element of the first polynucleotide.

In some embodiments, the binding element of the first polynucleotide is situated upstream (5') or downstream (3'), or in the open reading frame of the sequence encoding the therapeutic payload or prophylactic payload.

15 In some embodiments, the binding element of the first polynucleotide is situated upstream (5') or downstream (3') of a 5' UTR of the first polynucleotide. In some embodiments, the binding element of the first polynucleotide is situated upstream (5') or downstream (3') of a 3' UTR of the first polynucleotide. In some embodiments, the binding element of the first polynucleotide is situated downstream of a 3' UTR of the
20 first polynucleotide. In some embodiments, the binding element of the first polynucleotide is situated adjacent, e.g., next to, a Poly A tail.

In some embodiments, the binding element of the first polynucleotide is bound by the tether molecule of the second polynucleotide, e.g., an effector molecule further comprising a tether molecule.

25 In some embodiments, a tether molecule is chosen from a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, 15.5kd, or LARP7 or a variant or fragment thereof. In some embodiments, the binding element comprises a sequence which is bound, e.g., recognized, by the tether molecule. In some embodiments, the binding element comprises a sequence comprising a structure that is
30 bound, e.g., recognized, by the tether molecule.

In some embodiments, the binding element is chosen from a binding element provided in Table 1, e.g., MS2, PP7, BoxB, U1A hairpin, PRE, a kink-turn forming sequence, 7sk, or a variant or fragment thereof. In some embodiments, the binding element is MS2. In some embodiments, the binding element is PP7. In some
5 embodiments, the binding element is BoxB. In some embodiments, the binding element is U1A hairpin. In some embodiments, the binding element is PRE. In some embodiments, the binding element is a kink-turn forming sequence. In some embodiments, the binding element is 7SK.

In some embodiments, when the binding element is MS2 (e.g., wildtype MS2, or
10 a variant or fragment thereof) the tether molecule is MBP (e.g., wildtype MBP, a variant or fragment thereof).

In some embodiments, when the binding element is PP7 (e.g., wildtype PP7, or a variant or fragment thereof) the tether molecule is PCP (e.g., wildtype PCP, or a variant or fragment thereof).

15 In some embodiments, when the binding element is BoxB (e.g., wildtype BoxB, or a variant or fragment thereof) the tether molecule is Lambda N (e.g., wildtype Lambda N, or a variant or fragment thereof).

In some embodiments, when the binding element is U1A hairpin (e.g., wildtype U1A hairpin, or a variant or fragment thereof) the tether molecule is U1A (e.g.,
20 wildtype U1A, or a variant or fragment thereof).

In some embodiments, when the binding element is PRE (e.g., wildtype PRE, or a variant or fragment thereof) the tether molecule is PUF (e.g., wildtype PUF, or a variant or fragment thereof).

In some embodiments, when the binding element is a kink-turn forming
25 sequence the tether molecule is 15.5kd (e.g., wildtype 15.5kd, or a variant or fragment thereof).

In some embodiments, when the binding element is a 7sk sequence the tether molecule is LARP7 (e.g., wildtype LARP7, or a variant or fragment thereof).

Table 1: Exemplary binding elements and tether molecules

RNA binding protein (RBP)	RNA element	Recognition basis
PUF	PRE	Sequence
MBP	MS2	Structure
PCP	PP7	Structure
lambda N	boxB	Structure
U1A	U1A hairpin	Structure
15.5kd	Kink-turn forming sequence	Structure
LARP7	7sk	Structure

In some embodiments, the binding element comprises a sequence comprising 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90 or 100 nucleotides. In some embodiments, the binding element comprises a sequence comprising about 5-100, about 5-90, about 5-80, about 5-70, about 5-60, about 5-50, about 5-40, about 5-30, about 5-25, about 5-20, about 5-19, about 5-18, about 5-17, about 5-16, about 5-15, about 5-14, about 5-13, about 5-12, about 5-11, about 5-10, about 5-9, about 5-8, about 5-7 or about 5-6 nucleotides. In some embodiments, the binding element comprises a sequence comprising about 5-100, about 6-100, about 7-100, about 8-100, about 9-100, about 10-100, about 11-100, about 12-100, about 13-100, about 14-100, about 15-100, about 16-100, about 17-100, about 18-100, about 19-100, about 20-100, about 21-100, about 22-100, about 23-100, about 24-100, about 25-100, about 30-100, about 40-100, about 50-100, about 60-100, about 70-100, about 80-100, or about 90-100 nucleotides. In some embodiments, the binding element comprises a sequence comprising about 5-100, about 6-90, about 7-80, about 8-70, about 9-60, about 10-50, about 11-40, about 12-30, about 13-25, about 14-24, about 15-23, about 16-22, about 17-21, or about 18-20 nucleotides. In some embodiments, the binding element comprises a sequence comprising 19 nucleotides.

In some embodiments, the binding element comprises a binding element nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%,

95%, 96%, 97%, 98%, 99% or 100% identity thereof. In some embodiments, the binding element comprises a binding element sequence provided in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 154, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

5 In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, the binding element comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20 or 30 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises no more than 80, 70, 60, 50, 40 or 30 repeats of the sequence bound by the tether molecule of the
10 second polynucleotide. In some embodiments, the binding element comprises about 1-30, about 1-20, about 1-10, about 1-9, about 1-8, about 1-7, about 1-6, about 1-5, about 1-4, about 1-3, or about 1-2 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises about 1-30, about 2-30, about 3-30, about 4-30 about, 5-30 about, 6-30, about 7-30, about 8-30,
15 about 9-30, about 10-30, about 11-30, about 12-30, about 13-30, about 14-30, about 15-30, or about 20-30 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some embodiments, the binding element comprises about 1-30, about 2-20, about 3-15, about 4-14, about 5-13, about 6-12, about 7-11, or about 8-10 repeats of the sequence bound by the tether molecule of the second polynucleotide. In some
20 embodiments, the binding element comprises 6 repeats of the sequence bound by the tether molecule of the second polynucleotide.

In some embodiments of any of the systems, LNP compositions, methods or uses disclosed herein, each repeat is separated by a spacer sequence comprising 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50,
25 60, 70, 80, 90 or 100 nucleotides. In some embodiments, the spacer sequence comprises about 1-100, about 1-90, about 1-80, about 1-70, about 1-60, about 1-50, about 1-40, about 1-30, about 1-25, about 1-20, about 1-19, about 1-18, about 1-17, about 1-16, about 1-15, about 1-14, about 1-13, about 1-12, about 1-11, about 1-10, about 1-9, about 1-8, about 1-7, about 1-6, about 1-5, about 1-4, about 1-3, or about 1-2 nucleotides. In
30 some embodiments, the spacer sequence comprises about 1-100, about 2-100, about 3-100, about 4-100, about 5-100, about 6-100, about 7-100, about 8-100, about 9-100,

about 10-100, about 11-100, about 12-100, about 13-100, about 14-100, about 15-100, about 16-100, about 17-100, about 18-100, about 19-100, about 20-100, about 21-100, about 22-100, about 23-100, about 24-100, about 25-100, about 30-100, about 40-100, about 50-100, about 60-100, about 70-100, about 80-100, or about 90-100 nucleotides.

- 5 In some embodiments, the spacer sequence comprises about 1-100, about 2-90, about 3-80, about 4-70, about 5-60, about 6-50, about 7-40, about 8-40, about 9-30, about 10-25, about 11-24, about 12-23, about 13-22, about 14-21, about 15-20, about 16-19, about 17-18 nucleotides. In some embodiment, the spacer sequence comprises 20 nucleotides.

- In some embodiment, the spacer sequence comprises a spacer sequence
 10 provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

Table 2: Exemplary sequences of a binding element, a tether molecule, and/or an effector molecule.

SEQ ID NO	Sequence information	Sequence
3'UTR sequences		
1	3'v1.1_6xMS2 (MS2 sequences in bold and underline)	TGATAATAGGCTGGAGCCTCGGTGGCCTAGCTTCTTGCCCCTTGGCCTCCCCCAGCCCCCTCCTCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC <u>CAGGTATGATTACAACCTGT</u> GCTACTTTAGATGACTATT <u>GAAAGACCATTAGGCTTT</u> CTAACGGACGCTCACGAGAGAG <u>GGGGCACCAATAGGGCCCCACGGCACTTCAAAG</u> GTTTGG <u>CTTGGAGTAGTAACCCAAG</u> CAGCAACAGTTTGGACTTTTCGACCACCATCAGGGGTCCCACGTTGGGAACACGTAAC <u>CTCCTACTAACAAGAGGAG</u>
2	3'6xMS2_3'v1.1 (MS2 sequences in bold and underline)	TGATAATAG <u>CAGGTATGATTACAACCTGT</u> CTACTTTAGATGACTATT <u>GAAAGACCATTAGGCTTT</u> CTAACGGACGCTCACGAGAGAG <u>GGGCACCAATAGGGCCCCACGGCACTTCAAAGGTTTGGCTTGGAGTAGTAACCCAAG</u> CAGCAACAGTTTGGACTTT <u>CGGACCACCATCAGGGGTCCCACGTTGGGAACACGTAAC</u> <u>CTCCTACTAACAAGAGGAG</u> GCTGGAGCCTCGGTGGCCTAGCTTCTTGCCCCTTGGGCCTCCCCCA GCCCTCCTCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC
3	3'6xMS2 (MS2 sequences in bold and underline)	TGATAATAG <u>CAGGTATGATTACAACCTGT</u> CTACTTTAGATGACTATT <u>GAAAGACCATTAGGCTTT</u> CTAACGGACGCTCACGAGAGAG <u>GGGCACCAATAGGGCCCCACGGCACTTCAAAGGTTTGGCTTGGAGTAGTAACCCAAG</u> CAGCAACAGTTTGGACTTT <u>CGGACCACCATCAGGGGTCCCACGTTGGGAACACGTAAC</u> <u>CTCCTACTAACAAGAGGAG</u>
4	3'v1.1	TGATAATAGGCTGGAGCCTCGGTGGCCTAGCTTCTTGCCCCTTGGCCTCCCCCAGCCCCCTCCTCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC
154	3'UTR_24MS2	TGATAATAGGCGCGATGATCACACGCGCACGCGAAGCTTGTGTT CGCACGGCAAGCAGCAGCTGCCGTGGAGGCTTCCACGGCGAAATGT

SEQ ID NO	Sequence information	Sequence
		CCACGAATACGGGACAAATAATTGATATAGGGCCGGCCCGTAAGAG TACTACGGGATGATTCTCTTGACTTAAGTGCCGGACCACTAGGCCGG CAACGGACAGACAAAGAGGCTCTGGGACTACCAAGCCCAGTCACAC GGTACTTCAAGGCACCGGGAACAGTAGTCCCGGAGTGTTGACTTATA GAAGTAGTCCCATTATTAAAGGGACTTCGGGACTTCCCTTCTCGCTC CCGACCAGTAGGCGGGAAATGCGCCCTCCCGTACTGTCCTGCACCAT TAGGGCAGGTTCTTTGAGGCTCTCTCGCTCGTGTATGAATACAACAC GTGTGCGTCTGTCCCTCGTGACTCAGGACTACCGAGTCACAGCAA ACTTAAAGGATCGCACGGACGAACACGCCGTGACATAGCTAATCAG ACAATGGCACCAGTACCAACGGTGCACGTTGACACTCGAACGTGAT TGACAGGATTACCGTCAACTCTTCCACTAGCCCGTTACTGGCAAGGA TTACCTGCCACTGTTTTGTCTTAACCGGTTGACGTATCAACAGAACG TCTGGTAGCAGCCACGCTTACTGGGCCATTAATAAAGGCCCAAGGA ACTTAAAGCCAGTGCCCGGCACTATCAAGGCCGGGCATCATTGATTC GACGCCGGACGGAGGAGCACCCCGTCGGGCATCGCGGCCCTGCGGG TGCTAAACAGTAGTTAGCAGACATTTAGCGCCGGTTTGAAGCCCACG AGTACGGGGCTTAGCAGCTCCAGATACACGGCGGGAACCTACTAAGT CCCGTAAACTCTCCACTATCGCCGGGCCAATCAATAGATGGCC
5' UTR sequence		
5	5' v1.1	GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAC CCCGGCGCCGCCACC
Effector molecule or tether molecule sequences		
6	MBP (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITPIFATNSDCELVK AMQGLLKDGNIPIAIAANSIY
7	MBP (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTAC
8	eIF4GdN (aa)	IFASMQKPEGLPHISDVVLDKANKTPLRPLDPTLRQGINC GPDFTPSFAN LGRITLSTRGPPRGGPGGELPRGPAGLPRRSQQGPKEPRKIIATVLMT EDIKLNKA EKAWKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSILNK LTPQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMC RCLMALKVPTTEKPTVTVNFRKLLLNRCKEFEKDKDDDEVFEKKQKE MDEAATAEERGRLEELEEARDIARRSLGNIKFIGELFKLKM LTEAM HDCVVKLLKNHDEESLECLCRLTTIGKDLDFEKAKPRMDQYFNQMEK IIEKKTSSRIRFMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAEMEE HREHIKVQQLMAKGSDKRRGGPPGPISRG LPLVDDGGWNTVPISKGSR PIDTSRLTKITKPGSIDSNNQLFAPGGRLSWGKGSSGGSGAKPSDAASEA ARPATSTLNRFSALQQAVPTESDNRVVRSSLSRERGEKAGDRGDRL ERSERGGDRGDRLDRARTPATKRSFSKEVEERSRERPSQPEGLRKAASL TEDRDRGRDAVKREAALPPVSPLKAALSEELEKKSAMIEEYLHLNDM KEAVQCVQELASPSLLFIFVRHGVSTLERSAIAREHMQQLLHQLLCAG HLSTAQYYQGLYEILELAEDMEIDIPHVWLYLAELVTPILQEGGVPMGE LFREITKPLRPLGKAASLLLEILGLLCKSMGPKKVGTLWREAGLSWKEF LPEGQDIGAFVAEQKVEYTLGEESEAPGQRALPSEELNRQLEKLLKEGS SNQRVFDWIEANLSEQQIVSNTLVRALMTAVCYSAIIFETPLRVDVAVL

SEQ ID NO	Sequence information	Sequence
		KARAKLLQKYLCEQKELQALYALQALVVTLEQPPNLLRMFFDALYD EDVVKEDAFYSWESSKDPAEQQGKGVALKSVTAFFKWLREAEEDSDH N
9	eIF4GdN (nt)	ATATTCGCGAGCATGCAGAAGCCTGAAGGTCTGCCTCACATCAG CGACGTGGTGTCTGGACAAGGCCAACAAGACCCCTCTTAGACCTCTG GACCCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCC TTCTTTCGCAAACCTGGGCAGAACACCCTGAGCACCAGAGGCCCTC CTAGAGGTGGTCCCGGGCGGAGAACTCCCCAGGGGTCTGCCGGCCT GGGCCCTAGACGCTCCAGCAAGGTCTAGAAAGGAGCCAAGGAAG ATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAG CTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGCCGACAA GGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCT GTTCAAGAGAGTGAGAAGCATCTCAACAAGCTGACCCCTCAGATG TTCCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCG AGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTAT CTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTTGCC TGATGGCATTGAAGGTGCCAACCCAGGAGAAGCCTACTGTGACCGT CAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTCCGAG AAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAA ATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGGAG GAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCA ACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAG GCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGACG AAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGCAA GGACCTGGACTTCGAGAAGGCCAAAGCCTCGAATGGACCAGTACTTC AACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGA ATCAGATTTCATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAACT GGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAGA TCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAGG TGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGCGGGCC GCCCCGACCTCCTATCAGCAGAGGCCTTCTCTGGTAGACGACGGCG GCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACT TCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAA CCAGCTGTTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGT TCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCTG CCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAA GCTGTGCCGACCGAGAGCACCGACAACCGGCGAGTTGTGCAGAGAA GCAGCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATAGAGGTG ACCGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCGGC TTGACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTTCAGCAAGGA GGTCGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGACTC AGAAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGGAAGAGAC GCCGTGAAGAGGGAGGCGCACTGCCTCCTGTGAGCCCTCTGAAGG CGGCACTGAGCGAAGAAGAAGCTTGAGAAGAAGAGTAAGGCGATCAT CGAGGAGTACCTGCACCTGAACGACATGAAGGAGGCCGTGCAGTGC GTGCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCCGGCA CGGCGTGGAGTCCACACTGGAAAGATCTGCCATTGCTAGGGAGCAT ATGGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAAGTAC TGCCAGTACTATCAGGGTTTATACGAGATCCTCGAACTGGCCGAAG ACATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCGCAGAACTG GTGACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGAGTTGTT CAGAGAAATCACAAAGCCACTGCGCCCACTGGGCAAGGCGGCTAGT CTCCTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGGTCCAAA

SEQ ID NO	Sequence information	Sequence
		GAAGGTGGGCACCCTGTGGAGGGAAGCTGGACTCTCCTGGAAGGAA TTCCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGAACAGAA GGTTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGTACGCGG GCTCTGCCGAGTGAGGAGCTCAACAGACAACCTCGAGAAGCTCCTCA AGGAGGGATCCTCAAATCAGAGAGTGTTCGACTGGATCGAGGCCAA TCTCAGCGAGCAGCAAATCGTGAGCAACACGTTGGTTCGCGCTCTCA TGACAGCCGTGTGCTACTCAGCCATTATCTTCGAGACGCCTCTCCGC GTGGACGTGGCAGTGCTCAAGGCCCCGCGCTAAGCTGTTACAGAAGT ACCTGTGCGACGAGCAGAAGGAACTGCAGGGCCCTGTACGCCCTGCA AGCTCTGGTCGTCACTCGAGCAGCCTCCTAACCTGCTGAGAATGT TCTTCGACGCATTGTACGACGAGGACGTTGTGAAGGAAGACGCTTTC TACAGCTGGGAGTCTAGTAAGGATCCTGCGGAACAACAAGGCAAGG GCGTTGCTCTAAAGAGCGTGACCGCCTTCTTCAAGTGGCTGCGAGAG GCTGAGGAAGAGTCCGACCACAAC
10	MBP-eIF4GdN (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCS VRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVKAMQ GLLKDGNPIPSAIAANSYIYGGGGSIFASMOKPEGLPHISDVVLDKANKT PLRPLDPTRLQGINCGPDFTPSFANLGRITLSTRGPPRGGPGGELPRGPA GLGPRRSQQGPRKEPRKIIATVLMTEDIKLNKAIEKAWKPSSKRTAADKD RGEEDADGSKTQDLFRRVRSILNKLTPQMFMQQLMKQVTQLAIDTEERL KGVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTVTNFRKLL LNRCKEKEFEKDKDDDEVFEKKQKEMDEAATAEERGRLEELEEARDIA RRRSLGNIKFIGELFKLKMLTEAIMHDCVVKLLKNHDEESLECLCRLLT TIGKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDRGSN WVPRRGDQGPKTIDQIHKEAEMEEHREHIKVQQLMAKGS DKRRGGPPG PPISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKPGSIDSNNQLFAPG GRLSWGKGSSGSGAKPSDAASEAARPATSTLNRFSA LQQA VPTTESTD NRRVVQRSSLRERGEKAGDRGDRLE R SERGGDRGDR LDRARTPATKR SFSKEVEERSRERPSQPEGLRKAASLTEDRDRGRDAVKREAA LPPVSPL KAALSEEELEKKSKAIIIEYLHLNDMKEAVQCVQELASPSLLFIFVRHGV ESTLERSAIAREHMGQLLHQLLCAGHLSTAQYYQGLYEILELAEDMEID IPHVWLYLAELVTPILQEGGVPMGELFREITKPLRPLGKAASLLEILGLL CKSMGPKKVGT LWREAGLSWKEFLPEGQDIGAFVAEQKVEYTLGEESE APGQRALPSEELNRQLEKLLKEGSSNQRFVDWIEANLSEQQIVSNTLVR ALMTAVCYS AII FETPLRVDVAVLKARAKLLQKYL CDEQKELQALYAL QALVVTLEQPPNLLRMFFDALYDEDVVKEDAFYSWESSKDP AEQQGK GVALKS VTAFFKWLREAEESDHN
11	MBP-eIF4GdN (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGGTA CCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCGCC GAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGACCT GCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAAGTACACCATCAA GGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGGAG CTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATCGT GAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTAGC GCCATCGCCGCCAATTCAGGCATCTACGAGGCGGTGGAAGCATAT TCGCGAGCATGCAGAAGCCTGAAGGTCTGCCTCACATCAGCGACGT GGTGCTGGACAAGGCCAACAAGACCCCTCTTAGACCTCTGGACCCTA CCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCCTTCTTTC GCAAACCTGGGCAGAACCAACCCTGAGCACCAGAGGCCCTCCTAGAG GTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCTGGGCCCT AGACGCTCCAGCAAGGTCCTAGAAAGGAGCCAAGGAAGATCATCG CCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAGCTGAGAA GGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGC GACAAGGACAG

SEQ ID NO	Sequence information	Sequence
		AGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCTGTTTCAGA AGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAGATGTTCCAGC AGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCGAGGAGAG ACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTATCTCAGAGC CTAACTTCAGCGTGCCCTACGCCAACATGTGCCGTTGCCTGATGGCA TTGAAGGTGCCAACCACCGAGAAGCCTACTGTGACCGTCAATTTCCG TAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTTCGAGAAGGATAAG GACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAAATGGACGAG GCCGCCACCGCAGAGGAAAGGGGGCCGATTAAAGGAGGAGCTGGAG GAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCAACATCAAGT TCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAGGGCCATAATG CACGACTGCGTGTTGAAGCTACTGAAGAACCACGACGAAGAAAGCC TGGAGTGCCCTGTGCAGACTGCTGACCACCATCGGCAAGGACCTGGA CTTCGAGAAGGCAAAGCCTCGAATGGACCAGTACTTCAACCAGATG GAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGAATCAGATTCA TGCTGCAGGACGTACTGGACCTGCGCGGAAGCAACTGGGTGCCAAG GAGAGGGGACCAAGGACCAAGACCATCGACCATCCACAAGGA AGCGGAGATGGAGGAGCACAGAGAGCACATAAAGGTGCAGCAGCT TATGGCCAAGGGCAGCGACAAGCGAAGAGGGCGCCCGCCGACCT CCTATCAGCAGAGGCCTTCTCTGGTAGACGACGGCGCTGGAACA CCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACTTCCCGTCTT ACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAACCAGCTGT TCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGTTCCGGCGG ATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCTGCCAGACCT GCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAAGCTGTGCC GACCGAGAGCACCGACAACCGGCGAGTTGTGCAGAGAAGCAGCCTG AGCAGAGAGAGGGGCGAGAAAGCCGGCGATAGAGGTGACCGACTG GAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCGGCTTGACAGA GCCAGAACCCCTGCCACAAAGCGATCGTTTCAGCAAGGAGGTGCAAG AGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGACTCAGAAAGGC AGCCAGTCTAACAGAGGACCGCGACAGGGGAAGAGACGCCGTGAA GAGGGAGGCCGCACTGCCTCCTGTGAGCCCTCTGAAGGCGGCACTG AGCGAAGAAGAACTTGAGAAGAAGAGTAAGGCGATCATCGAGGAG TACCTGCACCTGAACGACATGAAGGAGGCCGTGCAGTGCGTGCAAG AGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCCGGCACGGCGTG GAGTCCACACTGGAAAGATCTGCCATTGCTAGGGAGCATATGGGCC AGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAAGTACTGCCAG TACTATCAGGGTTTATACGAGATCCTCGAACTGGCCGAAGACATGGA GATCGACATCCCTCACGTGTGGCTGTACCTCGCAGAACTGGTGACCC CTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGAGTTGTTTCAGAGA AATCACAAAGCCACTGCGCCCACTGGGCAAGGCGGCTAGTCTCCTTC TGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGGTCCAAAGAAGGT GGGCACCCTGTGGAGGGAAGCTGGACTCTCCTGGAAGGAATTCCTC CCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGAACAGAAGGTTG AGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGTGAGCGGGCTCT GCCGAGTGAGGAGCTCAACAGACAACCTCGAGAAGCTCCTCAAGGAG GGATCCTCAAATCAGAGAGTGTTGACTGGATCGAGGCCAATCTCA GCGAGCAGCAAATCGTGAGCAACACGTTGGTTTCGCGCTCTCATGAC AGCCGTGTGCTACTCAGCCATTATCTTCGAGACGCTCTCCGCGTGG ACGTGGCAGTGCTCAAGGCCCGCGCTAAGCTGTTACAGAAGTACCT GTGCGACGAGCAGAAGGAACTGCAGGCCCTGTACGCCCTGCAAGCT CTGGTCGTACACTCGAGCAGCCTCCTAACCTGCTGAGAATGTTCTT CGACGCATTGTACGACGAGGACGTTGTGAAGGAAGACGCTTTCTAC

SEQ ID NO	Sequence information	Sequence
		AGCTGGGAGTCTAGTAAGGATCCTGCGGAACAACAAGGCAAGGGCG TTGCTCTAAAGAGCGTGACCGCCTTCTTCAAGTGGCTGCGAGAGGCT GAGGAAGAGTCGGACCACAAC
12	MBP_eIF4G(fl) (aa)	MASNFTQFVLVDNNGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLDGNIPIPSAIAANSYIYGGGGSNKAPQSTGPPAPSPGLPQPAF PPGQTAPVVFSTPQATQMNTSPQRQHFYPSRAQPPSSAASRVQSAAPA RPGPAAHVYPAGSQVMMIPSIQSYSPASQGAAYYIPGQGRSTYVVPQTQQYP VQPGAPGFYPGA SPTEFGTYAGAYYPAQGVQQFPTGVAPTPLMNQPP QIAPKRERKTIRIRDPNQGGKDITEEIMSGARTASTPTPPQTGGGLEPQA NGETPQVAVIVRPDDRSQGAIIADRPGLPGPEHSPSESQSPSPSTPSPSPV LEPGSEPNLAVLSIPGDMTTIQMSVEESTPISRETGEPYRLSPEPTPLAEP ILEVEVTL SKVPPESEFSSSPLQAPTPLASHTVEIHEPNGMVPSLEDLEPEV ESSPELAPPPACPSSESPVPIAPTAQPEELLNGAPSPPAVDLSPVSEPEEQAK EVTASMAPPTIPSPATPATAPSATSPAQEEEMEEEEEEEEGEAGEAGEAES EKGGEEELLPESTPIPANLSQNLEAAAATQVAVSVPKRRRKIKELNKKE AVGDLDDAFKEANPAVPEVENQPPAGSNPGPESESGVPPRPPEEADETW DSKEDKIHNAENIQPGEQKYEYKSDQWKPLNLEEKRYDREFFLGFQFI FASMQKPEGLPHISDVLDKANKTPLRLDPTRLQGINCGPDFTPSFANL GRITTLSTRGPPRGPGGELPRGPAGLGPRRSQQGPRKEPRKIIATVLMTE DIKLNKAEKAWKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSILNKL TPQMFMQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCR CLMALKVPTTEKPTVTNFRKLLLNRCKEKEFEKDKDDDEVFEKKQKE MDEAATAEERGRLEKEELEEARDIARRRSLGNIFIGELFKLKMTEAIM HDCVVKLLKNHDEESLECLCRLTTIGKDLDFEAKPRMDQYFNQMEK IIEKKTSSRIRFMLQDVLDRGSNWVPRRGDQGPKTIDQIHKEAEMEE HREHIKVQQLMAKGS DKRRGGPPGPISRLPLVDDGGWNTVPISKGR PIDTSRLTKITKPGSIDSNNQLFAPGGRLSWGKGSSGGSGAKPSDAASEA ARPATSTLNRFSALQQAVPTESTDNRRVVQRSSLSRERGEKAGDRGDR ERSERGGDRGDRLDRLDRARTPATKRSFSKEVEERSRERPSQPEGLRKAASL TEDRDRGRDAVKREAAALPPVSPLKAALSEEELEKKS KAIIIEYLHLNDM KEAVQCVQELASPSLLFIFVRHGVESTLERSAIAREHMGQLLHQLLCAG HLSTAQYYQGLYEILELAEDMEIDIPHVWLYLAELVTPILQEGGVPMGE LFREITKPLRPLGKAASLLLEILGLLCKSMGPKKVGTLWREAGLSWKEF LPEGQDIGAFVAEQKVEYTLGEESEAPGQRALPSEELNRQLEKLLKEGS SNQRVFDWIEANLSEQQIVSNLVRALMTAVCYSAIFETPLRVDVAVL KARAKLLQKYL CDEQKELQALYALQALVVTLEQPPNLLRMFFDALYD EDVVKEDAFYSWESSKDPAEQQKGVALKSVTAFFKWLREAEESDH N
13	MBP_eIF4G(fl) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGCAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCAA CAAGGCCCCCTCAGAGCACCGGCCCTCCTCCTGCCCTAGCCCTGGCC TGCCTCAGCCTGCCTCCCTCCTGGCCAGACCGCCCCCTGTGGTGTTC GCACCCCTCAGGCCACCCAGATGAACACACCTAGCCAGCCTAGACA GCACTTCTACCCTAGCAGAGCCCAGCCTCCTAGCAGCGCCGCCAGCA GAGTGCAGTCCGCTGCACCTGCCAGACCTGGCCCTGCCGCCACCGTG TACCCTGCCGGCAGCCAGGTGATGATGATCCCGAGTCAAATCAGCTA

SEQ ID NO	Sequence information	Sequence
		CCCTGCAAGCCAGGGGCGCCTACTACATCCCGGGCCAGGGCAGAAGC ACCTACGTGGTGCCTACCCAGCAGTACCCTGTGCAGCCTGGCGCCCC TGGTTTCTATCCTGGCGCAAGCCCTACCGAGTTCGGAACCTACGCCG GCGCTTACTATCCAGCTCAGGGCGTGCAGCAGTTCCTACCGGCGTG GCCCCTACCCCTGTGCTGATGAATCAGCCACCTCAGATCGCCCCCTAA GCGCGAGCGCAAGACCATCAGAATCCGCGATCCTAACCAGGGCGGC AAGGACATCACCGAGGAGATCATGAGCGGCGCCAGAACAGCAAGTA CTCCAACCCCGCCACAAACCGGCGGCGGCGCTGGAGCCTCAAGCCAA CGGCGAGACGCCACAAGTGGCCGTGATCGTACGCCCTGACGACCGG AGCCAAGGTGCAATCATCGCCGATAGGCCTGGCCTCCAGGTCCGG AGCACAGCCCTAGCGAGTCCCAGCCGTCTTACCATCACCAACCCCT AGTCCATCCCCTGTTCTCGAACCAGGCAGCGAGCCTAACCTGGCCGT GCTGAGCATAACAGGTGACACCATGACCACCATCCAGATGAGCGTG GAGGAGAGCACCCCAATCAGCAGAGAACTGGAGAGCCTTACAGAC TGTCCCCAGAGCCGACCCCACTGGCCGAGCCAATACTGGAGGTGGA GGTGACCCTGAGCAAGCCTGTGCCTGAGAGCGAGTTTACGCTCCTCTC CACTGCAGGCCCCAACTCCTCTCGCAAGCCACACCGTGGAGATCCAC GAACCTAACGGCATGGTACCAAGCGAAGATCTTGAGCCAGAGGTCG AATCAAGCCAGAACTGGCCCCTCCACCTGCCTGCCCCTGCTGAATCT CCGGTCCCTATCGCTCCTACGGCACAGCCTGAGGAGCTGCTGAACGG TGCCCCGAGCCCTCCAGCAGTGGACTTATCCCCAGTATCAGAGCCTG AAGAACAGGCCAAGGAGGTAAGTGCCTCTATGGCGCCACCTACCAT ACCTTCGGCAACACCGGCTACAGCACCATCTGCGACTAGTCCGGCTC AGGAGGAGGAGATGGAGGAAGAAGAAGAGGAAGAGGAGGGCGAG GCCGGAGAGGCCGGTGAAGCCGAGTCCGAGAAGGGCGGCGAAGAA CTTCTCCCTCCAGAGTCAACTCCTATCCCTGCCAACCTTAGTCAGAAT CTGGAGGCCCGCCGCTACTCAGGTTGCAGTGAGCGTGCCAAAGA GACGTCGCAAGATCAAGGAGCTGAACAAGAAGGAGGCCGTGGGCG ACCTGCTGGACGCCTTCAAGGAGGCAAACCCGGCGGTGCCTGAAGT GGAGAATCAGCCTCCGGCCGGATCAAACCTGGTCTGAGAGTGAA GGCAGCGGCGTCCCACCAAGACCTGAAGAGGCTGACGAAACTTGGG ACAGCAAGGAGGACAAGATCCACAACGCCGAGAACATCCAGCCAG GCGAGCAGAAGTACGAGTACAAGAGCGACCAGTGGAAGCCTCTTAA CCTTGAAGAGAAGAAGAGATACGACAGAGAGTTCTGCTGGGCTTC CAGTTCATATTGCGGAGCATGCAGAAGCCTGAAGGTCTGCCTCACAT CAGCGACGTGGTGCTGGACAAGGCCAACAAGACCCCTCTTAGACCT CTGGACCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCAC TCCTTCTTTGCAAACTGGGCAGAACCCACCTGAGCACCAGAGGCC CTCCTAGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGG CCTGGGCCCTAGACGCTCCCAGCAAGGTCTAGAAAGGAGCCAAGG AAGATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACA AAGCTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGCA CAAGGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGA CCTGTTTCAAGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAG ATGTTCCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACA CCGAGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGC TATCTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTT GCCTGATGGCATTGAAGGTGCCAACCACCGAGAAGCCTACTGTGAC CGTCAATTTCCGTAACTGCTGCTGAACCGGTGCCAGAAAGAGTTTCG AGAAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAG AAATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGG AGGAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGG CAACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCG

SEQ ID NO	Sequence information	Sequence
		AGGCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGA CGAAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGC AAGGACCTGGACTTCGAGAAGGCCAAAGCCTCGAATGGACCAGTACT TCAACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCA GAATCAGATTCATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAA CTGGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCA GATCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAA GGTGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGGCGGC CCGCCCCGACCTCCTATCAGCAGAGGCCTTCCTCTGGTAGACGACGG CGGCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACA CTTCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAAC AACCAGCTGTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAA GTTCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGC TGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGC AAGCTGTGCCGACCGAGAGCACCGACAACCGGCGAGTTGTGCAGAG AAGCAGCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATAGAGG TGACCGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCG GCTTGACAGAGCCAGAACCCTGCCACAAGCGATCGTTTCAGCAAG GAGGTCTGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGA CTCAGAAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGGAAGA GACGCCGTGAAGAGGGAGGCGCACTGCCTCCTGTGAGCCCTCTGA AGGCGGCACTGAGCGAAGAAGAAGTCTGAGAAGAAGAGTAAGGCGA TCATCGAGGAGTACCTGCACCTGAACGACATGAAGGAGGCCGTGCA GTGCGTGCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCC GGCACGGCGTGGAGTCCACACTGGAAGATCTGCCATTGCTAGGGA GCATATGGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAA GTACTGCCCAGTACTATCAGGGTTTATACGAGATCCTCGAACTGGCC GAAGACATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCGCAG AACTGGTGACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGA GTTGTTTCAGAGAAATCACAAAGCCACTGCGCCCACTGGGCAAGGCG GCTAGTCTCCTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGG TCCAAAGAAGGTGGGCACCCTGTGGAGGGAAGCTGGACTCTCCTGG AAGGAATTCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGA ACAGAAGGTTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGT CAGCGGGCTCTGCCGAGTGAGGAGCTCAACAGACAACTCGAGAAGC TCCTCAAGGAGGGATCCTCAAATCAGAGAGTGTTGCACTGGATCGA GGCCAATCTCAGCGAGCAGCAAATCGTGAGCAACACGTTGGTTTCG GCTCTCATGACAGCCGTGTGCTACTCAGCCATTATCTTCGAGACGCC TCTCCGCGTGGACGTGGCAGTGCTCAAGGCCCGCGCTAAGCTGTTAC AGAAGTACCTGTGCGACGAGCAGAGAAGGAAGTGCAGGCCCTGTACGC CCTGCAAGCTCTGGTCGTCACTCGAGCAGCCTCCTAACCTGCTGA GAATGTTCTTCGACGCATTGTACGACGAGGACGTTGTGAAGGAAGA CGCTTTCTACAGCTGGGAGTCTAGTAAGGATCCTGCGGAACAACAA GGCAAGGGCGTTGCTCTAAAGAGCGTGACCGCCTTCTTCAAGTGGCT GCGAGAGGCTGAGGAAGAGTCGGACCACAAC
14	MBP_eIF4G-dN (623-1599, 3A) (aa)	MASNFTQFVLVDNNGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLDGNIPIPSAIAANSIYGGGGSIFASMQKPEGLPHISDVLDK ANKTPLRLDPTRLQGINCGPDFTPSFANLGRITLSTRGPPRGPGGELP RGPAGLGPRRSQQGPRKEPRKIIATVLMTEDIKLNKAEKAWKPSSKRTA ADKDRGEEDADGSKTQDLFRRVRSIANKATPQMAQQLMKQVTQLAID TEERLKGVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTVTNFF

SEQ ID NO	Sequence information	Sequence
		RKLLLNRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGRLEKEELE ARDIARRRSLGNIKFIGELFKLKMLTEAIMHDCVVKLLKNHDEESLECL CRLTTIGKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIRFMLQDVL LRGSNWVPRRGDQGPKTIDQIHKEAEMEEHREHIKVVQLMAKGSDKRR GGPPGPPISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKPGSIDSNNQ LFAPGGRLSWGKGSSGGSGAKPSDAASEAARPATSTLNRFSAALQQAVP TESTDNRRVVQRSSLSRERGEKAGDRGDR LERSERGGDRGDRDRARTPATKRSFSKEVEERSRERPSQPEGLRKAAS LTEDRDRGRDAVKREAAALPPVSPLKAALSEEELEKKSKAIIIEYLHLND MKEAVQCQVELASPSLLFIFVRHGVSTLERSAIAREHMGQLLHQLLCA GHLSTAQYYQGLYEILELAEDMEIDIPHVWLYLAELVTPILQEGGVPMG ELFREITKPLRPLGKAASLLEILGLLCKSMGPKKVGTWREAGLSWKE FLPEGQDIGAFVAEQKVEYTLGEESEAPGQRALPSEELNRQLEKLLKEG SSNQRVFDWIEANLSEQQIVSNTLVRALMTAVCYSAIIFETPLRVDVAVL KARAKLLQKYLCEQKELQALYALQALVVTLEQPPNLLRMFFDALYD EDVVKEDAFYSWESSKDPAEQQGKGVALK SVTAFFKWLREAEESDHN
15	MBP_eIF4G-dN (623-1599, 3A) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCAT ATTCGCGAGCATGCAGAAGCCTGAAGGTCTGCCTCACATCAGCGAC GTGGTGCTGGACAAGGCCAACAAGACCCCTCTTAGACCTCTGGACCC TACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCCTTCTT TCGCAAACCTGGGCAGAACCAACCTGAGCACCAGAGGCCCTCCTAG AGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCTGGGC CCTAGACGCTCCCAGCAAGGTCCTAGAAAGGAGCCAAGGAAGATCA TCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAGCTGA GAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCGCGGACAAGGA CAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCTGTTC AGAAGAGTGAGAAGCATCGCGAACAAGGCGACCCCTCAGATGGCGC AGCAGCTGATGAAGCAGGTGACGACGCTCGCCATCGACACCGAGGA GAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTATCTCA GAGCCTAACTTCAGCGTGCCCTACGCCAACATGTGCCGTTGCCTGAT GGCATTGAAGGTGCCAACCACCGAGAAGCCTACTGTGACCGTCAAT TTCCGTAACTGCTGCTGAACCGGTGCCAGAAAGAGTTGAGAAGG ATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAAATGG ACGAGGCCGCCACCGCAGAGGAAAGGGCCGATTAAAGGAGGAGC TGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCAACAT CAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAGGCC ATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGACGAAG AAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGCAAGGA CCTGGACTTCGAGAAGGCAAAGCCTCGAATGGACCAGTACTTCAAC CAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGAATC AGATTCATGCTGCAGGACGTAAGTGGACCTGCGCGGAAGCAACTGGG TGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAGATCCA CAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAGGTGCA GCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGCGGCCCGCC

SEQ ID NO	Sequence information	Sequence
		GGACCTCCTATCAGCAGAGGCCTTCCTCTGGTAGACGACGGCGGCTG GAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACTTCCC GTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAACCA GCTGTTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGTTCC GGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCTGCCA GACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAAGCT GTGCCGACCGAGAGCACCGACAACCGGCGAGTTGTGCAGAGAAGCA GCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATAGAGGTGACC GACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCGGCTTG ACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTTCAGCAAGGAGGT CGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGACTCAGA AAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGGAAGAGACGCC GTGAAGAGGGAGGCCGCACTGCCTCCTGTGAGCCCTCTGAAGGCGG CACTGAGCGAAGAAGAACTTGAGAAGAAGAGTAAGGCGATCATCGA GGAGTACCTGCACCTGAACGACATGAAGGAGGCCGTGCAGTGCCTG CAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCCGGCACGG CGTGGAGTCCACACTGGAAAGATCTGCCATTGTAGGGAGCATATG GGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAAGTACTGC CCAGTACTATCAGGGTTTATACGAGATCCTCGAACTGGCCGAAGACA TGGAGATCGACATCCCTCACGTGTGGCTGTACCTCGCAGAACTGGTG ACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGAGTTGTTTCA AGAAATCACAAAGCCACTGCGCCCACTGGGCAAGGCGGCTAGTCTC CTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGGTCCAAAGAA GGTGGGCACCTGTGGAGGGAAGCTGGACTCTCCTGGAAGGAATTC CTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGAACAGAAGG TTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGTGAGCGGGC TCTGCCGAGTGAGGAGCTCAACAGACAACCTCGAGAAGCTCCTCAAG GAGGGATCCTCAAATCAGAGAGTGTTGCGACTGGATCGAGGCCAATC TCAGCGAGCAGCAAATCGTGAGCAACACGTTGGTTCGCGCTCTCATG ACAGCCGTGTGCTACTCAGCCATTATCTTCGAGACGCCTCTCCGCGT GGACGTGGCAGTGCTCAAGGCCCGCGCTAAGCTGTTACAGAAGTAC CTGTGCGACGAGCAGAAGGAACTGCAGGCCCTGTACGCCCTGCAAG CTCTGGTCGTACACTCGAGCAGCCTCCTAACCTGCTGAGAATGTTC TTCGACGCATTGTACGACGAGGACGTTGTGAAGGAAGACGCTTTCTA CAGCTGGGAGTCTAGTAAGGATCCTGCGGAACAACAAGGCAAGGGC GTTGCTCTAAAGAGCGTGACCGCCTTCTTCAAGTGGCTGCGAGAGGC TGAGGAAGAGTCGGACCACAAC
16	MBP_eIF4G-dN2 (653-1599) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNIPIPSAIAANSIYGGGSDPTRLQGINCGPDFTPSFANLG RTTLSTRGPPRGGPGGELPRGPAGLGPRRSQQGPRKEPRKIIATVLMTE IKLNKAEKAWKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSILNKLT PQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCRC LMALKVPTTEKPTVTNFRKLLLNRCQKEFEKDKDDDEVFEKKQKEM DEAATAEERGRLEKEELEEARDIARRRSLGNIKFIGELFKLKMTEAIMHD CVVKLLKNHDEESLECLCRLLTTIGKDLDFEKAKPRMDQYFNQMEKIIK EKKTSSRIRFMLQDVLDRGSNWVPRRGDQGPKTIDQIHKEAEMEHR EHIKVQQLMAKGSDKRRGGPPGPISRGLPLVDDGGWNTVPISKGSRI DTSRLTKITKPGSIDSNNQLFAPGGRLSWGKGSSGSGAKPSDAASEAA RPATSTLNRFSALQQAVPTSTDNRVVRSSLSRERGEKAGDRGDRLE RSERGGDRGDRLDARTPATKRSFSKEVEERSRERPSQPEGLRKAASLT EDRDRGRDAVKREAALPPVSPLKAALSEEELEKKSKAIIIEYLHLNDMK

SEQ ID NO	Sequence information	Sequence
		EAVQCVQELASPSLLFIFVRHGVESTLERSAIAREHMGQLLHQLLCAGH LSTAQYYQGLYEILELAEDMEIDIPHVWLYLAELVTPILQEGGVPMGEL FREITKPLRPLG KAASLLLEILGLLCKSMGPKKVGTLWREAGLSWKEFLPEGQDIGAFVA EQKVEYTLGEESEAPGQRALPSEELNRQLEKLLKEGSSNQRVFDWIEAN LSEQQIVSNTLVRLMTAVCYSAIFETPLRVDVAVLKARAKLLQKYL DEQKELQALYALQALVVTLEQPPNLLRMFFDALYDEDVVKEDAFYSW ESSKDPAEQQGKGVALKSVTAFFKWLREAEEDSDHN
17	MBP_eIF4G-dN2 (653-1599) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGA CCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCCTT CTTTCGCAAACCTGGGCAGAACACCCTGAGCACCAGAGGCGCCTCCT AGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCTGG GCCCTAGACGCTCCCGAGCAAGGTCTAGAAAGGAGCCAAGGAAGAT CATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAGCT GAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGGACAAG GACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCTGT TCAGAAGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAGATGTT CCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCGAG GAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTATCT CAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTTGCCTG ATGGCATTGAAGGTGCCAACACCGAGAAAGCCTACTGTGACCGTCA ATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTCGAGAA GGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAAAT GGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGGAGGA GCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCAAC ATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAGGC CATAATGCACGACTGCGTGCTGAAGCTACTGAAGAACCACGACGAA GAAAGCCTGGAGTGCTGTGCAGACTGCTGACCACCATCGGCAAGG ACCTGGACTTCGAGAAGGCCAAAGCCTCGAATGGACCACTTCAA CCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGAAT CAGATTCATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAACTGG GTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAGATCC ACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAGGTGC AGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGCGGCCGCC CGGACCTCCTATCAGCAGAGGCCTTCTCTGGTAGACGACGGCGGCT GGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACTTCC CGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAACC AGCTGTTCCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGTTC CGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGTGCC AGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAAGC TGTGCCGACCGAGAGACCGACAACCGGCGAGTTGTGCAGAGAAGC AGCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATAGAGGTGAC CGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCGGCTT GACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTCAGCAAGGAGG TCGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGACTCAG

SEQ ID NO	Sequence information	Sequence
		AAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGGAAGAGACGC CGTGAAGAGGGAGGCCGCACTGCCTCCTGTGAGCCCTCTGAAGGCG GCACTGAGCGAAGAAGAAGCTTGAGAAGAAGAGTAAGGCGATCATCG AGGAGTACCTGCACCTGAACGACATGAAGGAGGCCGTGCAGTGCCT GCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCCGGCACG GCGTGGAGTCCACACTGGAAAGATCTGCCATTGCTAGGGAGCATAT GGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAAGTACTG CCCAGTACTATCAGGGTTTATACGAGATCCTCGAACTGGCCGAAGAC ATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCGCAGAACTGGT GACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGAGTTGTTCA GAGAAATCACAAGCCACTGCGCCCACTGGGCAAGGCGGCTAGTCT CCTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGGTCCAAAGA AGGTGGGCACCCTGTGGAGGGAAGCTGGACTCTCCTGGAAGGAATT CCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGAACAGAAG GTTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGTGAGCGGG CTCTGCCGAGTGAGGAGCTCAACAGACAACCTCGAAGAGTCTCTCAA GGAGGGATCCTCAAATCAGAGAGTGTTCGACTGGATCGAGGCCAAT CTCAGCGAGCAGCAAATCGTGAGCAACACGTTGGTTTCGCGCTCTCAT GACAGCCGTGTGCTACTCAGCCATTATCTTCGAGACGCCTCTCCGCG TGGACGTGGCAGTGCTCAAGGCCCGCGCTAAGCTGTTACAGAAGTA CCTGTGCGACGAGCAGAAGGAAGTGCAGGCCCTGTACGCCCTGCAA GCTCTGGTCGTCACTCGAGCAGCCTCCTAACCTGCTGAGAATGTT CTTCGACGCATTGTACGACGAGGACGTTGTGAAGGAAGACGCTTTCT ACAGCTGGGAGTCTAGTAAGGATCCTGCGGAACAACAAGGCAAGGG CGTTGCTCTAAAGAGCGTGACCGCCTTCTTCAAGTGGCTGCGAGAGG CTGAGGAAGAGTCGGACCACAAC
18	MBP_eIF4G-dN3 (653-1451) (aa)	MASNFTQFVLVDNGGTGDVTVPNSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLDGNIPIPSAIAANSIYGGGGSDPTRLQGINCGPDFTPSFANLG RTTLSTRGPPRGGPGGELPRGPAGLGPRRSQGPKEPRKIIATVLMTE IKLNKAEKAWKPSSKRTAADKDRGEEDADGSKTQDLFRVRVILNKL PQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCR LMALKVPTTEKPTVTNFRKLLNRCQKEFEKDKDDDEVFEKKQKEM DEAATAEERGLKEELEEARDIARRRSLGNIKFELFKLMLTEAIMHD CVVKLLKNHDEESLECLCRLTTIGKDLDFEKAAPRMDQYFNQMEKIIK EKKTSSRIRFMLQDVLDRGSNWVPRRGDQGPKTIDQIHKEAEMEHR EHIKVQQLMAKGSDDRGGPPGPPISRGLPLVDDGGWNTVPISKGSRI DTSRLTKITKPGSIDSNQLFAPGGRLSWGKGSSGGSGAKPSDAASEAA RPATSTLNRFSALQQAVPTESTDNRRVVQRSSLRERGEKAGDRGDRLE RSEGGDRGDRLDRAATPATKRSFSKEVEERSRERPSQPEGLRKAASLT EDRDRGRDAVKREAALPPVSPLKAALSEEELEKKSAAIIEYLHLNDMK EAVQCVQELASPSLLFIFVRHGVESTLERSAIAREHMGQLLHQLLCAGH LSTAQYYQGLYEILELAEDMEIDIPHVWLYLAELVTPILQEGGVPMGEL FREITKPLRPLG KAASLLLEILGLLCKSMGPKKVGTLWREAGLSWKEFLPEGQDIGAFVA EQKVEYTLGEES EAPGQRALPSEELNRQLEKLL
19	MBP_eIF4G-dN3 (653-1451) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG

SEQ ID NO	Sequence information	Sequence
		AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGA CCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCCTT CTTTCGCAAACCTGGGCAGAACCACCCTGAGCACCAGAGGGCCCTCCT AGAGGTGGTCCCGGGCGGAGAACTCCCCAGGGGTCTGCCGGCCTGG GCCCTAGACGCTCCCAGCAAGGTCTAGAAAAGGAGCCAAGGAAGAT CATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAGCT GAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGCCGACAAG GACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCTGT TCAGAAAGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAGATGTT CCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCGAG GAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTATCT CAGAGCCTAACTTCAGCGTGCCCTACGCCAACATGTGCCGTTGCCGTG ATGGCATTGAAGGTGCCAACCCAGAGAAGCCTACTGTGACCGTCA ATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTTCGAGAA GGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAAAT GGACGAGGCCGCCACCGCAGAGGAAAGGGGCCATTAAAGGAGGA GCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCAAC ATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAGGC CATAATGCACGACTGCGTGCTGAAGCTACTGAAGAACCACGACGAA GAAAGCCTGGAGTGCTGTGCAGACTGCTGACCACCATCGGCAAGG ACCTGGACTTCGAGAAGGCCAAAGCCTCGAATGGACCAGTACTTCAA CCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGAAT CAGATTCATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAACTGG GTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAGATCC ACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAGGTGC AGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGGCGGCCGCC CGGACCTCCTATCAGCAGAGGCCTTCCTCTGGTAGACGACGGCGGCT GGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACTTCC CGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAACC AGCTGTTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGTTC CGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCTGCC AGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAAGC TGTGCCGACCGAGAGCACCGACAACCGGCGAGTTGTGCAGAGAAGC AGCCTGAGCAGAGAGAGGGGGCGAGAAAGCCGGCGATAGAGGTGAC CGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCGGCTT GACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTCAGCAAGGAGG TCGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGACTCAG AAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGGAAGAGACGC CGTGAAGAGGGAGGCCGCACTGCCTCCTGTGAGCCCTCTGAAGGCG GCACTGAGCGAAGAAGAAGCTTGAGAAGAAGAGTAAGGCGATCATCG AGGAGTACCTGCACCTGAACGACATGAAGGAGGCCGTGCAGTGCGT GCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCCGGCACG GCGTGAGTCCACACTGGAAAGATCTGCCATTGCTAGGGAGCATAT GGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAAGTACTG CCCAGTACTATCAGGGTTTATACGAGATCCTCGAACTGGCCGAAGAC ATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCGCAGAACTGGT GACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGAGTTGTTCA GAGAAATCACAAAGCCACTGCGCCCACTGGGCAAGGCGGCTAGTCT CCTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGGTCCAAAGA AGGTGGGCACCCTGTGGAGGGAAGCTGGACTCTCCTGGAAGGAATT CCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGAACAGAAG

SEQ ID NO	Sequence information	Sequence
		GTTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGTCAGCGGGCTCTGCCGAGTGAGGAGCTCAACAGACAACTCGAGAAGCTCCTC
20	MBP_eIF4G-mid1 (674-1079) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLDGNIPIPSAIAANSYIYGGGSGRTTLSTRGPPRGGPGGELPR GPAGLGPRRSQQGPRKEPRKIIATVLMTEDIKLNKAEKAWKPSSKRTAA DKDRGEEDADGSKTQDLFRRVRSILNKLTPQMFQQLMKQVTQLAIDTE ERLKGVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTVTVNFR KLLLNRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGLKEELEEA RDIARRRSLGNIKFIGELFKLKLMLTEAIMHDCVVKLLKNHDEESLECLC RLLTTIGKDLDFEKAAPRMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDL RGSNWVPRRGDQGPKTIDQIHKEAEMEEHREHIKVQQLMAKGSDKR RGGPPGPPISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKPGSID
21	MBP_eIF4G-mid1 (674-1079) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGACAACGGCGG TACCGGAGACGTGACCGTGCCCTTCTAAGTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGG CAGAACCACCTGAGCACCAGAGGCCCTCTAGAGGTGGTCCCGGC GGAGAACTCCCCAGGGGTCTGCGGCCTGGGCCCTAGACGCTCCC AGCAAGGTCTAGAAAGGAGCCAAGGAAGATCATCGCCACCGTGCT GATGACCGAGGACATCAAGCTGAACAAAGCTGAGAAGGCCTGGAAG CCTAGCAGCAAGAGAACCGCCGCCGACAAGGACAGAGGCGAGGAG GACGCCGACGGATCCAAGACCCAGGACCTGTTTCAAGAGTGAGAA GCATCCTCAACAAGCTGACCCCTCAGATGTTCCAGCAGCTGATGAAG CAGGTGACGCAGCTCGCCATCGACACCGAGGAGAGACTGAAGGGCG TGATCGACCTGATCTTTGAGAAGGCTATCTCAGAGCCTAACTTCAGC GTGGCCTACGCCAACATGTGCCGTTGCCTGATGGCATTGAAGGTGCC AACCACCGAGAAGCCTACTGTGACCGTCAATTTCCGTAAACTGCTGC TGAACCGGTGCCAGAAAGAGTTTCGAGAAGGATAAGGACGACGACG AGGTCTTCGAGAAGAAACAGAAAGAAATGGACGAGGCCGCCACCGC AGAGGAAAGGGGCCGATTAAAGGAGGAGCTGGAGGAGGCCAGAGA CATCGCCAGACGGCGTTCTCTGGGCAACATCAAGTTCATAGGTGAGC TGTTCAAGCTAAAGATGCTCACCAGAGGCCATAATGCACGACTGCGTG GTGAAGCTACTGAAGAACCACGACGAAGAAAGCCTGGAGTGCTGT GCAGACTGCTGACCACCATCGGCAAGGACCTGGACTTCGAGAAGGC AAAGCCTCGAATGGACCAGTACTTCAACCAGATGGAGAAGATTATC AAGGAGAAGAAGACCAGCAGCAGAATCAGATTCATGCTGCAGGACG TACTGGACCTGCGCGGAAGCAACTGGGTGCCAAGGAGAGGGGACCA AGGACCAAAGACCATCGACCATCCACAAGGAAGCGGAGATGGA GGAGCACAGAGAGCACATAAAGGTGCAGCAGCTTATGGCCAAGGGC AGCGACAAGCGAAGAGGCGGCCCGCCGGACCTCCTATCAGCAGAG GCCTTCTCTGGTAGACGACGGCGGCTGGAACACCGTGCTATCTCT AAGGGCTCCAGACCTATCGACACTTCCCGTCTTACCAAGATCACCAA GCCAGGATCTATTGAC
22	MBP_eIF4G-mid2 (653-1130) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLDGNIPIPSAIAANSYIYGGGSDPTRLQGINCGPDFTPSFANLG

SEQ ID NO	Sequence information	Sequence
		RTTLSTRGPPRGGPGGELPRGPAGLGPRRSQQGPRKEPRKIIATVLMTED IKLNKAEKAWKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSILNKLT PQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCRC LMALKVPTTEKPTVTNFRKLLLNRQCKEFEKDKDDDEVFEKKQKEM DEAATAEERGRLEKEELEEARDIARRRSLGNIKFIGELFKLKMLETEAIMHD CVVKLLKNHDEESLECLCRLTTIGKDLDFEKAKPRMDQYFNQMEKIIK EKKTSSRIRFMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAEMEEHR EHIKVQQLMAKGSDKRRGGPPGPISRGLPLVDDGGWNTVPISKGSRI DTSRLTKITKPGSIDSNNQLFAPGGRLSWGKGSSGGSGAKPSDAASEAA RPATSTLNRFSALQAV
23	MBP_eIF4G-mid2 (653-1130) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGA CCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCCTT CTTTCGCAAACCTGGGCAGAACCAACCCTGAGCACCAGAGGCCCTCCT AGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCTGG GCCCTAGACGCTCCCAGCAAGGTCTAGAAAAGGAGCCAAGGAAGAT CATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAGCT GAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGACAAG GACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCTGT TCAGAAGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAGATGTT CCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCGAG GAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTATCT CAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTTGCCTG ATGGCATTGAAGGTGCCAACCAACCGAGAAGCCTACTGTGACCGTCA ATTTCCGTAACTGCTGCTGAACCGGTGCCAGAAAGAGTTTCGAGAA GGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAAAT GGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGGAGGA GCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCAAC ATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAGGC CATAATGCACGACTGCGTGTTGAAGCTACTGAAGAACCACGACGAA GAAAGCCTGGAGTGCTGTGCAGACTGCTGACCACCATCGGCAAGG ACCTGGACTTCGAGAAGGCAAAGCCTCGAATGGACCAGTACTTCAA CCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGAGAAT CAGATTGCTGCTGCAGGACGTACTGGACCTGCGCGGAAGCAACTGG GTGCCAAGGAGAGGGGACCAAGGACCAAGAGACCATCGACCAGATCC ACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAGGTGC AGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGCGGCCCGCC CGGACCTCCTATCAGCAGAGGCCTTCTCTGGTAGACGACGGCGGCT GGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACTTCC CGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAACC AGCTGTTCCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGTTC CGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCTGCC AGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAAGC TGTG

SEQ ID NO	Sequence information	Sequence
24	MBP_eIF4G-mid3 (712-1130) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMTIPIFATNSDCELIVK AMQGLLKDGNIPIPSAIAANSIYGGGGSPRKIIATVLMTEDIKLNKA EKA WKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSILNKLTPQMFQQLM KQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCRCLMALKVPT TEKPTVTNVFRKLLLNRQCQKEFEKDKDDDEVFEKKQKEMDEAATAEER GRLKEELEEARDIARRRSLGNIKFIFELFKLKMLTEAIMHDCVVKLLKN HDEESLECLCRLTTIGKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIR FMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAEEMEEHREHIKVVQQL MAKGSDKRRGGPPGISRLPLVDDGGWNTVPISKGSRPIDTSRLTKIT KPGSIDSNNQLFAPGGRLSWGKGSSGSGSAKPSDAASEAARPATSTLNR FSALQQAV
25	MBP_eIF4G-mid3 (712-1130) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCCC AAGGAAGATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTG AACAAAGCTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACCGCC GCCGACAAGGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACC CAGGACCTGTTCAAGAAGAGTGAGAAGCATCCTCAACAAGCTGACCC CTCAGATGTTCCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATC GACACCGAGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGA AGGCTATCTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGC CGTTGCCTGATGGCATTGAAGGTGCCAACACCGAGAAGCCTACTGT GACCGTCAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGT TCGAGAAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGA AAGAAATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAA AGGAGGAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCT GGGCAACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTC ACCGAGGCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACC ACGACGAAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCATCAT CGGCAAGGACCTGGACTTCGAGAAGGCAAGCCTCGAATGGACCAG TACTTCAACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCA GCAGAATCAGATTCATGCTGCAGGACGTACTGGACCTGCGCGGAAG CAACTGGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGA CCAGATCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACAT AAAGGTGCAGCAGCTTATGGCCAAGGGCAGGACGACAAGCGAAGAGG CGGCGCGCCGACCTCCTATCAGCAGAGGCCTTCCTCTGGTAGACG ACGGCGGCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATC GACACTTCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAG CAACAACCAGCTGTTCCGCCCAGGAGGAAGACTTAGCTGGGGCAAG GGAAGTTCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCG AGGCTGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTG CAGCAAGCTGTG
26	MBP_eIF4G-C1 (1080-1599) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMTIPIFATNSDCELIVK AMQGLLKDGNIPIPSAIAANSIYGGGGSSNNQLFAPGGRLSWGKGSSG

SEQ ID NO	Sequence information	Sequence
		GSGAKPSDAASEAARPATSTLNRFSAALQQAVPTESTDNRRVVQRSSLR ERGEKAGDRGDRRLERSERGGDRGDRLDRA RTPATKRSFSKEVEERSRE RPSQPEGLRKAASLTEDRDRGRDAVKREAALPPVSPLKAALSEEELEKK SKAIIIEYLHLNDMKEAVQCVQELASPSLLFIFVRHGVESTLERSAIARE HMGQLLHQLLCAGHLSTAQYYQGLYEILELAEDMEIDIPHVWLYLAEL VTPILQEGGVPMGELFREITKPLRPLGKAASLLEILGLLCKSMGPKKVG TLWREAGLSWKEFLPEGQDIGAFVAEQKVEYTLGEESEAPGQRALPSEE LNRQLEKLLKEGSSNQRVFDWIEANLSEQQIVSNLVRALMTAVCYSAI IFETPLRVDVAVLKARAKLLQKYLCDQKELQALYALQALVVTLEQPP NLLRMFFDALYDEDVVKEDAFYSWESSKDPAEQQGKGVALKSVTAFF KWLREAEESDHN
27	MBP_eIF4G-C1 (1080-1599) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTA ACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCAG CAACAACCAGCTGTTCCGCCAGGAGGAAGACTTAGCTGGGGCAAG GGAAGTTCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCCGCCAGCG AGGCTGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTG CAGCAAGCTGTGCCGACCGAGAGCACCGACAACCGGCGAGTTGTGC AGAGAAGCAGCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATA GAGGTGACCGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCG ACCGGCTTGACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTCAG CAAGGAGGTCGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGA AGGACTCAGAAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGG AAGAGACGCCGTGAAGAGGGAGGCCGACTGCCTCCTGTGAGCCCT CTGAAGGCGGCACTGAGCGAAGAAGAACTTGAGAAGAAGAGTAAG GCGATCATCGAGGAGTACCTGCACCTGAACGACATGAAGGAGGCCG TGCAGTGCGTGCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTC GTCCGGCACGGCGTGGAGTCCACACTGGAAAGATCTGCCATTGCTA GGGAGCATATGGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCAC CTAAGTACTGCCAGTACTATCAGGGTTTATACGAGATCCTCGAACT GGCCGAAGACATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCG CAGAACTGGTGACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGG TGAGTTGTTTCAGAGAAATCACAAAGCCACTGCGCCCACTGGGCAAG GCGGCTAGTCTCCTTCTGGAGATTCTCGGCCCTGCTCTGTAAGAGCAT GGGTCCAAAGAAGGTGGGCACCCTGTGGAGGGAAGCTGGACTCTCC TGGAAGGAATTCCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGC CGAACAGAAAGGTTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCG GGTCAGCGGGCTCTGCCGAGTGAGGAGCTCAACAGACAACCTCGAGA AGCTCCTCAAGGAGGGATCCTCAAATCAGAGAGTGTTGCACTGGAT CGAGGCCAATCTCAGCGAGCAGCAAATCGTGAGCAACACGTTGGTT CGCGCTCTCATGACAGCCGTGTGCTACTCAGCCATTATCTTCGAGAC GCCTCTCCGCGTGGACGTGGCAGTGCTCAAGGCCCGCGCTAAGCTGT TACAGAAGTACCTGTGCGACGAGCAGAAGGAAGTGCAGGCCCTGT CGCCCTGCAAGCTCTGGTCGTCACACTCGAGCAGCCTCCTAACCTGC TGAGAATGTTCTTCGACGCATTGTACGACGAGGACGTTGTGAAGGA AGACGCTTTCTACAGCTGGGAGTCTAGTAAGGATCCTGCGGAACAA

SEQ ID NO	Sequence information	Sequence
		CAAGGCAAGGGCGTTGCTCTAAAGAGCGTGACCGCCTTCTTCAAGTG GCTGCGAGAGGCTGAGGAAGAGTCGGACCACAAC
28	MBP_eIF4G-C2 (1080-1451) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNPIPSAIAANSYGGGGSSNNQLFAPGGRLSWGKSSG GSGAKPSDAASEAARPATSTLNRFSAALQQA VPTSTDNRRVVQRSSLR ERGEKAGDRGDRLESLERGGDRGDRLDRA RPTATKRSFSKEVEERSRE RPSQPEGLRKAASLTEDRDRGRDAVKREAAALPPVSPLKAALSEELEKK SKAIIIEYLHLNDMKEAVQCVQELASPSLLFIFVRHGVSTLERSAIARE HMGQLLHQLLCAGHLSTAQYYQGLYEILELAEDMEIDIPHVWLYLAEL VTPILQEGGVPMGELFREITKPLRPLGKAASLLLEILGLLCKSMGPKKV GTLWREAGLSWKEFLPEGQDIGAFVAEQKVEYTLGEESEAPGQRALPSEE LNRQLEKLL
29	MBP_eIF4G-C2 (1080-1451) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTA ACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAAGTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCAG CAACAACCAGCTGTTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAG GGAAGTTCGGCGGATCCGGCGCCAAGCCTTCGACGCCGCCAGCG AGGCTGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTG CAGCAAGCTGTGCCGACCGAGAGCACCGACAACCGGCGAGTTGTGC AGAGAAGCAGCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATA GAGGTGACCGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCG ACCGGCTTGACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTCAG CAAGGAGGTGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGA AGGACTCAGAAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGG AAGAGACGCCGTGAAGAGGGAGGCCGCACTGCCTCCTGTGAGCCCT CTGAAGGCGGCACTGAGCGAAGAAGAACTTGAGAAGAAGAGTAAG GCGATCATCGAGGAGTACCTGCACCTGAACGACATGAAGGAGGCCG TGCAAGTGCCTGCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTC GTCCGGCACGGCGTGGAGTCCACACTGGAAAGATCTGCCATTGCTA GGGAGCATATGGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCAC CTAAGTACTGCCAGTACTATCAGGGTTTATACGAGATCCTCGAACT GGCCGAAGACATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCG CAGAACTGGTGACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGG TGAGTTGTTTCAGAGAAATCACAAAGCCACTGCGCCCACTGGGCAAG GCGGCTAGTCTCCTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCAT GGGTCCAAAGAAGGTGGGCACCCTGTGGAGGGAAGCTGGACTCTCC TGGAAGGAATTCCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGC CGAACAGAAGGTTGAGTACACCTGGGAGAGGAATCGGAAGCGCCG GGTCAGCGGGCTCTGCCGAGTGAGGAGCTCAACAGACAACTCGAGA AGCTCCTC
30	MBP-CTIF (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNPIPSAIAANSYGGGGSENSSAASASSEAGSSRSQEIEE LERFIDSYVLEYQVQGLLADKTEGDGESERTQSHISQWTADCSEPLDSS

SEQ ID NO	Sequence information	Sequence
		CSFSRGRAPPQQNGSKDNSLDMLGTDIWAANTFDSFSGATWDLQPEKL DFTQFHRKVRHTPKQPLPHIDREGCGKGKLEDGDGINLNDIEKVLP QGYHPMPHEVEIAHTKKLFRRRRNRDRRQQRPPGGNKPPQGHGHDHQP SAKHNRDHQKSYQGGAPHPSGRPTHGYSQNRWRHHGNMKHPPGD KGEAGAHRNAKETMTIENPKLEDTAGDTGHSSLEAPRSPDTLAPVASER LPPQQSGGPEVETKRKDSILPERIGERPKITLLQSSKDRLRRRLKEKDEV AVETTTTPQQNKMDKLEILNSMRNNSSD VDTKL TTFMEEAQNSTNSEE MLGEIVRTIYQKAVSDRSFAFTAAKLCDKMAFMVEGTFKFRSLLL NML QKDFTVREELQQQDVERWLGFI TFLCEVFGTMRSSSTGEPFRVLVCP IYTCLRELLQSQDVKEDAVLCCSMELQSTGRLLLEEQLPEMMTELLASARDK MLCPSESMLTRSLLLEVIELHANSWNPLTPPITQYYNRTIQKLTA
31	MBP-CTIF (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCAGCATGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGA GAACAGTAGCGCCGCTAGCGCGAGCAGCGAAGCCGGGAGCAGCCG GTCTCAGGAGATCGAGGAGCTGGAGCGGTTTCATCGACAGCTACGTG CTGGAGTACCAGGTGCAGGGCCTGCTGGCCGACAAGACCGAGGGCG ACGGCGAGAGCGAGCGGACCCAGAGCCACATCAGCCAGTGGACCGC CGACTGCAGCGAGCCCCTGGACTCTAGCTGCTCATTAGTCGTGGAC GGGCCCCCTCCACAGCAGAACGGCAGCAAGGACAACAGCCTGGACAT GCTGGGCACCGACATCTGGGCCGCCAACACCTTCGACAGCTTCAGCG GCGCCACCTGGGATCTGCAGCCCAGAGAAGCTGGACTTTACCCAGTTC CACCGGAAGGTGCGGCACACTCCCAAGCAGCCCCTGCCCCACATCG ATCGGGAGGGCTGCGGCAAGGGCAAGCTGGAAGACGGCGACGGCA TCAACCTGAACGACATCGAGAAGGTGCTGCCTGCCTGGCAGGGCTA CCACCCCATGCCCCACGAGGTGGAGATCGCCCACACCAAGAAGCTG TTCCGGCGACGACGCAACGACCGGCGTAGGCAGCAACGGCCGCCTG GAGGGAACAAGCCCCAGCAGCACGGAGACCACAGCCCCGGTAGCGC CAAGCACAACCGGGACCACCAGAAGAGCTACCAGGGCGGAAGCGC ACCACACCCCTCGGGCAGACCCACCCACCACGGCTACAGCCAGAAC CGGCGGTGGCATCACGGTAACATGAAGCACCCACCCGGCGACAAAG GAGAGGCCGGCGCTCACCGTAACGCCAAGGAGACCATGACCATCGA GAACCCCAAGCTGGAGGATACCGCCGGCGATACGGGTCACAGCAGC CTGGAGGCACCGCGGTCTCCCGACACCTGGCACCCGTGGCCAGCG AACGGCTGCCACCCCAACAGAGCGGCGCCCTGAGGTTGAGACCAA GCGGAAGGACAGCATCCTGCCCGAACGGATCGGTGAGCGGCCCAAG ATCACCTTACTGCAGAGTAGCAAGGACCGGCTGAGACGGCGGCTGA AGGAGAAGGACGAGGTGGCCGTGGAGACAACCACTCCCCAGCAGA ACAAGATGGACAAGCTGATCGAGATCCTGAACAGCATGCGGAACAA CAGCAGCGACGTGGACACCAAGCTGACCACCTTCATGGAGGAGGCC CAGAACAGCACCAACAGCGAGGAGATGCTGGGCGAGATCGTGCGGA CCATCTACCAGAAGGCCGTGAGCGACCGGAGCTTCGCCTTCACCGCC GCCAAGCTGTGCGACAAGATGGCCCTGTTTCATGGTGGAGGGCACCA AGTTCCGGAGCTTACTGCTGAATATGCTGCAGAAGGACTTCACCGTG CGGGAGGAGCTGCAGCAGCAGGACGTGGAGCGGTGGCTGGGCTTCA TCACCTTCCTGTGCGAGGTGTTCCGGCACCATGCGGAGCAGCACCGGC GAACCCCTCCGGGTGCTGGTGTGCCCCATCTACACCTGCCTGCGGGA

SEQ ID NO	Sequence information	Sequence
		GTTGCTGCAGAGCCAGGACGTGAAGGAGGACGCCGTGCTGTGCTGC AGCATGGAACTGCAGAGCACTGGCCGGCTGCTGGAGGAGCAGCTGC CCGAGATGATGACCGAGCTGCTCGCTAGCGCCCGGGACAAGATGCT GTGCCCCAGCGAGAGCATGCTGACCCGGAGCCTGCTTCTGGAGGTG ATCGAGCTGCACGCCAACAGCTGGAATCCCCTGACCCCTCCCATCAC CCAGTACTACAACCGGACCATCCAGAAGCTGACCGCC
32	MBP--CTIF(379-579) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNIPIPSAIAANSIYGGGSLNSMRNNSDVTDLTTFMEE AQNSTNSEEMLGEIVRTIYQKAVSDRSFAFTAACLCDKMLFMVEGTK FRSLLLNLQKDFTVREELQQQDVERWLGFIITLCEVFGTMRSSSTGEPF RVLVCPITYTCLRELLQSQDVKEDAVLCCSMELQSTGRLLLEEQLPEMMTE LLASARDKMLCPSESMLTRSLLEVIELHANSWN
33	MBP--CTIF(379-579) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCAATTCAGGCATCTACGGAGGCGGTGGAAGCCT GAACAGCATGAGAAACAACAGCAGCGACGTGGACACCAAGCTGACC ACCTTCATGGAGGAGGCCAGAACAGCACCAACAGCGAGGAGATGC TGGGCGAGATCGTGAGAACCATCTACCAGAAGGCCGTGAGCGACAG AAGCTTCGCCTTCACCGCCGCCAAGCTGTGCGACAAGATGGCCCTGT TCATGGTGGAGGGCACCAAGTTCAGAAGCCTGCTGCTGAACATGCT GCAGAAGGACTTCACCGTGAGAGAGGAGCTGCAGCAGCAGGACGTG GAGCGATGGCTGGGCTTCATCACCTTCCTGTGCGAGGTGTTGGGCAC CATGAGAAGCAGCACCGGCGAGCCTTTCAGAGTGCTGGTGTGCCCT ATCTACACCTGCCTGAGAGAGCTGCTGCAGAGCCAGGACGTGAAGG AGGACGCCGTGCTGTGCTGCAGCATGGAGCTGCAGAGCACCGGCAG ACTGCTGGAGGAGCAGCTGCCTGAGATGATGACCGAGCTGCTGGCC AGCGCCAGAGACAAGATGCTGTGCCCTAGCGAGAGCATGCTGACTA GATCCCTACTGTTGGAGGTGATCGAGCTGCACGCCAACAGCTGGAA C
34	MBP- CTIF(365-598) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNIPIPSAIAANSIYGGGSLNSMRNNSDVTDLTTFMEE AQNSTNSEEMLGEIVRTIYQKAVSDRSFAFTAACLCDKMLFMVEGTK FRSLLLNLQKDFTVREELQQQDVERWLGFIITLCEVFGTMRSSSTGEPF RVLVCPITYTCLRELLQSQDVKEDAVLCCSMELQSTGRLLLEEQLPEMMTE LLASARDKMLCPSESMLTRSLLEVIELHANSWN
35	MBP- CTIF(365-598) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC

SEQ ID NO	Sequence information	Sequence
		GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCAC CACCCCTCAGCAGAACAGATGGACAAGCTGATCGAGATCCTGAAC AGCATGAGAAACAACAGCAGCGACGTGGACACCAAGCTGACCACCT TCATGGAGGAGGCCAGAACAGCACCAACAGCGAGGAGATGCTGGG CGAGATCGTGAGAACCATCTACCAGAAGGCCGTGAGCGACAGAAGC TTCGCCTTCACCGCCGCCAAGCTGTGCGACAAGATGGCCCTGTTCAT GGTGGAGGGCACCAAGTTCAGAAGCCTGCTGCTGAACATGCTGCAG AAGGACTTCACCGTGAGAGAGGAGCTGCAGCAGCAGGACGTGGAGA GGTGGCTGGGCTTCATCACCTTCCTGTGCGAGGTGTTCCGGCACCATG AGAAGCAGCACCGGCGAGCCTTTCAGAGTGCTGGTGTGCCCTATCTA CACCTGCCTGAGAGAGCTGCTGCAGAGCCAGGACGTGAAGGAGGAC GCCGTGCTGTGCTGCAGCATGGAGCTGCAGAGCACCGGCAGACTGC TGGAGGAGCAGCTGCCTGAGATGATGACCGAGCTGCTGGCCAGCGC CAGAGACAAGATGCTGTGCCCTAGCGAGAGCATGCTGACGCGCAGC CTCCTGCTGGAGGTGATCGAGCTGCACGCCAACAGCTGGAACCCTCT GACCCCTCCTATCACCCAGTACTACAACAGAACCATCCAGAAGCTGA CCGCC
36	MBP-eIF4G-mid4(752-993) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELIVK AMQGLLKDGNPIPSAIAANSYIYGGGGSADGSKTQDLFRRVRSILNKLT PQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCRC LMALKVPTTEKPTVTNFRKLLLNRCQKEFEKDKDDDEVFEKKQKEM DEATAAEERGRLEEELEEARDIARRRSLGNIKFIGELFKLKMTEAIMHD CVVKLLKNHDEESLECLCRLTTIGKDLDFEKAAPRMDQYFNQMEKIIK EKKTSRIRFMLQDVLDLRGSNWVPR
37	MBP-eIF4G-mid4(752-993) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGC CGACGGCAGCAAGACCCAGGACCTGTTTCAAGAGTGAGAAGCATC CTGAACAAGCTGACCCCTCAGATGTTCCAGCAGCTGATGAAGCAGG TGACCCAGCTGGCCATCGACACCGAGGAGAGACTGAAGGGCGTGAT CGACCTGATCTTCGAGAAGGCCATCAGCGAGCCTAACTTCAGCGTGG CCTACGCCAACATGTGCCGGTGCCTGATGGCCCTGAAGGTGCCTACC ACCGAGAAGCCTACCGTGACCGTGAACCTCAGAAAGCTGCTGCTGA ACCGGTGCCAGAAGGAGTTTCGAGAAGGACAAGGACGACGACGAGG TGTTTCGAGAAGAAGCAGAAGGAGATGGACGAGGCCGCCACCGCCGA GGAGAGAGGCAGACTGAAGGAGGAGCTGGAGGAGGCCAGAGACAT CGCCAGAAGAAGAAGCCTGGGCAACATCAAGTTTATCGGCGAGCTG TTCAAGCTGAAGATGCTGACCGAGGCCATCATGCACGACTGCGTGGT GAAGCTGCTGAAGAACCACGACGAGGAGAGCCTGGAGTGCCTGTGC AGACTGCTGACCACCATCGGCAAGGACCTGGATTTTCGAGAAGGCCGA AGCCTAGAATGGACCAGTACTTCAACCAGATGGAGAAGATCATCAA GGAGAAGAAGACCAGCAGCAGAATCAGATTATGCTGCAGGACGTG CTGGACCTGAGAGGCAGCAACTGGGTGCCTAGA

SEQ ID NO	Sequence information	Sequence
38	MBP-LacZ (control) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNIPIPSAIAANSIGYGGGSSFTLTNKNVIFVAGLGGIGLDT SKELLKRDPVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRP SQQLRSLNGEWRFAWFPAPAEVPESWLECDLPEADTVVVPNSWQMHG YDAPIYTNVTYPITVNPFPVPTENPTGCYSLTFNVDESWLQEGQTRIIFD GVNSAFHLWCNGRWVGYGQDSRLPSEFDLSAFLRAGENRLAVMVLRL WSDGSYLEDQDMWRMSGIFRDVSLHHPKPTTQISDFHVATRNDDFSRA VLEAEVQMCCELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGGYA DRVTLRLNVENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFRE VRIENGLLLLNGKPLLIRGVNRHEHPLHGGVMDEQTMVQDILLMKQN NFNAVRCSHYPNHPLWYTLCDRYGLYVVDANIETHGMVPMNRLTDD PRWLPAMSERVTRMVQRDRNHPSVIIWSLGNESGHGANHDALYRWIKS VDPSPRPVQYEGGGADTTATDIICPMYARVDEDQFPFAPVKWSIKKWLS LPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGGFVWDVVD QSLIKYDENGPNPWSAYGGDFGDTNPDRQFCMNGLVFADRTPHPALTEA KHQQQFFQFRLSGQTIEVTSEYLFRRHSDNELLHWMVALDGGKPLASGEV PLDVAPQKQLIELPELPQPEASAGQLWLTVRVVPNATAWSEAGHISA WQQWRLAENLSVTLPAASHAIPHLTSEMDFCIELGNKRWQFNRQSGF LSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSEATRDPNAWVERWKA AGHYQAEAAALLQCTADTLADAVLITTAHAWQHOGKTLFISRKTYRIDG SGQMAITVDVEVASDTPHPARIGLNCQLAQVAERVNWLGLGPQENYPD RLTAACFDRWDLPLSDMYTPYVFPSENGLRCTRELNYGPHQWRGDF QFNISRYSQQQLMETSHRHLLHAEEGTWLNIDGFHMGIGGDDSWSPSV SAELQLSAGRYHYQLVWCQK
39	MBP-LacZ (control) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCTC GTTACGCTAACGAACAAGAACGTCATCTTCGTAGCGGGACTTGCGG GTATCGGCCTAGACACGTCGAAGGAAGTAAAGCGTGACCCGGT AGTCCTCCAACGTCGCGATTGGGAGAACCCGGGCGTAACGCAACTA AACCGTCTTGCGGCGCACCCGCCGTTTTCGCTCGTGCGGTAACGCGA GGAGGCGCGAACGGATCGTCCGTCGCAACAACTACGTTTCGCTCAAC GGGGAGTGGCGCTTCGCGTGGTTCCCGGCGCCGAGGCGGTACCGG AGTCGTGGCTCGAGTGCATCTACCGGAGGCGGACCGGTCTGCTCGT ACCGTCGAAGTGGCAAATGCACGGTTACGACGCGCCGATATACACG AACGTCACGTACCCGATAACGGTAAACCCGCGTTTCGTCCCGACGG AGAACCCGACGGGGTGCTACTCGCTAACGTTCAACGTTGACGAGTC GTGGTTGCAAGAGGGTCAAACGCGTATCATATTTCGACGGTGTAAGT CGGCGTTCCACCTGTGGTGCAACGGGCGCTGGGTAGGGTACGGCCA AGACTCGCGTCTACCGTCGGAGTTTCGACCTATCGGCGTTCTACGAG CGGGTGAGAACCGGCTAGCGGTCATGGTCTACGTTGGTCGGACGG TTCGTACCTCGAGGACCAAGACATGTGGCGAATGTCGGGTATCTTCC GCGACGTATCGCTCCTACACAAGCCGACGACGCAAATCTCGGACTTC CACGTCGCGACGCGTTTCAACGACGATTTCTCGCGGGCAGTCCTAGA GGCGGAGGTCCAAATGTGCGGGGAGCTACGTGACTACCTCCGTGTC ACGGTATCGCTCTGGCAAGGTGAGACGCAAGTAGCGTCGGGTACGG

SEQ ID NO	Sequence information	Sequence
		CGCCGTTTCGGCGGTGAGATCATCGACGAGCGTGGTGGGTACGCGGA CCGTGTAACGCTACGTCTAAACGTGCGAGAACCCGAAGCTCTGGTCGG CGGAGATCCCGAACCTATAACCGTGCGGTCTGTCGAGCTACATACGGC GGACGGGACGCTAATAGAGGCGGAAGCGTGCGACGTCGGGTTTCGA GAGGTTCGTATAGAGAACGGGCTGCTACTTCTAAACGGGAAGCCGT TGCTCATACGTGGTGTCAACCGTCACGAGCACCACCCGCTACACGGT CAAGTAATGGACGAGCAAACGATGGTACAAGACATCCTACTAATGA AGCAGAACAACTTCAACGCGGTACGCTGTTTCGCATTACCCGAACCAT CCGTTGTGGTACACGCTTTGCGACCGGATACGGTCTATACGTCGTA CGAGGCGAACATAGAGACGCACGGGATGGTACCGATGAATCGCCTA ACGGACGACCCGCGTTGGCTACCGGCGATGTCGGAGCGAGTCACGC GTATGGTCCAACGGGACCGTAACCACCCGTCGGTAATAATCTGGTCG CTAGGCAACGAATCGGGGACGGGGCGAACACGACGCGCTATACC GTTGGATCAAGTCGGTAGACCCGTCGCGTCCGGTACAATACGAAGG TGGCGGTGCGGACACGACGGCGACGGACATCATCTGCCCGATGTAT GCGCGCGTCGACGAAGACCAACCGTTCGCGGCGGTACCGAAGTGGT CGATCAAGAAGTGGCTCTCGTTGCCGGGTGAAACGCGTCCGTTGATA CTTTGCGAGTACGCGCACGCGATGGGCAACTCGTTGGGTGGGTTCCG GAAGTACTGGCAGGCGTTCCGTCAATACCCGCGTCTACAGGGTGGGT TCGTCTGGGACTGGGTAGACCAATCGCTAATCAAGTACGACGAGAA CGGCAACCCGTGGTCGGCGTACGGTGGGGACTTCGGGGACACGCCG AACGACCGCCAATTCTGTATGAACGGCCTAGTCTTCGCGGACCGAAC GCCGCACCCGGCGTTGACGGAGGCGAAGCATCAACAACAATTCTTC CAATTCCGTCTATCGGGGCAAACGATCGAGGTAACGTCGGAGTACTT GTTCCGGCACTCGGACAACGAGCTACTACACTGGATGGTAGCTACTA GACGGCAAGCCGCTAGCGTCGGGAGAAGTCCCTTTGGACGTCGCGC CGCAAGGTAAGCAACTAATCGAGCTACCGGAGCTACCGCAACCGGA GTCGGCGGGTCAACTGTGGTTGACGGTCCGTGTCGTTCAACCGAACG CGACGGCGTGGTCGGAGGCGGGTCACATCTCGGCGTGGCAGCAGTG GCGTCTAGCGGAGAACCTCTCGGTACGCTACCGGCGGCGTTCGCAC GCGATACCGCATCTAACGACGTCCGAGATGGACTTCTGCATCGAGTT GGGGAACAAGAGGTGGCAGTTCAACCGTCAATCGGGATTCTATCG CAAATGTGGATAGGTGACAAGAAGCAACTACTAACGCCGCTACGTG ATCAGTTCACGCGTGCTCCGCTAGACAACGACATAGGTGTTTCGGAG GCGACGCGTATAGACCCGAACGCGTGGGTGGAGCGGTGGAAGGCGG CGGGGCACTACCAAGCGGAGGCGGCGCTACTACAGTGCACGGCGGA CACGCTAGCGGACGCGGTATTGATCACGACGGCGCACGCGTGGCAA CACCAGGGGAAGACGCTATTATCTCGCGTAAGACGTACCGTATCG ACGGTTCGGGCCAAATGGCGATCACGGTCGACGTAGAGGTAGCGTC GGACACGCCGCATCCGGCGCGCATCGGTCTAACTGCCAACTAGCG CAAGTAGCGGAGCGGTGTAACACTGGCTAGGGCTAGGGCCGCAAGAGA ACTATCCGGACCGCCTAACGGCGGCGTGCTTCGACCGTTGGGACCTA CCGCTTTTCGGACATGTATAACCCGCTACGTCTTCCCGTCGGAGAACGG GTTGAGGTGCGGGACGCGCGAGCTAAACTACGGGCCGACCAAGTGG CGAGGGGACTTCCAATTCAACATATCGCGTTACTCGCAACAACAAC AATGGAGACGTCGCACCGTCACCTACTACACGCGGAGGAGGGGACG TGGCTAAACATCGACGGGTTCCACATGGGCATAGGTGGGGACGACT CGTGGTCGCCGTCGGTCTCGGCGGAGCTCCAACTCTCGGCGGGTCGT TACCATTACCAACTAGTTTGGTGCCAGAAG
40	eIF4G-mid2 (aa)	DPTRLQGINCGPDFTPSFANLGRITLSTRGPPRGGPGGELPRGPAGL GPRRSQQGPRKEPRKIIATVLMTEDIKLNKAEKAWKPSSKRTAADKDR GEEDADGSKTQDLFRRVRSILNKLTPQMFQQLMKQVTQLAIDTEERLK

SEQ ID NO	Sequence information	Sequence
		GVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTVTNFRKLLL NRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGRLEKEELEEARDIAR RRSLGNIKFIGELFKLKMLTEAIMHDCVVKLLKNHDEESLECLCRLTTI GKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDLRGSNW VPRRGDQGPKTIDQIHKEAEMEEREHREHIKVQQLMAKGSDKRRGGPPGP PISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKPGSIDSNNQLFAPGG RLSWGKGSSGGSGAKPSDAASEAARPATSTLNRFSALQQAV
41	eIF4G-mid2 (nt)	GACCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCAC TCCTTCTTTTCGCAAACCTGGGCAGAACACCCTGAGCACCAGAGGCC CTCCTAGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTTGCCGG CCTGGGCCCTAGACGCTCCCAGCAAGGTCTAGAAAAGGAGCCAAGG AAGATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACA AAGCTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCGCCGA CAAGGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGA CCTGTTTCAAGAAGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAG ATGTTCCAGCAGCTGATGAAGCAGGTGACGACGATCTTTGAGAAGGC CCGAGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGC TATCTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTT GCCTGATGGCATTGAAGGTGCCAACCCACCGAGAAGCCTACTGTGAC CGTCAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTCCG AGAAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAG AAATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGG AGGAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGG CAACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCG AGGCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGA CGAAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGC AAGGACCTGGACTTCGAGAAGGCAAAGCCTCGAATGGACCAGTACT TCAACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCA GAATCAGATTCATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAA CTGGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCA GATCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAA GGTGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGGCGGC CCGCCCCGACCTCCTATCAGCAGAGGCCTTCCTCTGGTAGACGACGG CGGCTGGAACACCGTGCTATCTCTAAGGGCTCCAGACCTATCGACA CTTCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAAC AACCAGCTGTTGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAA GTTCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGC TGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGC AAGCTGTG
42	eIF4G(dN) 623-1599 (aa)	IFASMQKPEGLPHISDVVLDKANKTPRLPLDPTRLQGINCGPDTFSP ANLGRITTLSTRGPPRGPGGELPRGPAGLGPRRSQQGPRKEPRKIIATVL MTEDIKLNKAEKAWKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSIL NKLTPQMFFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYAN MCRCLMALKVPTTEKPTVTNFRKLLLNRQKEFEKDKDDDEVFEKK QKEMDEAATAEERGRLEKEELEEARDIARRSLGNIKFIGELFKLKMLTE AIMHDCVVKLLKNHDEESLECLCRLTTIGKDLDFEKAKPRMDQYFNQ MEKIIKEKKTSSRIRFMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAE MEEHREHIKVQQLMAKGSDKRRGGPPGPPIISRGLPLVDDGGWNTVPIS KGSRPIDTSRLTKITKPGSIDSNNQLFAPGGRLSWGKGSSGGSGAKPSDA ASEAARPATSTLNRFSALQQAVPTTESTDNRRVVQRSSLRERGEKAGDR

SEQ ID NO	Sequence information	Sequence
		GDRLERSERGGDRGDRLDRARTPATKRSFSKEVEERSRERPSQPEGLRK AASLTEDRDRGRDAVKREAAALPPVSPLKAALSEELEKKSKAIIIEYLHL NDMKEAVQCVQELASPSLLFIFVRHGVESTLERSAIAREHMGQLLHQLL CAGHLSTAQYYQGLYEILELAEDMEIDIPHVWLYLAELVTPILQEGGVP MGELFREITKPLRPLGKAASLLLEILGLLCKSMGPKKVGTLWREAGLSW KEFLPEGQDIGAFVAEQKVEYTLGEESEAPGQRALPSEELNRQLEKLLK EGSSNQRVFDWIEANLSEQQIVSNLVRALMTAVCYSAIIFETPLRVDVA VLKARAKLLQKYLCDQKELQALYALQALVVTLEQPPNLLRMFFDAL YDEDVVKEDAFYSWESSKDPAEQQGKGVALKSVTAFKWLREAEES DHN
43	eIF4G(dN) 623-1599 (nt)	ATATTCGCGAGCATGCAGAAGCCTGAAGGTCTGCCTCACATCAG CGACGTGGTGTCTGGACAAGGCCAACAAGACCCCTCTTAGACCTCTG GACCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCC TTCTTTCGCAAACCTGGGCAGAACCCCTGAGCACCAGAGGCCCTC CTAGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCT GGGCCCTAGACGCTCCCAGCAAGGTCCTAGAAAGGAGCCAAGGAAG ATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAG CTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACCAGCCGCGACAA GGACAGAGGCGAGGAGGACGCGCAGGATCCAAGACCCAGGACCT GTTCAGAAGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAGATG TTCCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCG AGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTAT CTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTTGCC TGATGGCATTGAAGGTGCCAACCCAGAGAAGCCTACTGTGACCGT CAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTCGAG AAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAA ATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGGAG GAGCTGGAGGAGGCCAGAGACATCGCCAGACGCGCTTCTCTGGGCA ACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAG GCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGACG AAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGCAA GGACCTGGACTTCGAGAAGGCAAAGCCTCGAATGGACCAGTACTTC AACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGA ATCAGATTTCATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAACT GGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAGA TCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAGG TGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGGCGGCC GCCCCGACCTCCTATCAGCAGAGGCCTTCTCTGGTAGACGACGGCG GCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACT TCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAA CCAGCTGTTCCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGT TCCGGCGGATCCGGCGCCAAGCCTTCCGACGCGCCGACGAGGCTG CCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAA GCTGTGCCGACCGAGAGCACCGACAACCGCGCAGTTGTGCAGAGAA GCAGCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATAGAGGTG ACCGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCGGC TTGACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTCAGCAAGGA GGTGCAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGACTC AGAAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGGAAGAGAC GCCGTGAAGAGGGAGGCCGCACTGCCTCCTGTGAGCCCTCTGAAGG CGGCACTGAGCGAAGAAGAACTTGAGAAGAAGAGTAAGGCGATCAT CGAGGAGTACCTGCACCTGAACGACATGAAGGAGGCCGTGCAGTGC GTGCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCGGCA

SEQ ID NO	Sequence information	Sequence
		CGGCGTGGAGTCCACACTGGAAAGATCTGCCATTGCTAGGGAGCAT ATGGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAAGTAC TGCCAGTACTATCAGGGTTTATACGAGATCCTCGAACTGGCCGAAG ACATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCGCAGAACTG GTGACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGAGTTGTT CAGAGAAATCACAAAGCCACTGCGCCCACTGGGCAAGGCGGCTAGT CTCCTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGGTCCAAA GAAGGTGGGCACCCTGTGGAGGGAAGCTGGACTCTCCTGGAAGGAA TTCCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGAACAGAA GGTTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGTGAGCGG GCTCTGCCGAGTGAGGAGCTCAACAGACAACTCGAGAAGCTCCTCA AGGAGGGATCCTCAAATCAGAGAGTGTTCGACTGGATCGAGGCCAA TCTCAGCGAGCAGCAAATCGTGAGCAACACGTTGGTTCGCGCTCTCA TGACAGCCGTGTGCTACTCAGCCATTATCTTCGAGACGCCTCTCCGC GTGGACGTGGCAGTGCTCAAGGCCCGCGCTAAGCTGTTACAGAAGT ACCTGTGCGACGAGCAGAAAGGAAGTGCAGGCCCTGTACGCCCTGCA AGCTCTGGTTCGTACACTCGAGCAGCCTCCTAACCTGTGAGAATGT TCTTCGACGATTGTACGACGAGGACGTTGTGAAGGAAGACGCTTTC TACAGCTGGGAGTCTAGTAAGGATCCTGCGGAACAACAAGGCAAGG GCGTTGCTCTAAAGAGCGTGACCGCCTTCTCAAGTGGCTGCGAGAG GCTGAGGAAGAGTCGGACCACAAC
44	eIF4G(dN)- mid_mutant (aa)_3A	IFASMQKPEGLPHISDVVLDKANKTPLRPLDPTRLQGINCGPDFTPSF ANLGRITTLSTRGPPRGPGGELPRGPAGLGPRRSQQGPRKEPRKIIATVL MTEDIKLNKAEKAWKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSIA NKATPQMAQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYAN MCRCLMALKVPTTEKPTVTNFRKLLLNRQCQKEFEKDKDDDEVFEKK QKEMDEAATAEERGRLEKEELEEARDIARRSLGNIKFIGELFKLKMLTE AIMHDCVVKLLKNHDEESLECLCRLTTIGKDLDFEAKAPRMDQYFNQ MEKIIKEKKTSSRIRFMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAE MEEHREHIK VQQLMAKGSDKRRGGPPGPPISRGLPLVDDGGWNTVPIS KGSRPIDTSRLTKITKPGSIDSNNQLFAPGGRLSWGKGSSGSGAKPSDA ASEAARPATSTLNRFSAALQQA VPTESDNRVVRSSLSRERGEKAGDR GDRLESERGGDRGDRLD RARTPATKRSFSKEVEERSRERPSQPEGLRK AASLTEDRDRGRDAVKREAAALPPVSPLKAALSEELEKKS KAIIEEYLHL NDMKEAVQCQELASPSLLFIFVRHGVESTLERSAIAREHMGQLLHQLL CAGHLSTAQYYQGLYEILELAEDMEIDIPHVWLYLAELVTPILQEGGVP MGELFREITKPLRPLGKAASLLLEILGLLCKSMGPKKVGTLWREAGLSW KEFLPEGQDIGAFVAEQKVEYTLGEESEAPGQRALPSEELNRQLEKLLK EGSSNQRVFDWIEANLSEQQIVSNTLVRLMTAVCYS AIIFETPLRVDVA VLKARAKLLQKYL CDEQKELQALYALQALVVTLEQPPNLLRMFFDAL YDEDVVKEDAFYSWESSKDP AEQQGKGVALKS VTAFFKWLREAEES DHN
45	eIF4G(dN)- mid_mutant (nt)_3A	ATATTCGCGAGCATGCAGAAGCCTGAAGGTCTGCCTCACATCAG CGACGTGGTGTGGACAAGGCCAACAAGACCCCTCTTAGACCTCTG GACCCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCC TTCTTTTCGCAAACCTGGGCAGAACCAACCCTGAGCACCAGAGGCCCTC CTAGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCT GGGCCCTAGACGCTCCAGCAAGGTCCTAGAAAGGAGCCAAGGAAG ATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAG CTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACCGCCGCCGACAA GGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCT GTTCAGAAGAGTGAGAAGCATCGCGAACAAGGCGACCCCTCAGATG

SEQ ID NO	Sequence information	Sequence
		GCGCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCG AGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTAT CTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTTGCC TGATGGCATTGAAGGTGCCAACACCGAGAAGCCTACTGTGACCGT CAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTCGAG AAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAA ATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGGAG GAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCA ACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGAG GCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGACG AAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGCAA GGACCTGGACTTCGAGAAGGCCAAAGCCTCGAATGGACCAGTACTTC AACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGA ATCAGATTCATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAACT GGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAGA TCCACAAGGAAGCGGAGATGGAGGAGCACAGAGACACATAAAGG TGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGCGGCCCC GCCCCGACCTCCTATCAGCAGAGGCCCTTCTCTGTTAGACGACGGCG GCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACT TCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAA CCAGCTGTTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGT TCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCTG CCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAA GCTGTGCCGACCGAGAGCACCGACAACCGGCGAGTTGTGCAGAGAA GCAGCCTGAGCAGAGAGAGGGGCGAGAAAGCCGGCGATAGAGGTG ACCGACTGGAGAGAAGCGAGAGAGGAGGTGATAGAGGCGACCGGC TTGACAGAGCCAGAACCCCTGCCACAAAGCGATCGTTTCAGCAAGGA GGTCGAAGAGAGGTCCAGGGAGCGCCCTAGCCAGCCGGAAGGACTC AGAAAGGCAGCCAGTCTAACAGAGGACCGCGACAGGGGAAGAGAC GCCGTGAAGAGGGAGGCCGCACTGCCTCCTGTGAGCCCTCTGAAGG CGGCACTGAGCGAAGAAGAAGCTTGAGAAGAAGAGTAAGGCGATCAT CGAGGAGTACCTGCACCTGAACGACATGAAGGAGGCCGTGCAGTGC GTGCAAGAGCTCGCGAGCCCATCACTGCTGTTTCATCTTCGTCCGGCA CGGCGTGGAGTCCACACTGGAAAGATCTGCCATTGCTAGGGAGCAT ATGGGCCAGTTGTTGCACCAATTGCTTTGCGCCGGCCACCTAAGTAC TGCCCACTACTATCAGGGTTTATACGAGATCCTCGAACTGGCCGAAG ACATGGAGATCGACATCCCTCACGTGTGGCTGTACCTCGCAGAACTG GTGACCCCTATCCTGCAGGAGGGCGGCGTTCCAATGGGTGAGTTGTT CAGAGAAATCACAAGCCACTGCGCCCACTGGGCAAGGCGGCTAGT CTCCTTCTGGAGATTCTCGGCCTGCTCTGTAAGAGCATGGGTCCAAA GAAGGTGGGCACCCTGTGGAGGGAAGCTGGAAGTCTCTGGAAGGAA TTCCTCCCTGAGGGTCAGGACATCGGCGCCTTCGTGGCCGAACAGAA GGTTGAGTACACCCTGGGAGAGGAATCGGAAGCGCCGGGTACGCGG GCTCTGCCGAGTGAGGAGCTCAACAGACAACCTCGAGAAGCTCCTCA AGGAGGGATCCTCAAATCAGAGAGTGTTCGACTGGATCGAGGCCAA TCTCAGCGAGCAGCAAATCGTGAGCAACACGTTGGTTTCGCGCTCTCA TGACAGCCGTGTGCTACTCAGCCATTATCTTCGAGACGCCTCTCCGC GTGGACGTGGCAGTGCTCAAGGCCCGCGCTAAGCTGTTACAGAAGT ACCTGTGCGACGAGCAGAAGGAACTGCAGGCCCTGTACGCCCTGCA AGCTCTGGTCGTACACTCGAGCAGCCTCCTAACCTGCTGAGAATGT TCTTCGACGCATTGTACGACGAGGACGTTGTGAAGGAAGACGCTTTC TACAGCTGGGAGTCTAGTAAGGATCCTGCGGAACAACAAGGCAAGG

SEQ ID NO	Sequence information	Sequence
		GCGTTGCTCTAAAGAGCGTGACCGCCTTCTTCAAGTGGCTGCGAGAG GCTGAGGAAGAGTTCGGACCACAAC
46	eIF4GminMut (aa)	DPTRLQGINCGPDFTPSFANLGRITLSTRGPPRGPGGELPRGPAGL GPRRSQQGPRKEPRKIIATVLMTEDIKLNKAEKAWKPSSKRTAADKDR GEEDADGSKTQDLFRRVRSIANKATPQMAQQLMKQVTQLAIDTEERLK GVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTVTVNFRKLLL NRCQKEFEKDKDDDEVFEKKQKEMDEAATAERGRLEKELEEARDIAR RRSLGNIKFIFELFKLKMLTEAIMHDCVVKLLKNHDEESLECLCRLLTTI GKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDRGSNW VPRRGDQGPKTIDQIHKEAEMEEHREHIKVQQLMAKGSDKRRGGPPGP PISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKPGSIDSNNQLFAPGG RLSWGKGSSGGSGAKPSDAASEAARPATSTLNRFSALQQAV
47	eIF4GminMut (nt)	GACCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCAC TCCTTCTTTTCGAAACCTGGGCAGAACCCCTGAGCACCAGAGGCC CTCCTAGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGG CCTGGGCCCTAGACGCTCCCAGCAAGGTCCTAGAAAGGAGCCAAGG AAGATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACA AAGCTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGCA CAAGGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGA CCTGTTCAGAAGAGTGAGAAGCATCGCCAACAAGGCGACCCCTCAG ATGGCGCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACA CCGAGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGC TATCTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTT GCCTGATGGCATTGAAGGTGCCAACCCAGGAGAAGCCTACTGTGAC CGTCAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTCCG AGAAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAG AAATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGG AGGAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGG CAACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCG AGGCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGA CGAAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGC AAGGACCTGGACTTCGAGAAGGCAAAGCCTCGAATGGACCAGTACT TCAACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCA GAATCAGATTATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAA CTGGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCA GATCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAA GGTGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGGCGGC CCGCCCCGACCTCCTATCAGCAGAGGCCTTCTCTGGTAGACGACGG CGGCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACA CTTCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAAC AACCAGCTGTTCCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAA GTTCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGC TGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGC AAGCTGTG
48	PABP (aa)	NPSAPSYPMASLYVGDLHPDVTEAMLYEKFSPAGPILSIRVCRDMIT RRSLGYAYVNFQQPADAERALDTMNFVVIKGPVRIMWSQRDPSLRKS GVGNIFIKNLDSIDNKALYDTFSAFGNLSCKVVCDENGSKGYGFVHF ETQEAAERAIEKMNGMLLNDRKVFVGRFKSRKEREAELEGARAKEFTNV YIKNFGEDMDDERLKDLEFGKFGPALSVKVMTDESGKSKGFGVFSERH EDAQKAVDEMNGKELNGKQIYVGRAQKKVERQTELKRKFEQMKQDRI TRYQGVNLYVKNLDDGIDDERLRKEFSPFGTITSAKVMMEGGRSKGFG FVCFSSPEEATKAVTEMNGRIVATKPLYVALAQRKEERQAHLTNQYMQ

SEQ ID NO	Sequence information	Sequence
		RMASVRAVPNPVINPYQPAPPSGYFMAAIPQTQNRAAYYPPSQIAQLRP SPRWTAQGARPFPFQNMPGAIRPAAPRPPFSTMRPASSQVPRVMSTQRV ANTSTQTMGPRPAAAAAATPAVRTVPQYKYAAGVRNPQQHLNAQPQ VTMQQPAVHVQGQEPLTASMLASAPPQEQQMLGERLFPLIQAMHPTL AGKITGMLLEIDNSELLHMLESPESLRSKVDEAVAVLQAHQAKEAAQK AVNSATGVPTV
49	PABP (nt)	AACCCAGCGCCCCAAGCTACCCCATGGCCAGCCTGTACGTGGG CGACCTGCACCCCGACGTTACCGAGGCCATGCTGTACGAGAAATTTT CTCCCGCTGGCCCCATCCTGAGCATCCGGGTGTGCCGGGACATGATT ACCCGGCGCTCTCTGGGGTACGCCTACGTGAACTTCCAGCAGCCCGC AGACGCCGAACGGGCCCTGGACACCATGAACTTCGACGTGATCAAG GGCAAGCCCGTGC GGATCATGTGGAGCCAGCGGGACCCAAGCCTGA GAAAGAGCGGCGTGGGCAACATCTTCATCAAGAATCTGGACAAGAG CATCGACAACAAGGCCCTGTACGACACCTTCAGCGCCTTCGGCAATA TCTTAAGCTGCAAGGTGGTGTGCGACGAGAACGGCAGCAAGGGCTA CGGCTTCGTGCACTTCGAAACACAAGAAGCCGCAGAGCGGGCCATC GAGAAGATGAACGGCATGCTGCTGAACGACCGGAAGGTGTTCTGTGG GCCGTTCAAGAGCCGGAAGGAGCGGGAGGCTGAACTCGTGCCCCG GGCCAAGGAGTTCACCAACGTGTACATCAAGAACTTCGGCGAGGAC ATGGACGACGAGCGGTTGAAGGACCTGTTTCGGCAAGTTCGGCCCCG CCCTGTCTGTGAAGGTGATGACCGACGAGAGCGGCAAGTCAAGGG TTTCGGCTTTGTGAGCTTCGAGCGGCACGAGGACGCACAGAAGGCC GTGGACGAAATGAACGGTAAGGAGCTTAACGGTAAGCAGATCTACG TTGGCCGGGCCCCAGAAGAAGGTGGAGCGGCAGACCGAGCTGAAGCG GAAGTTCGAGCAGATGAAGCAGGACCGGATCACCCGGTACCAGGGC GTGAACCTATACGTGAAGAACCTGGACGACGGCATCGACGACGAAC GGCTGCGGAAGGAATTCAGCCCCCTTCGGCACCATCACAGCGCCAA GGTAATGATGGAGGGTGGCCGGTCAAGGGATTTCGGTTTCGTGTGCT TCAGCAGCCCCGAGGAGGCCACCAAGGCCGTTACTGAGATGAACGG CCGGATCGTAGCCACAAAGCCACTTTACGTCGCCCTGGCCCAACGCA AGGAGGAACGGCAGGCCCACCTGACCAACCAGTACATGCAGCGGAT GGCCAGCGTGCGGGCAGTGCCCAACCCCGTGATCAACCCCTACCAA CCCGCAACCACCAAGCGGCTACTTCATGGCCGCCATCCCTCAGACCCA GAACCGGGCCGCCTACTACCCACCCAGCCAGATCGCCAGTTGCGG CCTAGTCCCCGCTGGACGGCTCAAGGAGCACGACCCACCCCTTCCA GAACATGCCCCGAGCCATCAGACCAGCTGCCCCTAGGCCACCCCTTCA GCACCATGCGGCCCGCCTCAAGCCAGGTGCCCCGGGTGATGAGCAC CCAGCGGGTGGCCAAACACCAGCACCCAGACCATGGGCCCTAGACCA GCAGCTGCTGCGGCTGCTGCCACTCCTGCCGTCCGGACCGTGCCACA GTACAAGTACGCCGCCGGCGTTTCGGAACCCACAGCAGCACCTGAAC GCCCAGCCCCAGGTGACCATGCAGCAGCCTGCCGTGCACCTGCAGG GCCAGGAGCCCCGTACCGCCAGCATGTTAGCAAGCGCCCTCCACA GGAGCAGAAGCAGATGCTGGGCGAAAGGCTGTTTCCCCTGATCCAA GCGATGCATCCACCCCTGGCCGGCAAGATCACCGGTATGCTGCTGGA GATCGATAATAGCGAGCTGCTGCACATGCTGGAGAGCCCCGAGAGT CTTAGATCTAAGGTGGACGAGGCCGTGGCCGTGCTCCAGGCTCACCA GGCAAAGGAGGCCGCCAGAAAGCCGTGAACCTCGCCACAGGCGTG CCCACCGTG
50	MBP-PABP (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLDGPNIPSAIAANSYIYGGGGSNPSAPSYPMASLYVGDLHPDV TEAMLYEKFSPAGPILSIRVCRDMITRSLGYAYVNFQQPADAERALDT MNFDVIKGPVRIMWSQRDPSLRKSGVGNIFIKNLDSIDNKALYDTFS

SEQ ID NO	Sequence information	Sequence
		AFGNILSCKVVCDENGSKGYGFVHFETQEAAERAIEKMNGMLLNDRKV FVGRFKSRKEREAELEGARAKEFTNVYIKNFGEDMDDERLKDLFGKF ALSVKVMTDESGKSKGFGFVSFERHEDAQKAVDEMNGKELNGKQIYV GRAQKKVERQTELKRKFEQMKQDRITRYQGVNLYVKNLDDGIDDERL RKEFSPFGTITSAKVMMEGGRSKGFGFVCFSSPEEATKAVTEMNGRIVA TKPLYVALAQRKEERQAHLTNQYMQRMASVRAVPNPVINPYQPAPPSG YFMAAIPQTQNRAAYPPSQIAQLRPSRWTAQGARPHPFQNPMPGAIRP AAPRPPFSTMRPASSQVPRVMSTQRVANTSTQTMGPRPAAAAAATPA VRTVPQYKYAAGVRNPQQHLNAQPQVTMQQPAVHVQGGQEPLTASML ASAPPQEQKQMLGERLFLIQAMHPTLAGKITGMLLEIDNSELLHML PESLRSKVDEAVAVLQAHQAKEAAQKAVNSATGVPTV
51	MBP-PABP (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCAA CCCCAGCGCCCCAAGCTACCCCATGGCCAGCCTGTACGTGGGCGACC TGACCCCCGACGTTACCGAGGCCATGCTGTACGAGAAATTTCTCCC GCTGGCCCCATCCTGAGCATCCGGGTGTGCCGGGACATGATTACCCG GCGCTCTCTGGGGTACGCCTACGTGAACTTCCAGCAGCCCGCAGACG CCGAACGGGGCCCTGGACACCATGAACTTCGACGTGATCAAGGGCAA GCCCCGTGCGGATCATGTGGAGCCAGCGGGACCCAAGCCTGAGAAAG AGCGGCGTGGGCAACATCTTCATCAAGAATCTGGACAAGAGCATCG ACAACAAGGCCCTGTACGACACCTTCAGCGCCTTCGGCAATATCTTA AGCTGCAAGGTGGTGTGCGACGAGAACGGCAGCAAGGGCTACGGCT TCGTGCACTTCGAAACACAAGAAGCCGCAGAGCGGGCCATCGAGAA GATGAACGGCATGCTGCTGAACGACCGGAAGGTGTTCTGTGGGCCGG TTCAAGAGCCGGAAGGAGCGGGAGGCTGAACTCGGTGCCCCGGGCCA AGGAGTTCACCAACGTGTACATCAAGAACTTCGGCGAGGACATGGA CGACGAGCGGTTGAAGGACCTGTTTCGGCAAGTTCGGCCCCGCCCTGT CTGTGAAGGTGATGACCGACGAGAGCGGCAAGTCGAAGGGTTTCGG CTTTGTGAGCTTCGAGCGGCACGAGGACGCACAGAAGGCCGTGGAC GAAATGAACGGTAAGGAGCTTAACGGTAAGCAGATCTACGTTGGCC GGGCCCAGAAGAAGGTGGAGCGGCAGACCGAGCTGAAGCGGAAGT TCGAGCAGATGAAGCAGGACCGGATCACCCGGTACCAGGGCGTGAA CCTATACGTGAAGAACCTGGACGACGGCATCGACGACGAACGGCTG CGGAAGGAATTCAGCCCCCTTCGGCACCATCACCAAGCGCCAAGGTAA TGATGGAGGGTGGCCGGTTCGAAGGGATTCTGTTCTGTGTGCTTCAGC AGCCCCGAGGAGGCCACCAAGGCCGTTACTGAGATGAACGGCCGGA TCGTAGCCACAAAGCCACTTACGTGCGCCTGGCCCAACGCAAGGA GGAACGGCAGGCCACCTGACCAACCAGTACATGCAGCGGATGGCC AGCGTGCGGGCAGTGCCCCAACCCCGTGATCAACCCCTACCAACCCG CACCACCAAGCGGCTACTTCATGGCCGCCATCCCTCAGACCCAGAAC CGGGCCGCTACTACCCACCCAGCCAGATCGCCAGTTGCGGCCTAG TCCCCGCTGGACGGCTCAAGGAGCACGACCCACCCCTTCCAGAAC ATGCCCCGAGCCATCAGACCAGCTGCCCTAGGCCACCCCTCAGCAC CATGCGGCCCCGCTCAAGCCAGGTGCCCCGGGTGATGAGCACCCAG CGGGTGGCCAAACACCAGCACCCAGACCATGGGCCCTAGACCAGCAG CTGCTGCGGCTGCTGCCACTCCTGCCGTCCGGACCGTGCCACAGTAC AAGTACGCCCGCGGCGTTCGGAACCCACAGCAGCACCTGAACGCC

SEQ ID NO	Sequence information	Sequence
		AGCCCCAGGTGACCATGCAGCAGCCTGCCGTGCACGTGCAGGGCCA GGAGCCCCTGACCGCCAGCATGTTAGCAAGCGCCCCTCCACAGGAG CAGAAGCAGATGCTGGGCGAAAGGCTGTTTCCCCTGATCCAAGCGA TGCATCCCACCCTGGCCGGCAAGATCACCGGTATGCTGCTGGAGATC GATAATAGCGAGCTGCTGCACATGCTGGAGAGCCCCGAGAGTCTTA GATCTAAGGTGGACGAGGCCGTGGCCGTGCTCCAGGCTCACCAGGC AAAGGAGGCCGCCAGAAAGCCGTGAACGCCACAGGCGTGCC ACCGTG
52	TENT4A (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNPIPSAIAANSYGGGGSDPRVAWIQPEQKGPANALWM QIWETSQGVGRGGSGFASYFCLNSPALDTAAAAGAAGRSGGGLPALP AASPPPPGPTAPAALPPALLTALGPAAEGARRLHKSPSLSSSSSSSSNAE SGTESPGCSSSSSSASLGRPGGGRGGAFFNFADGAPSAPGTANGHPGR GPAPAGSPSQHQFHPGRRKREKASTYGLNYLLSGSRAAALSGGGGPG AQAPRPGTPWKSRAVSPGIQGLHEEIDFYNFMSPCPEEAAMRREVVKRI ETVVKDLWPTADVQIFGSFSTGLYLPTSDIDL VVFGKWERPPLQLEQA LRKHNVAEPCSIKVLDAKATVPIIKLTDQETEVKVDISFNMETGVRAEFI KNYMKKYSLLPYLILVLKQFLQRLDNEVFTGGISSYSLILMAISFLQLH PRIDARRADENLGMILLVEFFELYGRNFNYLKTGIRIKEGGAYIAKEEIM KAMTSGYRPSMLCIEDPLLPNDVGRSSYGAMQVKQVFDYAYIVLSHA VSPLARSYPNRDAESTLGRIIKVTQEVIDYRRWIKKWSKAHPSGMD SRIKIKERIAATCNGEQTQNPESPYGQRLTSLSSPQLSSGSSASSVSSL SGSDVDSSTPPCTTPSVYQFSLQAPAPLMAGLPTALPMPGKQPPTTSRT LIMTTNNQTRFTIPPPTLGVAPVPCRQAGVEGTASLKAVHHMSSPAIPSA SPNPLSSPHLYHKQHNGMKLSMKGSHGHTQGGGYSSVSGGVRPPVG NRGHHQYNRTGWRRKHTHTRDSLPSLSR
53	TENT4A (nt)	GATCCCCGGGTGGCCTGGATCCAGCCCAGCAGAAGGGCCCCCG CAACGCACTGTGGATGCAGATCTGGGAGACCTCTCAGGGCGTAGGG AGGGGAGGAAGCGGCTTCGCCAGCTACTTCTGCCTGAACAGCCCCG CTCTGGATACGGCAGCCGAGCAGGAGCCGCTGGGAGAGGCTCTGG AGGCCTGGGACCAGCACTGCCCGCTGCTTACCACCACCTCCAGGCC CTACCGCACCCGAGCCCTTCCACCCGCCCTGCTGACCGCTCTGGGC CCAGCTGCAGAGGGTGCTCGGCGGCTGCACAAGAGCCCCAGCCTGA GCAGCAGCAGTAGCAGCTCCAGCAGCAACGCCGAGAGCGGCACAGA GAGCCCCGGCTGCAGCAGCAGCAGCTCCAGTTCCGCAAGCCTCGGC AGACCAGGCGGTGGCAGAGGCGGTGCATTCTTCAACTTCGCTGACG GAGCGCCAAGCGCTCCAGGGACAGCCAACGGACACCCAGGCCCAAG AGGTCCGGCACCCGCTGGAAGCCCCAGCCAGCACCACTTCCACCCC GGCAGACGGAAGCGGAGAACAAAGGCCAGCACCTACGGCCTGAACT ACCTGCTGAGCGGCTCTCGAGCCGCTGCCCTCTCTGGTGGCGGTGGC CCTGGTGCTCAGGCCCCAAGACCTGGCACCCCGTGGAAGTCACGGG CCTACAGCCCCGGTATCCAGGGCCTGCACGAGGAGATCATCGACTTC TACAACTTCATGAGCCCCTGTCCCGAGGAAGCCGCTATGCGGAGGG AGGTGGTGAAGCGGATCGAGACCGTGGTGAAGGACCTGTGGCCAC CGCAGACGTGCAGATCTTCGGCAGCTTCAGCACCGGCCTGTACCTGC CCACCAGCGACATCGACCTGGTGGTGTTCGGCAAGTGGAACGCCC TCCCCTGCAGCTGCTGGAGCAGGCCCTGCGGAAGCACACGTGGCC GAGCCCTGCAGCATCAAGGTGCTGGACAAGGCCACCGTGCCCATCA TCAAGCTGACCGACCAGGAGACCGAGGTGAAGGTGGACATCAGCTT CAACATGGAGACCGGCGTGCCTGCAGCAGAGTTTCAAGAAGTAC ATGAAGAAGTACAGCCTGCTGCCCTACCTGATCCTGGTGCTGAAGCA GTTCTGCTGCAGCGGGACCTGAACGAGGTGTTACCGGCGGCATCA

SEQ ID NO	Sequence information	Sequence
		GCAGCTACAGCCTGATCCTCATGGCCATCAGCTTCCTGCAGCTGCAC CCGCGGATCGACGCAAGGCGGGCCGACGAGAACCTGGGCATGCTGC TGGTGGAGTTCTTCGAGCTGTACGGCCGGAACCTCAACTACCTGAAG ACCGGCATCCGGATCAAGGAGGGCGGCGCCTACATCGCCAAGGAGG AGATCATGAAGGCCATGACCAGCGGCTACCGGCCCAGCATGCTGTG CATCGAGGACCCTCTGCTGCCCCGGCAACGACGTGGGCCGGAGCAGC TACGGCGCCATGCAGGTGAAGCAGGTGTTCTGACTACGCCTACATCGT GCTGAGCCACGCCGTGAGCCCACTGGCCCCGGAGCTACCCCAACCGG GACGCCGAGAGCACCTGGGCCGGATCATCAAGGTGACCCAGGAGG TGATCGACTACCGGCGGTGGATCAAGGAGAAGTGGGGCAGCAAGGC CCACCCATCTCCCGGCATGGACAGCCGGATCAAGATCAAGGAGCGG ATCGCCACCTGCAACGGCGAGCAGACCCAGAACCGGGAGCCCCGAAA GCCCCCTACGGCCAGCGTCTGACCCTGAGCCTGAGCTCTCCCCAGCTG CTGAGCAGCGGCAGCAGCGCCAGCAGCGTGAGCAGCCTGAGCGGCA GCGACGTGGACAGCGACACCCCAACCCTGCACCACCCCAAGCGTGTA CCAGTTCTCACTGCAGGGCCCCCTGCTCCGCTGATGGCTGGCCTGCCTA CCGCCCTGCCCATGCCACGCGGCAAGCCCCAGCCCAACCCAGCCG GACCCGTGATCATGACCACCAACAACAGACCCGGTTACCATCCAC CACCAACCCTGGGCGTGGCTCCTGTGCCTTGTCTGGCAGGCCGGAGTG GAGGGCACCGCCAGCCTGAAGGCCGTGCACCACATGAGCAGCCAG CCATCCCCAGCGCCAGCCCCAACCCTCTGAGCAGCCACACCTGTAC CACAAGCAGCACAACGGCATGAAGCTGAGCATGAAGGGCAGCCACG GACACACCCAGGGCGGAGGGTACTCAAGCGTTGGGAGCGGAGGGGT ACGGCCTCCAGTGGGCAACCGGGGCCACCACCAGTACAACAGGACC GGCTGGAGACGGAAGAAGCACACCCACACCCGGGACTCTCTGCCCCG TGAGCCTGAGCCGG
54	MBP-TENT4A (aa)	DPRVAWIQPEQKGPANALWMQIWETSQGVGRGSGFASYFCLNSP ALDTAAAAGAAAGRSGGLPALPAASPPPPGPTAPAALPPALLTALGPA AEGARRLHKSPSLSSSSSSSSNAESGTESPGCSSSSSSASLGRPGGGRG GAFFNFADGAPSAPGTANGHPGPRGPAPAGSPSQHQFHPGRRKRENKA STYGLNYLLSGSRAAALSGGGGPGAQAPRPGTPWKSRAYSPIQGLHE EIIDFYNFMSPCPEEAAMRREVVKRIETVVKDLWPTADVQIFGSFSTGLY LPTSDIDL VVFGKWERPPLQLLEQALRKHNVAEPCSIKVLDDKATVPILKL TDQETEVKVDISFNMETGVRAAEFIKNYMKKYSLLPYLILVLKQFLLR DLNEVFTGGISSYSLILMAISFLQLHPRIDARRADENLGMILLVEFFELYG RNFNYLKTGIRIKEGGAYIAKEEIMKAMTSGYRPSMLCIEDPLLPNDV GRSSYGAMQVKQVFDYAYIVLSHAVSPLARSYPNRDAESTLGRIIKVTQ EVIDYRRWIKKWKSKAHPSPGMDSRIKIKERIA TCNGEQTQNPES YQRLTLSSLSPQLSSGSSASSVSSLGSDVSDSTPPCTTPSVYQFSLQA PAPLMAGLPTALPMPSPGKPQPTTSRTLIMTTNNQTRFTIPPPTLGVA PVP CRQAGVEGTASLKA VHHMSSPAIPSPNPPLSSPHLYHKQHNGMKLSM KGSHTGHTQGGGYSSVSGGVRPPVGNRGHHQYNRTGWRRKKHHTHR DSLPSLSR
55	MBP-TENT4A (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGA TCCCCGGGTGGCCTGGATCCAGCCCGAGCAGAAGGGCCCCGCCAAC GCACTGTGGATGCAGATCTGGGAGACCTCTCAGGGCGTAGGGAGGG

SEQ ID NO	Sequence information	Sequence
		GAGGAAGCGGCTTCGCCAGCTACTTCTGCCTGAACAGCCCCGCTCTG GATACGGCAGCCGCAGCAGGAGCCGCTGGGAGAGGCTCTGGAGGCC TGGGACCAGCACTGCCCGCTGCTTCACCACCACCTCCAGGCCCTACC GCACCCGCAGCCCTTCCACCCGCCCTGCTGACCGCTCTGGGCCCAGC TGCAGAGGGTGCTCGGCGGCTGCACAAGAGCCCCAGCCTGAGCAGC AGCAGTAGCAGCTCCAGCAGCAACGCCGAGAGCGGCACAGAGAGCC CCGGCTGCAGCAGCAGCAGCTCCAGTTCCGCAAGCCTCGGCAGACC AGGCGGTGGCAGAGGCGGTGCATTCTTCAACTTCGCTGACGGAGCG CCAAGCGCTCCAGGGACAGCCAACGGACACCCAGGCCCAAGAGGTC CGGCACCCGCTGGAAGCCCCAGCCAGCACCAGTTCCACCCCGGCAG ACGGAAGCGGGAGAACAAGGCCAGCACCTACGGCCTGAACTACCTG CTGAGCGGCTCTCGAGCCGCTGCCCTCTCTGGTGGCGGTGGCCCTGG TGCTCAGGCCCCAAGACCTGGCACCCCGTGGAAAGTCACGGGCCTAC AGCCCCGGTATCCAGGGCCTGCACGAGGAGATCATCGACTTCTACA ACTTCATGAGCCCCCTGTCCCAGGGAAGCCGCCATGCGGAGGGAGGT GGTGAAGCGGATCGAGACCGTGGTGAAGGACCTGTGGCCCCACGCCA GACGTGCAGATCTTCGGCAGCTTCAGCACCGGCCTGTACCTGCCCCAC CAGCGACATCGACCTGGTGGTGTTCGGCAAGTGGGAACGCCCTCCCC TGCAGCTGCTGGAGCAGGCCCTGCGGAAGCACAACTGGCCGAGCC CTGCAGCATCAAGGTGCTGGACAAGGCCACCGTGCCCATCATCAAG CTGACCGACCAGGAGACCGAGGTGAAGGTGGACATCAGCTTCAACA TGGAGACCGGCGTGCCTGCAGCAGAGTTCATCAAGAACTACATGAA GAAGTACAGCCTGCTGCCCTACCTGATCCTGGTGTGAAGCAGTTCC TGCTGCAGCGGGACCTGAACGAGGTGTTACCGGCGGCATCAGCAG CTACAGCCTGATCCTCATGGCCATCAGCTTCTGAGCTGCACCCGC GGATCGACGCAAGGCGGGCCGACGAGAACCTGGGCATGCTGCTGGT GGAGTTCTTCGAGCTGTACGGCCGGAACCTCAACTACCTGAAGACCG GCATCCGGATCAAGGAGGGCGGCGCCTACATCGCCAAGGAGGAGAT CATGAAGGCCATGACCAGCGGCTACCGGCCCAGCATGCTGTGCATC GAGGACCCTCTGCTGCCCGCAACGACGTGGGCGGAGCAGCTACG GCGCCATGCAGGTGAAGCAGGTGTTGACTACGCCTACATCGTGCTG AGCCACGCCGTGAGCCCACTGGCCCGGAGCTACCCCAACCGGGACG CCGAGAGCACCCCTGGGCCGATCATCAAGGTGACCCAGGAGGTGAT CGACTACCGGCGGTGGATCAAGGAGAAGTGGGGCAGCAAGGCCAC CCATCTCCCGGCATGGACAGCCGGATCAAGATCAAGGAGCGGATCG CCACCTGCAACGGCGAGCAGACCCAGAACCAGGAGCCCGAAAGCCC CTACGGCCAGCGTCTGACCCTGAGCCTGAGCTCTCCCCAGCTGCTGA GCAGCGGCAGCAGCGCCAGCAGCGTGAGCAGCCTGAGCGGCAGCGA CGTGGACAGCGACACCCACCCCTGCACCACCCCAAGCGTGTACCAG TTCTCACTGCAGGCCCCCTGCTCCGCTGATGGCTGGCCTGCCTACCG CCTGCCCCATGCCCCAGCGGCAAGCCCCAGCCACACCGCCGACC CTGATCATGACCACCAACAACCAGACCCGGTTCACCATCCACCACC AACCCCTGGGCGTGGCTCCTGTGCCTTGTGCGCAGGCCGGAGTGAG GGCACCGCCAGCCTGAAGGCCGTGCACCACATGAGCAGCCAGCCA TCCCCAGCGCCAGCCCCAACCCCTCTGAGCAGCCACACCTGTACCAC AAGCAGCACAAACGGCATGAAGCTGAGCATGAAGGGCAGCCACGGA CACACCCAGGGCGGAGGGTACTCAAGCGTTGGGAGCGGAGGGGTAC GGCCTCCAGTGGGCAACCGGGGCCACCACAGTACAACAGGACCGG CTGGAGACGGAAGAAGCACACCCACACCCGGGACTCTTGCCCGTG AGCCTGAGCCGG
56	TENT4B (aa)	DPRIAWFQPEQLGPSNSLWMQIWETTQGLRNLYFNHHCHSSGGAS GGGGSSSSSTATGSGSSSTGSPGGAASAPAPAPAGMYRSGERLLGSHA LPAEQRDFLPLETTNNNNNHHPGAWARRAGSSASSPPSASSSPHPSAA

SEQ ID NO	Sequence information	Sequence
		VPAADPADSASGSSNKRKRDNKASTYGLNYSLLQPSGGRAAGGGRAD GGGVVYSGTPWKRRNYNQGVVGLHEEISDFYEYMSPRPEEEKMRMEV VNRIESVIKELWPSADVQIFGSFKTGLYLPTSDIDL VVFGKWENLPLWTL EEALRKHKVADEDSVKVLDKATVPIIKLTDSFTEVKVDISFNVQNGVRA ADLIKDFTKKYPVLPYLVLVLKQFLLQRDLNEVFTGGIGSYSLFLMAVS FLQLHPREDACIPNTNYGVLLIEFFELYGRHFNYLKTGIRIKDGGSYVAK DEVQKNMLDGYRPSMLYIEDPLQPGNDVGRSSYGAMQVKQAFDYAY VVLSHAVSPIAKYYPNNETESILGRIIRVTDEVATYRDWISKQWGLKNR PEPSCNGPVSSSSATQSSSSDVSDATPCKTPKQLLCPSTGNRVGSQDV SLESSQAVGKMQSTQTTNTSNSTNKSQHGHSARLFRSSSKGFQGTQTSH GSLMTNKQHOGKSNNQYYHGKKRKHKRDAPLSDLCR
57	TENT4B (nt)	GATCCCCGGATCGCCTGGTTCCAGCCGAGCAGCTAGGCCCAAG CAACAGCCTGTGGATGCAGATCTGGGAGACCACCCAGGGCCTGCGG AACCTGTACTTCAACCACCACTGTACAGCTCTGGCGGCGCAAGTGG CGGAGGCGGCAGTAGCAGCAGCTCTTCAACAGCAACCGGAGGCAGC GGCTCGTCCACCGGCAGTCCGGGTGGAGCCGCCAGTGCCCCAGCTCC TGCCCCAGCAGGCATGTACCGGAGCGGCGAACCGCTGTGGGCAGC CACGCTCTGCCTGCCGAGCAGCGGGACTTCTTGCCCTGGAGACCAC CAATAACAACAACAATCACCAACCAACCGGGGCTTGGGCTCGACGA GCCGGCTCCAGTGCAAGCAGTCCACCTAGCGCCAGCAGCTCTCCACA CCCCAGCGCCGCTGTTCCAGCAGCCGATCCCGCCGACAGTGCCAGCG GCAGCAGCAACAAGCGGAAGCGGGACAACAAGGCCAGCACCTACG GCCTGAACTACAGCCTGCTGCAACCTTCTGGAGGCCGTGCAGCTGGC GGAGGTAGAGCCGACGGTGGCGGTGTGGTGTACAGCGGCACGCCCT GGAAGCGGCGGAATAACAACCAGGGCGTGGTGGCCTGCACGAGGA GATCAGCGACTTCTACGAGTACATGAGCCCACGGCCCGAAGAGGAG AAGATGCGGATGGAGGTGGTGAACCGGATCGAGAGCGTGATCAAGG AGCTGTGGCCTAGCGCCGACGTGCAGATCTTCGGCAGCTTCAAGACC GGCCTGTACCTGCCCACCAGCGACATCGACCTGGTGGTGTTCGGCAA GTGGGAGAACCTGCCCCTGTGGACCCTGGAGGAGGCCCTGCGGAAA CACAAGGTGGCCGACGAGGACAGCGTGAAGGTGCTGGACAAGGCCA CCGTGCCCATCATCAAGCTGACCGACAGCTTACCGAGGTGAAGGT GGACATCAGCTTCAACGTGCAGAACGGCGTGAGGGCAGCCGACCTG ATCAAGGACTTCACCAAGAAGTACCCCGTGCTGCCCTACCTGGTGCT GGTGCTGAAGCAGTTCTGCTGCAGCGGGACCTGAACGAGGTGTTC ACCGGCGGCATCGGCAGCTACAGCCTGTTCTGATGGCCGTGAGCTT CCTGCAGCTGCATCCCCGTGAGGACGCCTGCATCCCCAACACCAACT ACGGCGTGCTGCTGATCGAGTTCTTCGAGCTGTACGGCCGGCACTTC AACTACCTGAAGACCGGCATCCGGATCAAGGACGGCGGCAGCTACG TGGCCAAGGACGAGGTGCAGAAGAACATGCTGGACGGCTACCGGCC CAGCATGCTGTACATCGAGGACCCACTGCAGCCCCGGAACGACGTG GGCCGAGCAGCTACGGCGCCATGCAGGTGAAGCAGGCCTTCGACT ACGCCTACGTGGTGCTGAGCCACGCCGTGAGCCCCATCGCCAAGTAC TACCCCAACAACGAGACCGAGAGCATCCTGGGCGCGATCATCCGGG TGACCGACGAGGTGGCCACCTACCGGGACTGGATCAGCAAGCAGTG GGGCTGAAGAACCGGCCCGAGCCTAGCTGCAACGGCCCCGTGAGC TCAAGCAGCGCCACCCAGAGCAGCAGCAGCGACGTGGACAGCGACG CCACTCCCTGCAAGACCCCTAAGCAGCTGCTGTGCCGTCCCAGCACC GGCAACCGGGTGGGCAGCCAGGACGTGAGCCTGGAGAGCAGCCAG GCCGTGGGCAAGATGCAGAGCACCCAGACCACCAACACAGCAACA GCACCAACAAGAGCCAACACGGCAGCGCACGGCTGTTCCGGAGCAG CAGCAAGGGCTTCCAGGGCACCAACCAGACCAGCCACGGCAGCCTG ATGACCAACAAGCAGCACCAGGGCAAGAGCAACAACCAAGTACTACC

SEQ ID NO	Sequence information	Sequence
		ACGGCAAGAAGCGGAAGCACAAAGCGGGACGCTCCCCTGAGCGACCTGTGCCGG
58	MBP-TENT4B (aaa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNPIPSAIAANSYIYGGGGSDPRIAWFQPEQLGPSNSLWMQ IWETTQGLRNL YFNHHCHSSGGASGGGGSSSSSSTATGSGSSTGSPGG AASAPAPAPAGMYRSGERLLGSHALPAEQRDFLPLETTNNNNHHQPG AWARRAGSSASSPPSASSSPHPSAAVPAADPADSASGSSNKRKRDNKAS TYGLNYSLLQPSGGRAAGGGRADGGGVVYSGTPWKRRNYNQGVVGL HEEISDFYEYMSRPEEEKMRMEVVNRIESVIKELWPSADVQIFGSFKTG LYLPTSDIDL VVFGK WENLPLWTLEEALRHKVVADEDSVKVLDKATVP IHLTDSFTEVKVDISFNVQNGVRAADLIKDFTKYPVLPYLVLVLKQFL LQRDLNEVFTGGIGSYSLFLMAVSFLQLHPREDACIPNTNYGVLLIEFFE LYGRHFNYLKTGIRIKDGGSYVAKDEVQKNMLDGYRPSMLYIEDPLQP GNDVGRSSYGAMQVKQAFDYAYVVLSHAVSPIAKYYPNNETESILGRII RVTDEVATYRDWISKQWGLKNRPEPSCNGPVSSSSATQSSSDVDSDAT PCKTPKQLLCPSTGNRVGSQDVSLESSQAVGKMQSTQTTNTSNSTNKS QHGSARLFRSSSKGFQGTQTQSHGSLMTNKQHQGKSNNQYYHGKCRK HKRDAPLSDLCR
59	MBP-TENT4B (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGA TCCCCGGATCGCCTGGTTCCAGCCCCGAGCAGCTAGGCCCAAGCAAC AGCCTGTGGATGCAGATCTGGGAGACCACCCAGGGCCTGCGGAACC TGTA CTTC AACCACCACTGTACAGCTCTGGCGGCGCAAGTGGCGGA GGCGGCAGTAGCAGCAGCTCTTCAACAGCAACCGGAGGCAGCGGCT CGTCCACCGGCAGTCCGGGTGGAGCCGCCAGTGCCCCAGCTCCTGCC CCAGCAGGCATGTACCGGAGCGGCGAACGGCTGCTGGGCAGCCACG CTCTGCCTGCCGAGCAGCGGGACTTCCTGCCCCCTGGAGACCACCAAT AACAACAACAATCACCACCAACCCGGGGCTTGGGCTCGACGAGCCG GCTCCAGTGCAAGCAGTCCACCTAGCGCCAGCAGCTCTCCACACCCC AGCGCCGCTGTTCCAGCAGCCGATCCCGCCGACAGTGCCAGCGGCA GCAGCAACAAGCGGAAGCGGGACAACAAGGCCAGCACCTACGGCCT GAACTACAGCCTGCTGCAACCTTCTGGAGGCCGTGCAGCTGGCGGA GGTAGAGCCGACGGTGGCGGTGTGGTGTACAGCGGCACGCCCTGGA AGCGGCGGAAC TACAACCAGGGCGTGGTGGGCCTGCACGAGGAGAT CAGCGACTTCTACGAGTACATGAGCCCACGGCCCCGAAGAGGAGAAG ATGCGGATGGAGGTGGTGAACCGGATCGAGAGCGTGATCAAGGAGC TGTGGCCTAGCGCCGACGTGCAGATCTTCGGCAGCTTCAAGACCGGC CTGTACCTGCCCCACCAGCGACATCGACCTGGTGGTGTTCGGCAAGTG GGAGAACCTGCCCCCTGTGGACCCTGGAGGAGGCCCTGCGGAAACAC AAGGTGGCCGACGAGGACAGCGTGAAGGTGCTGGACAAGGCCACCG TGCCCATCATCAAGCTGACCGACAGCTTACCGAGGTGAAGGTGGA CATCAGCTTCAACGTGCAGAACGGCGTGAGGGCAGCCGACCTGATC AAGGACTTACCAAGAAGTACCCCGTGCTGCCCTACCTGGTGCTGGT GCTGAAGCAGTTCCTGCTGCAGCGGGACCTGAACGAGGTGTTACCC GGCGGCATCGGCAGCTACAGCCTGTTCTGATGGCCGTGAGCTTCCT GCAGCTGCATCCCCGTGAGGACGCCTGCATCCCCAACCAACTACG

SEQ ID NO	Sequence information	Sequence
		GCGTGCTGCTGATCGAGTTCTTCGAGCTGTACGGCCGGCACTTCAAC TACCTGAAGACCGGCATCCGGATCAAGGACGGCGGCAGCTACGTGG CCAAGGACGAGGTGCAGAAGAACATGCTGGACGGCTACCGGCCAG CATGCTGTACATCGAGGACCCACTGCAGCCCGGCAACGACGTGGGC CGGAGCAGCTACGGCGCCATGCAGGTGAAGCAGGCCTTCGACTACG CCTACGTGGTGCTGAGCCACGCCGTGAGCCCCATCGCCAAGTACTAC CCCAACAACGAGACCGAGAGCATCCTGGGCCGGATCATCCGGGTGA CCGACGAGGTGGCCACCTACCGGGACTGGATCAGCAAGCAGTGGGG CCTGAAGAACCGGCCCGAGCCTAGCTGCAACGGCCCCGTGAGCTCA AGCAGCGCCACCCAGAGCAGCAGCAGCAGCGACGTGGACAGCGACGCCA CTCCCTGCAAGACCCCTAAGCAGCTGCTGTGCCGTCCCAGCACCGGC AACCGGGTGGGCAGCCAGGACGTGAGCCTGGAGAGCAGCCAGGCCG TGGGCAAGATGCAGAGCACCCAGACCACCAACACCAGCAACAGCAC CAACAAGAGCCAACACGGCAGCGCACGGCTGTTCCGGAGCAGCAGC AAGGGCTTCCAGGGCACCCAGACCAGCCACGGCAGCCTGATGA CCAACAAGCAGCACCAGGGCAAGAGCAACAACCAGTACTACCACGG CAAGAAGCGGAAGCACAAAGCGGGACGCTCCCCTGAGCGACCTGTGC CGG
60	eIF4G-fl (aa)	NSQPQTRSPFFQRPQIQPPRATIPNSSPSIRPGAQTPTAVYQANQHIM MVNHLMPYPVPQGPQYQIPQYRHSGPPYVGPPQYQPVQPPGPGPFYPG PGPGDFPNAYGTFYPSQPVYQSAPIIVPTQQQPPPAKREKKTIRDPNQ GGKDITEEIMSGGSRNPTPIGRPTSTPTPPQQLPSQVPEHSPVYGTVE SAHLAASTPVTAASDQKQEEKPKPDPVLKSPSPVLRLVLSGEKKEQEQQ TSETTAIVSIAELPLPPSPTTVSSVARSTIAAPTSSALSSQIFTTAIDDRCE LSSPREDTIPIPSLTSCETSDPLPTNENDDDICKKPCSVAPNDIPLVSSTN LINEINGVSEKLSATESIVEIVKQEVLPVLTLEILEENPPEEMKLECIPAPIT PSTVPSFPPTPTPPASPPHTPVIVPAAATTVSSPSAAITVQRVLEEDESIRT CLSEDAKEIQNKIEVEADGQTEEILDSQNLNSRRSPVPAQIAITVPKTKW KPKDRTRTTEEMLEAELELKAEELSIDKVLESEQDKMSQGFHPERDPS DLKKVKAVEENGEEAEVVRNGAESVSEGEIDANSGSTDSSGDGVTFPF KPESWKPTDTEGKKQYDREFLLDFQFMPACIQKPEGLPPISDVVLKIN QPKLPMRITLDPRLPRGPDFTPAFADFGRQTPGGRGVPLNVGSSRSQP GQRREPRKIITVSVKEDVHLKKAENAWKPSQKRDSQADDPENIKTQELF RKVRSILNKLTPQMFMNQLMKQVSGLTVDTEERLKGVIDLVFEKAIDEPS FSVAYANMCRCLVTLKVPMAADKPGNTVNFRKLLNRCQKEFEKDKAD DDVFEKKQKELEAASAPEERTRLHDELEEAKDKARRRSIGNIKFIGELFK LKMLTEAIMHDCVVKLLKNHDEESLECLCRLTTIGKDLDFEAKPRM DQYFNQMEKIVKERKTSSRIRFMLQDVIDLRLCNWVSRRADQGPKTIEQ IHKEAKIEEQEQRKVQQLMTKEKRRPGVQRVDEGGWNTVQGAKNRSR VLDPSKFLKITKPTIDEKIQLVPAQLGSWGKGSSGGAKASETDALRSSA SSLNRFSAALQPPAPSGSTPSTPVEFDSRRITLTSRGSMDGREKNDKPLPSAT ARPNTFMRGGSSKDLLDNQSQEEQRREMLETVKQLTGGVDVERNSTEA ERNKTRESAKPEISAMSAHDKAALSEEELERKSKSIIDEFLHINDFKEAM QCVEELNAQGLLHVFRVGVVESTLERSQITRDHMGQLLYQLVQSEKLS KQDFFKGFSETLELADDMAIDIPHIWLYLAELVTPMLKEGGISMRELIE FSKPLLPGVGRAGVLLSEILHLLCKQMSHKKVGALWREADLSWKDFLPE GEDVHNFLLEQKLDIESDSPCSSEALSKKELSAEELYKRLEKLIIDKAN DEQIFDWVEANLDEIQMSSPTFLRALMTAVCKAAIIADSSTFRVDTAVIK QRPILLKYLDSDTEKELQALYALQASIVKLDQPANLLRMFFDCLYDEE VISEDIFYKWESSKDPAEQNGKGVALKSVTAFFTWLREAESEEDN
61	eIF4G-fl (nt)	AACAGCCAACCCAGACACGGAGCCCCTTCTCCAGCGGCCCA GATCCAGCCACCACGGCAACCATCCCCAACAGCAGCCCCAGCATC CGACCCGGCGCACAGACACCCACCGCCGTGTACCAGGCCAACCAGC

SEQ ID NO	Sequence information	Sequence
		ACATCATGATGGTGAACCACCTGCCCATGCCCTACCCCGTGCCACAA GGCCCTCAGTACTGCATCCCACAGTACCGGCACAGCGGTCTCCCTA CGTGGGGCCGCCACAGCAGTACCCCGTTAGCCTCCTGGACCGGGA CCCTTTTATCCGGGACCTGGACCCGGAGACTTCCCCAACGCCTACGG CACACCCTTCTACCCAGCCAGCCAGTGTACCAGAGCGCGCCCATCA TCGTGCCCACCCAGCAGCAGCCTCCACCCGCCAAGCGGGAGAAGAA GACCATCCGGATCCGGGACCCCAACCAGGGCGGCAAGGACATCACC GAGGAGATCATGAGCGGAGGCGGCAGTCGGAACCCACACCGCCCA TCGGCAGACCCACCAGCACCCCAACTCCACCACAGCAGCTGCCAG CCAGGTGCCCAGCAGACAGCCCCGTGGTGTACGGCACCGTGGAAGC GCCCACCTGGCCGCTAGCACCCCTGTGACCGCCGCCAGCGACCAGA AGCAGGAGGAGAAGCCCAAACCCGACCCCGTGCTGAAGTCTCCAG CCCCGTGCTGCGTCTGGTGTGAGCGGCGAGAAGAAGGAGCAGGAG GGCCAGACCAGCGAGACAACCGCCATCGTGAGCATCGCCGAGCTGC CACTGCCTCCCAGTCCCACCACCGTGAGCAGCGTGCGCCGGAGCACC ATCGCAGCGCCACAAGCAGCGCCCTGAGCAGCAGCCCATCTTCA CCACCGCCATCGACGACCGGTGCGAGCTGAGCAGTCCCCGGGAGGA CACCATCCCCATCCCCAGCCTGACCAGCTGCACCGAGACCAGCGACC CACTGCCCACCAACGAGAACGACGACGACATCTGCAAGAAGCCCTG CAGCGTGGCACCCAAACGACATCCCGCTGGTGAGCAGCACCAACCTG ATCAACGAGATCAACGGCGTTAGCGAGAACTGTCTGCCACCGAGA GCATCGTGGAGATCGTCAAGCAGGAGGTGCTGCCCTGACCCTGGA GCTAGAGATCCTGGAGAACCCTCCAGAGGAGATGAAGCTGGAGTGC ATCCCCGCCCAATCACCCCGAGCACCGTGCCAGCTTTCTCCAAC CCCACCCACCCACCAGCCAGTCCACCTCACACTCCCCTGATCGTGC CTGCCGCCGCTACCACCGTGTCTAGCCCCAGCGCCGCCATCACCGTC CAGCGGGTACTGGAGGAGGACGAGAGCATCCGGACCTGCCTGAGCG AAGACGCCAAGGAGATCCAGAACAAGATCGAGGTGGAGGCCGACG GCCAGACCGAGGAAATCCTGGACAGCCAGAACCTGAATTCCCGGCG AAGCCCTGTGCCCCGACAGATCGCCATCACAGTGCCCAAGACCTGG AAGAAGCCCAAGGACCGGACCCGGACCACAGAGGAAATGCTGGAG GCCGAGCTGGAGCTGAAGGCCGAGGAAGAGCTGAGCATCGACAAG GTGCTGGAGAGCGAGCAGGACAAGATGAGCCAGGGCTTCCACCCTG AGCGGGACCCAGCGACCTGAAGAAGGTGAAGGCCGTGGAGGAGA ACGGCGAGGAGGCAGAGCCTGTGCGGAACGGGGCCGAGAGCGTGA GCGAGGGCGAGGGCATCGACGCCAACAGCGGCAGCACTGACAGCA GCGGCGACGGCGTGACCTTCCCCTTCAAGCCCCAAAGCTGGAAGCC CACCGACACCGAAGGAAAGAAGCAGTACGACCGGGAGTTCTGTCTG GACTTCCAGTTCATGCCCCGCTGCATCCAGAAACCCGAGGGGCTGCC ACCAATCAGCGACGTGGTGTGCTGGACAAGATCAACCAGCCCAAGCTT CCCATGCGGACCCCTGGATCCTCGGATCCTCCAGAGGCCCCGACTT CACTCCCGCCTTCGCCGACTTCGGCAGGCAGACCCAGGTTGGTGGG GTGTGCCCTGCTGAACGTGGGCAGCCGTGCGAGCCAGCCTGGCCA ACGGCGGGAACCCCGGAAGATCATCACCGTGAGCGTGAAGGAGGAC GTGCATCTGAAGAAGGCCGAGAACGCCTGGAAACCCAGCCAGAAGC GGGACAGCCAAGCCGACGACCCCGAGAACATCAAGACCCAGGAGCT GTTCCGGAAGGTTGCGAGCATCCTGAACAAGCTGACACCCAGATG TTCAACCAGCTGATGAAGCAGGTGAGCGGCCTGACCGTGGACACCG AGGAGCGGCTGAAGGGCGTGATCGACCTGGTGTGAGAAAGCCAT TGACGAGCCCAGCTTCAGCGTGGCCTACGCCAACATGTGCCGGTGCC TGGTGACCCTGAAGGTGCCCATGGCCGACAAGCCCGGCAACACCGT GAACTTCCGGAAGCTGCTGCTGAACCGGTGCCAGAAGGAGTTCGAG AAGGACAAGGCAGACGACGACGTGTTGAGAAGAAGCAGAAAGAG

SEQ ID NO	Sequence information	Sequence
		TTAGAGGCCGCAAGCGCCCCTGAGGAGCGGACCCGGCTGCACGACG AGCTGGAGGAGGCCAAAGACAAGGCCCGGCGGAGGAGCATCGGCA ACATCAAGTTCATCGGCGAGCTGTTCAAGCTGAAGATGCTGACCGA GGCCATCATGCACGACTGCGTGGTGAAGCTGCTGAAGAACCACGAC GAGGAGAGCCTGGAGTGCCTGTGCCGGCTGCTGACCACCATCGGCA AGGACCTGGACTTCGAGAAGGCCAAGCCCCGGATGGACCAGTACTT CAACCAGATGGAGAAGATCGTGAAGGAGCGGAAGACCAGCAGCCG GATCCGGTTCATGCTGCAGGACGTTATCGACCTGCGGGCTGTGCAACT GGGTGAGCCGGAGGGCTGACCAGGGCCCCAAGACCATCGAGCAGAT CCACAAGGAGGCTAAGATCGAGGAGCAGGAAGAGCAGCGGAAGGT GCAGCAGCTGATGACCAAAGAGAAGCGGAGGGCTGGCGTGACGCA GTGGACGAGGGCGGCTGGAACACCGTGCAGGGCGCCAAGAACAGCC GGGTGCTGGACCCAGCAAGTTCCTGAAGATCACCAAGCCCACCAT CGACGAGAAGATCCAGCTGGTGCCCAAGGCCAGCTCGGCAGCTGG GGCAAGGGCAGCTCAGGTGGCGCCAAGGCCAGCGAGACCGACGCCC TGAGAAGCAGCGCCAGCAGCCTGAACCGGTTCTCTGCCCTGCAGCCT CCTGCCCCTAGCGGCTCCACCCCAAGCACACCCGTGGAGTTCGACAG CCGGCGGACCTTGACGACCGGGGCTCCATGGGTGCGGAGAAGAAC GACAAGCCCTTACCCAGCGCAACCGCTCGGCCCAACACCTTCATGCG GGGCGGCAGCTCCAAGGACCTGCTGGACAACCAGAGCCAGGAGGAG CAGCGGCGGGAGATGCTGGAGACCGTGAAGCAGCTGACCGGAGGCG TGGACGTGGAGCGGAACAGCACCGAGGCCGAGCGGAACAAGACCC GGGAGAGCGCCAAGCCCCGAGATCAGCGCCATGAGCGCCACGACAA GGCCGCCCTGAGCGAGGAGGAGCTGGAACGGAAGAGCAAGAGCAT CATCGACGAGTTCCTGCACATCAACGACTTCAAGGAGGCCATGCAGT GCGTGGAGGAGCTGAACGCCCAGGGCCTGCTGCACGTGTTCTGTGCG GGTGGGCGTGGAGAGCACCTTGAGCGGAGCCAGATCACCCGGGAC CACATGGGCCAGCTGCTGTACCAACTGGTGCAGAGCGAGAAGCTGA GCAAGCAGGACTTCTTCAAGGGCTTCAGCGAGACTCTGGAGCTGGC CGACGACATGGCCATCGACATTCCCCACATCTGGCTGTACCTGGCCG AGCTTGTGACCCCAATGCTGAAGGAGGGCGGCATCAGCATGCGGGA GCTGACCATCGAGTTCAGCAAGCCCCCTCTTGCCAGTCGGCAGAGCAG GAGTGCTGCTGAGCGAGATCCTGCACCTGCTGTGCAAGCAGATGAG CCACAAGAAGGTGGGCGCCCTGTGGAGGGAGGCCGACCTGAGCTGG AAGGACTTCCTGCCCCAGGGCGAAGACGTGCACAACCTTCTTGCTGG AGCAGAAGCTGGACTTCATCGAGAGCGACAGCCCCCTGCAGCAGCGA GGCCCTGAGCAAGAAGGAAGTGAAGCGCCGAGGAGCTGTACAAGCGG CTGGAGAAGCTGATCATCGAGGACAAGGCCAACGACGAGCAGATCT TCGACTGGGTGGAGGCCAACCTGGACGAGATCCAGATGAGCAGCCC CACCTTCCTGCGGGCCCTGATGACCGCCGTGTGCAAGGCCGCCATCA TCGCCGACAGCAGCACCTTCCGGGTGGACACCGCCGTGATCAAGCA GCGGGTGCCCATCCTGCTGAAGTACCTGGACAGCGACACCGAGAAG GAGCTGCAGGCCCTGTACGCCCTGCAGGCCAGCATCGTGAAGCTGG ACCAGCCCGCCAACCTGCTGCGGATGTTCTTCGACTGCCTGTACGAC GAAGAGGTGATCAGCGAGGACGCCTTCTACAAGTGGGAGAGCAGCA AGGACCCCGCCGAGCAGAACGGCAAGGGCGTGGCCCTGAAGAGCGT GACCGCCTTCTTACCTGGCTGCGAGAGGCCGAGGAGGAGAGCGAG GACAAC
62	eIF4G (590-1129) (aa)	GEQKYEYKSDQWKPLNLEEKKRYDREFLLGFQFIFASMQKPEGLPH ISDVVLDKANKTPLRPLDPTRLQGINCPDFTPSFANLGRITLSTRGPPR GGPGGELPRGPAGLGPRRSQQGPRKEPRKIIATVLMTEDIKLNKAEKAW KPSSKRATAADKDRGEEDADGSKTQDLFRRVRSILNKLTPQMFQQLMKQ VTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEK

SEQ ID NO	Sequence information	Sequence
		PTVTVNFRKLLLNRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGR LKEELEEARDIARRRSLGNIKFIGELFKLKMLTEAIMHDCVVKLLKNHD EESLECLCRLTTIGKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIRFM LQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAEMEEHREHIKVQQLMA KGSDKRRGGPPGPPISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKP GSIDSNNQLFAPGGRLSWGKGSSGGSGAKPSDAASEAARPATSTLNRFS ALQQA
63	eIF4G (590-1129) (nt)	GGCGAGCAGAAGTACGAGTACAAGAGCGACCAAGTGGGAAGCCTC TGAACCTGGAGGAGAAGAAGAGATACGACAGAGAGTTCTGCTGGG CTTCCAGTTCATATTCGCGAGCATGCAGAAGCCTGAAGGTCTGCCTC ACATCAGCGACGTGGTGTGGACAAGGCCAACAAGACCCCTCTTAG ACCTCTGGACCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACT TCACTCCTTCTTTCGAAACCTGGGCAGAACCACCTGAGCACCAGA GGCCCTCCTAGAGGTGGTCCCGGCGGAGAAGTCCCCAGGGGTCTG CCGGCCTGGGCCCTAGACGCTCCAGCAAGGTCTAGAAAGGAGCC AAGGAAGATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTG AACAAAGCTGAGAAGGCCTGGAAGCCTAGACGACGAAGAACCAGCC GCCGACAAGGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACC CAGGACCTGTTTTCAGAAGAGTGAGAAGCATCCTCAACAAGCTGACCC CTCAGATGTTCCAGCAGCTGATGAAGCAGGTGACGACGCTCGCCATC GACACCGAGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGA AGGCTATCTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGC CGTTGCCTGATGGCATTGAAGGTGCCAACCACCGAGAAGCCTACTGT GACCGTCAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGT TCGAGAAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGA AAGAAATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAA AGGAGGAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCT GGGCAACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTC ACCGAGGCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACC ACGACGAAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCAT CGGCAAGGACCTGGACTTCGAGAAGGCAAAGCCTCGAATGGACCAG TACTTCAACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCA GCAGAATCAGATTATGCTGCAGGACGTACTGGACCTGCGCGGAAG CAACTGGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGA CCAGATCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACAT AAAGGTGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGG CGGCCCCGCCGACCTCCTATCAGCAGAGGCCCTTCTCTGGTAGACG ACGGCGGCTGGAACACCGTGCTATCTCTAAGGGCTCCAGACCTATC GACACTTCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAG CAACAACCAGCTGTTCCGCCAGGAGGAAGACTTAGCTGGGGCAAG GGAAGTTCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCG AGGCTGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTG CAGCAAGCT
64	eIF4G3 (aa)	NSQPQTRSPFFQRPQIQPPRATIPNSSPSIRPGAQTPTAVYQANQHIM MVNHLPMYPVPQGPQYQIPQYRHSPPYVGPPQYQPVQPPGPGPFYPG PGPGDFPNAYGTFYPSQPQYQSAPIIVPTQQQPPPAKREKKTIRIRDPNQ GGKDITEEIMSGGSRNPTPIGRPTSTPTPPQQLPSQVPEHSPVYGTVE SAHLAASTPVTAASDQKQEEKPKPDPVLKSPSPVLRVLVLSGEKKEQEQQ TSETTAIVSIAELPLPPSPTTVSSVARSTIAAPTSSALSSQIFTTAAIDDRCE LSSPREDTIPIPSLTCTETSDPLPTNENDDDICKKPCSVAPNDIPLVSSTN LINEINGVSEKLSATESIVEIVKQEVLPVLTLEILEENPPEEMKLECIPAPIT PSTVPSFPPTPTPPASPPHTPVIVPAAATTVSSPSAAITVQRVLEEDESIRT CLSEDAKEIQNKIEVEADGQTEEILDSQNLNSRRSPVPAQIAITVPKTKW

SEQ ID NO	Sequence information	Sequence
		KPKDRTRTTEEMLEAELELKAAEELSIDKVLESEQDKMSQGFHPERDPS DLKKVKAVEENGEEAEPVRNGAESVSEGEIDANSSTSSGDGVTFFP KPESWKPTDTEGKKQYDREFLLDFQFMPACIQKPEGLPPISDVVLDKIN QPKLPMRTLDPRLPRGPDFTPAFADFGRQTPGGRGVPLLVGSRRSQP GQRREPRKIITVSVKEDVHLKKAENAWKPSQKRDSQADDPENIKTQELF RKVRSILNKLTPQMFNQLMKQVSGLTVDTEERLKGVIDLVFEKAIDEPS FSVAYANMCRCLVTLKVPMAADKPGNTVNFRKLLLNRCQKEFEKDKAD DDVFEKKQKELEAASAPEERTRLHDELEEAKDKARRRSIGNIKFIGELFK LKMLTEAMHDCVVKLLKNHDEESLECLCRLLTIGKDLDFEAKAPRM DQYFNQMEKIVKERKTSSRIRFMLQDVIDLRLCNWVSRRADQGPKTIEQ IHKEAKIEEQEEQRKVQQLMTKEKRRPGVQRVDEGGWNTVQGAKNRSR VLDPSKFLKITKPTIDEKIQLVPAQLGSWGKGSSGGAKASETDALRSSA SSLNRFSAHQPPAPSGSTPSTPVEFDSRRTLTSRGSMDGREKNDKPLPSAT ARPNTFMRGGSSKDLLDNQSQEEQRREMLETVKQLTGGVDVERNSTE ERNKTRESAKPEISAMSAHDKAALSEEELERKSKSIIDEFLHINDFKEAM QCVEELNAQGLLHVFRVGVESTLERSQITRDHMGQLLYQLVQSEKLS KQDFFKGFSETLELADDMADIPHIWLYLAELVTPMLKEGGISMREL TIEFSKPLLPVGRAGVLLSEILHLLCKQMSHKVGVALWREADLSWKDFLPE GEDVHNFLLEQKLDFIEDSPCSSEALSKKELSAEELYKRLEKLIEDKAN DEQIFDWVEANLDEIQMSSPTFLRALMTAVCKAAIIADSSTFRVDTAVIK QRVPIILLKYLDSDTEKELQALYALQASIVKLDQPANLLRMFFDCLYDEE VISEDIFYKWESSKDPAEQNGKGVALKSVTAFFTWLREAEESSEDN
65	eIF4G3 (nt)	AGCAACAGCCAACCCAGACACGGAGCCCCTTCTTCCAGCGGCC CCAGATCCAGCCACCACGGGCAACCATCCCCAACAGCAGCCCCAGC ATCCGACCCGGCGCACAGACACCCACCGCCGTGTACCAGGCCAACC AGCACATCATGATGGTGAACACCTGCCCATGCCCTACCCCGTGCCA CAAGGCCCTCAGTACTGCATCCCACAGTACCGGCACAGCGGTCTCTC CTACGTGGGGCCGCCACAGCAGTACCCCGTTACGCCTCTGGACCGG GACCCTTTTATCCGGGACCTGGACCCGGAGACTTCCCCAACGCCTAC GGCACACCCTTCTACCCAGCCAGCCAGTGTACCAGAGCGCGCCCAT CATCGTGCCCAACCAGCAGCAGCCTCCACCCGCCAAGCGGGAGAAG AAGACCATCCGGATCCGGGACCCCAACCAGGGCGGCAAGGACATCA CCGAGGAGATCATGAGCGGAGGCGGCAGTCGGAACCCACACCGCC CATCGGCAGACCCACCAGCACCCCAACTCCACCACAGCAGCTGCCC AGCCAGGTGCCCAGACACAGCCCCGTGGTGTACGGCACCGTGGA GCGCCACCTGGCCGCTAGCACCCCTGTGACCGCCGCCAGCGACCA GAAGCAGGAGGAGAAGCCCAACCCGACCCCGTGCTGAAGTCTCCC AGCCCCGTGCTGCGTCTGGTGTGAGCGGCGAGAAGAAGGAGCAGG AGGGCCAGACCAGCGAGACAACCGCCATCGTGAGCATCGCCGAGCT GCCACTGCCTCCCAGTCCCACCACCGTGAGCAGCGTGCCCCGGAGC ACCATCGCAGCGCCACAAGCAGCGCCCTGAGCAGCCAGCCCATCT TCACCACCGCCATCGACGACCGGTGCGAGCTGAGCAGTCCCCGGGA GGACACCATCCCCATCCCCAGCCTGACCAGCTGCACCGAGACACG GACCCACTGCCCACCAACGAGAACGACGACGACATCTGCAAGAAGC CCTGCAGCGTGGCACCCAAACGACATCCCCTGGTGTGAGCAGCACCA CCTGATCAACGAGATCAACGGCGTTAGCGAGAACTGTCTGCCACC GAGAGCATCGTGGAGATCGTCAAGCAGGAGGTGCTGCCCCGTGACC TGGAGCTAGAGATCCTGGAGAACCCTCCAGAGGAGATGAAGCTGGA GTGCATCCCCGCCCAATCACCCCGAGACCGTGCCAGCTTTCTCTC CAACCCACCCACCCACCCAGCCAGTCCACCTCACACTCCCGTGATC GTGCCTGCCGCCGCTACCACCGTGTCTAGCCCCAGCGCCGCCATCAC CGTCCAGCGGGTACTGGAGGAGGACGAGAGCATCCGGACCTGCCTG AGCGAAGACGCCAAGGAGATCCAGAACAAGATCGAGGTGGAGGCC

SEQ ID NO	Sequence information	Sequence
		GACGGCCAGACCGAGGAAATCCTGGACAGCCAGAACCTGAATTCCC GGCGAAGCCCTGTGCCCCGACAGATCGCCATCACAGTGCCCAAGAC CTGGAAGAAGCCCAAGGACCGGACCCGGACCACAGAGGAAATGCTG GAGGCCGAGCTGGAGCTGAAGGCCGAGGAAGAGCTGAGCATCGAC AAGGTGCTGGAGAGCGAGCAGGACAAGATGAGCCAGGGCTTCCACC CTGAGCGGGACCCAGCGACCTGAAGAAGGTGAAGGCCGTGGAGGA GAACGGCGAGGAGGCAGAGCCTGTGCGGAACGGGGCCGAGAGCGT GAGCGAGGGCGAGGGCATCGACGCCAACAGCGGCAGCACTGACAG CAGCGGCGACGGCGTGACCTTCCCCCTTCAAGCCCGAAAGCTGGAAG CCCACCGACACCGAAGGAAAGAAGCAGTACGACCGGGAGTTCTCTGC TGGACTTCCAGTTCATGCCCCGCTGCATCCAGAAACCCGAGGGGCTG CCACCAATCAGCGACGTGGTGCTGGACAAGATCAACCAGCCCAAGC TGCCCATGCGGACCCCTGGATCCTCGGATCCTTCCAGAGGCCCGGAC TTCCTCCCGCCTTCGCGGACTTCGGCAGGCAGACCCAGGTGGTTCG GGGTGTGCCCCTGCTGAACGTGGGCAGCCGTCGGAGCCAGCCTGGC CAACGGCGGGAACCCCGGAAGATCATCACCGTGAGCGTGAAGGAGG ACGTGCATCTGAAGAAGGCCGAGAACGCCTGGAAACCCAGCCAGAA GCGGGACAGCCAAGCCGACGACCCCGAGAACATCAAGACCCAGGA GCTGTTCCGGAAGGTTTCGGAGCATCCTGAACAAGCTGACACCCAG ATGTTCAACCAGCTGATGAAGCAGGTGAGCGGCCTGACCGTGGACA CCGAGGAGCGGCTGAAGGGCGTGATCGACCTGGTGTGTTGAGAAAGC CATTGACGAGCCCAGCTTCAGCGTGGCCTACGCCAACATGTGCCGT GCCTGGTGACCCTGAAGGTGCCCATGGCCGACAAGCCCGGCAACAC CGTGAACCTCCGGAAGCTGCTGCTGAACCGGTGCCAGAAGGAGTTC GAGAAGGACAAGGCAGACGACGACGTGTTTCGAGAAGAAGCAGAAA GAGTTAGAGGCCGCAAGCGCCCCCTGAGGAGCGGACCCGGCTGCACG ACGAGCTGGAGGAGGCCAAAGACAAGGCCCGGCGGAGGAGCATCG GCAACATCAAGTTCATCGGCGAGCTGTTCAAGCTGAAGATGCTGACC GAGGCCATCATGCACGACTGCGTGGTGAAGCTGCTGAAGAACCACG ACGAGGAGAGCCTGGAGTGCCTGTGCCGGCTGCTGACCACCATCGG CAAGGACCTGGACTTCGAGAAGGCCAAGCCCCGGATGGACCAGTAC TTCAACCAGATGGAGAAGATCGTGAAGGAGCGGAAGACCAGCAGCC GGATCCGGTTCATGCTGCAGGACGTTATCGACCTGCGGCTGTGCAAC TGGGTGAGCCGGAGGGCTGACCAGGGCCCCAAGACCATCGAGCAGA TCCACAAGGAGGCTAAGATCGAGGAGCAGGAAGAGCAGCGGAAGG TGCAGCAGCTGATGACCAAGAGAAGCGGAGGCCTGGCGTGCAGCG AGTGGACGAGGGCGGCTGGAACACCGTGACGGGCGCCAAGAACAG CCGGGTGCTGGACCCAGCAAGTTCCTGAAGATCACCAAGCCCCACC ATCGACGAGAAGATCCAGCTGGTGCCCAAGGCCAGCTCGGCAGCT GGGGCAAGGGCAGCTCAGGTGGCGCCAAGGCCAGCGAGACCCGACG CCCTGAGAAGCAGCGCCAGCAGCTGAACCGGTTCTCTGCCCTGCAG CCTCCTGCCCCTAGCGGCTCCACCCCAAGCACACCCGTGGAGTTCGA CAGCCGGCGGACCCTGACCAGCCGGGGCTCCATGGGTGCGGAGAAG AACGACAAGCCCTTACCCAGCGCAACCGCTCGGCCCAACACCTTCAT GCGGGGCGGCAGCTCCAAGGACCTGCTGGACAACCAGAGCCAGGAG GAGCAGCGGCGGGAGATGCTGGAGACCGTGAAGCAGCTGACCGGA GGCGTGGACGTGGAGCGGAACAGCACCGAGGCCGAGCGGAACAAG ACCCGGGAGAGCGCCAAGCCCGAGATCAGCGCCATGAGCGCCACG ACAAGGCCGCCCTGAGCGAGGAGGAGCTGGAACGGAAGAGCAAGA GCATCATCGACGAGTTCCTGCACATCAACGACTTCAAGGAGGCCATG CAGTGCGTGGAGGAGCTGAACGCCAGGGCCTGCTGCACGTGTTTCG TGCGGGTGGGCGTGGAGAGCACCCCTGGAGCGGAGCCAGATACCCG GGACCACATGGGCCAGCTGCTGTACCAACTGGTGCAGAGCGAGAAG

SEQ ID NO	Sequence information	Sequence
		CTGAGCAAGCAGGACTTCTTCAAGGGCTTCAGCGAGACTCTGGAGCT GGCCGACGACATGGCCATCGACATTCCCCACATCTGGCTGTACCTGG CCGAGCTTGTGACCCCAATGCTGAAGGAGGGCGGCATCAGCATGCG GGAGCTGACCATCGAGTTCAGCAAGCCCCTCTTGCCAGTCGGCAGA GCAGGAGTGCTGCTGAGCGAGATCCTGCACCTGCTGTGCAAGCAGA TGAGCCACAAGAAGGTGGGCGCCCTGTGGAGGGAGGCCGACCTGAG CTGGAAGGACTTCTTGCCCGAGGGCGAAGACGTGCACAACCTTCTTG TGGAGCAGAAGCTGGACTTCATCGAGAGCGACAGCCCCTGCAGCAG CGAGGCCCTGAGCAAGAAGGAACCTGAGCGCCGAGGAGCTGTACAAG CGGCTGGAGAAGCTGATCATCGAGGACAAGGCCAACGACGAGCAGA TCTTCGACTGGGTGGAGGCCAACCTGGACGAGATCCAGATGAGCAG CCCCACCTTCTGCGGGCCCTGATGACCGCCGTGTGCAAGGCCGCCA TCATCGCCGACAGCAGCACCTTCCGGGTGGACACCGCCGTGATCAA GCAGCGGGTGGCCATCCTGCTGAAGTACCTGGACAGCGACACCGAG AAGGAGCTGCAGGCCCTGTACGCCCTGCAGGCCAGCATCGTGAAGC TGGACCAGCCCCGCCAACCTGCTGCGGATGTTCTTCGACTGCCTGTAC GACGAAGAGGTGATCAGCGAGGACGCCTTCTACAAGTGGGAGAGCA GCAAGGACCCCGCCGAGCAGAACGGCAAGGGCGTGGCCCTGAAGA GCGTGACCGCCTTCTTCACCTGGCTGCGAGAGGCCGAGGAGGAGAG CGAGGACAAC
66	eIF4G3-mid (aa)	RTLDPRILPRGPDFTPAFADFGRTQTPGGRGVPLLNVGSRRSQPGQRR EPRKIITVSVKEDVHLKKAENAWKPSQKRDSQADDPENIKTQELFRKVR SILNKLTPQMFNQLMKQVSLTVDTTEERLKGVIDLVFEKAIDEPSFSVA YANMCRCLVTLKVPMDKPGNTVNFRLKLLNRCQKEFEKDKADDDVF EKKQKELEAASAPEERTRLHDELEEAKDKARRRSIGNIKFIGELFKLKM LTEAIMHDCVVLLKNHDEESLECLCRLTTIGKDLDFEKAAPRMDQY FNQMEKIVKERKTSSRIRFMLQDVIDLRLCNWVSRRADQGPKTIEQIHK EAKIEEQEEQRKVQQLMTKEKRRPGVQVRVDEGGWNTVQGAKNRVLDP PSKFLKITKPTIDEKIQLVPAQLGSWGKGS
67	eIF4G3-mid (nt)	AGCCGGACCCTGGATCCTCGGATCCTTCCCAGAGGCCCCGACTT CACTCCCGCCTTCGCCGACTTCGGCAGGCAGACCCAGGTGGTCGGG GTGTGCCCTGCTGAACGTGGGCAGCCGTCCGAGCCAGCCTGGCCA ACGGCGGGAACCCCGGAAGATCATCACCGTGAGCGTGAAGGAGGAC GTGCATCTGAAGAAGGCCGAGAACGCCTGGAAACCCAGCCAGAAGC GGGACAGCCAAGCCGACGACCCCGAGAACATCAAGACCCAGGAGCT GTTCCGGAAGGTTCCGAGCATCCTGAACAAGCTGACACCCAGATG TTCAACCAGCTGATGAAGCAGGTGAGCGGCCTGACCGTGGACACCG AGGAGCGGTGAAGGGCGTGATCGACCTGGTGTGTTGAGAAAGCCAT TGACGAGCCCAGCTTCAGCGTGGCCTACGCCAACATGTGCCGGTGCC TGGTGACCCTGAAGGTGCCCATGGCCGACAAGCCCGGCAACACCGT GAACTTCCGGAAGCTGCTGCTGAACCGGTGCCAGAAGGAGTTCGAG AAGGACAAGGCAGACGACGACGTGTTGAGAAGAAGCAGAAAGAG TTAGAGGCCGCAAGCGCCCCTGAGGAGCGGACCCGGCTGCACGACG AGCTGGAGGAGGCCAAAGACAAGGCCCGCGGAGGAGCATCGGCA ACATCAAGTTCATCGGCGAGCTGTTCAAGCTGAAGATGCTGACCGA GGCCATCATGCACGACTGCGTGGTGAAGCTGCTGAAGAACCACGAC GAGGAGAGCCTGGAGTGCCTGTGCCGGCTGCTGACCACCATCGGCA AGGACCTGGACTTCGAGAAGGCCAAGCCCGGATGGACCAGTACTT CAACCAGATGGAGAAGATCGTGAAGGAGCGGAAGACCAGCAGCCG GATCCGGTTCATGCTGCAGGACGTTATCGACCTGCGGCTGTGCAACT GGGTGAGCCGGAGGGCTGACCAGGGCCCCAAGACCATCGAGCAGAT CCACAAGGAGGCTAAGATCGAGGAGCAGGAAGAGCAGCGGAAGGT GCAGCAGCTGATGACCAAAGAGAAGCGGAGGCCTGGCGTGACGCA

SEQ ID NO	Sequence information	Sequence
		GTGGACGAGGGCGGCTGGAACACCGTGCAGGGCGCCAAGAACAGCC GGGTGCTGGACCCCAGCAAGTTCCTGAAGATCACCAAGCCCACCAT CGACGAGAAGATCCAGCTGGTGCCCAAGGCCAGCTCGGCAGCTGG GGCAAGGGCAGCTCAGGT
68	MBP-eIF4G- mid2mut (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNPIPSAIAANSIYGGGGSDPTRLQGINCGPDFTPSFANLG RTTLSTRGPPRGGPGGELPRGPAGLGPRRSQQGPRKEPRKIIATVLMTED IKLNKAEKAWKPSSKRTAADKDRGEEDADGSKTQDLFRRVRSILNKLT PQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCRC LMALKVPTTEKPTVTNFRKLLLNRCQKEFEKDKDDDEVFEKKQKEM DEAATAEERGRLEKEELEEARDIARRRSLGNIAIGELFKLKMTEAIMH DCVVKLLKNHDEESLECLCRLTTIGKDLDFEKAAPRMDQYFNQMEKII KEKKTSSRIRFMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAEMEEH REHIKVQQLMAKGSCKRRGGPPGPISRLPLVDDGGWNTVPISKGSRP IDTSRLTKITKPGSIDSNQLFAPGGRLSWGKGSSGGSGAKPSDAASEAA RPATSTLNRFSALQQAV
69	MBP-eIF4G- mid2mut (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAGCCAGGCCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGA CCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCCTT CTTTCGCAAACCTGGGCAGAACACCCTGAGCACCAGAGGCCCTCTT AGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCTGG GCCCTAGACGCTCCCAGCAAGGTCTAGAAAGGAGCCAAGGAAGAT CATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAGCT GAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACC GCCCGGCAAG GACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCTGT TCAGAAGAGTGAGAAGCATCTCAACAAGCTGACCCCTCAGATGTT CCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCGAG GAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTATCT CAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTTGCCTG ATGGCATTGAAGGTGCCAACCAACCGAGAAGCCTACTGTGACCGTCA ATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTCGAGAA GGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGAAAT GGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGGAGGA GCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCAAC ATCAAGGCGATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCAGG CCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGACGA AGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGCAAG GACCTGGACTTCGAGAAGGCAAGCCTCGAATGGACCAGTACTTCA ACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGAA TCAGATTGATGCTGCAGGACGTAAGTGGACCTGCGCGGAAGCAACTG GGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAGAT CCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAGGT GCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGCGGCCCCG CCCGGACCTCCTATCAGCAGAGGCCTTCTCTGTTAGACGACGGCGG CTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACTT CCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACAAC

SEQ ID NO	Sequence information	Sequence
		CAGCTGTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGTT CCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCTGC CAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAAG CTGTG
70	MBP- eIF4G (590-1129) (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNPIPSAIAANSYIYGGGSGEQKYEYKSDQWKPLNLEEK KRYDREFLLGFQIFASMQKPEGLPHISDVVLDKANKTPLRPLDPTRLQ GINCGPDFTPSFANLGRITLSTRGPPRGPGGELPRGPAGLGPRRSQQGP RKEPRKIIATVLMTEDIKLNKAEKAWKPSSKRTAADKDRGEEDADGSK TQDLFRRVRSILNKLTPQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAI SEPNFSVAYANMCRCLMALKVPTTEKPTVTNFRKLLLNRQCKEFEKD KDDDEVFEKKQKEMDEAATAEERGLKEELEEARIRRRSLGNIKFIG ELFKLKMLTEAIMHDCVVKLLKNHDEESLECLCRLTTIGKDLDFEKAK PRMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDLRGSNWVPRRGDQGP TIDQIHKEAEEMEEHREHIKVQQLMAKGS DKRRGGPPGPPIRGLPLVDD GGWNTVPISKGSRPIDTSRLTKITKPGSIDSNQLFAPGGRLSWGKGSSG GSGAKPSDAASEAARPATSTLNRFSALQQA
71	MBP- eIF4G (590-1129) (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAACCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCGG CGAGCAGAAAGTACGAGTACAAGAGCGACCAAGTGAAGCCTCTGAAC CTGGAGGAGAAGAAGAGATACGACAGAGAGTTCTCTGCTGGGCTTCC AGTTCATATTTCGCGAGCATGCAGAAGCCTGAAGGTCTGCCTCACATC AGCGACGTGGTGTGCTGGACAAGGCCAACAAGACCCCTCTTAGACCTC TGGACCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCACT CCTTCTTTTCGCAAACCTGGGCAGAACCCACCTGAGCACCAGAGGCC TCCTAGAGGTGGTCCCAGGCGGAGAACTCCCCAGGGGTCTGCGCGC CTGGGCCCTAGACGCTCCAGCAAGGTCTAGAAAGGAGCCAAGGA AGATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACAA AGCTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACCGCCGCCGAC AAGGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGAC CTGTTTCAAGAGTGAGAAGCATCTCAACAAGCTGACCCCTCAGA TGTTCCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACAC CGAGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCT ATCTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTTG CCTGATGGCATTGAAGGTGCCAACCACCGAGAAGCCTACTGTGACC GTCAATTTCCGTAAACTGCTGCTGAACCGGTGCCAGAAAGAGTTTCA GAAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAGA AATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGGA GGAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGGC AACATCAAGTTCATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCGA GGCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACGAC GAAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGGCA AGGACCTGGACTTCGAGAAGGCAAAGCCTCGAATGGACCAGTACTT CAACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCAG AATCAGATTGATGCTGCAGGACGTAAGTGGACCTGCGCGGAAGCAAC TGGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCAG

SEQ ID NO	Sequence information	Sequence
		ATCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAAG GTGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGCGGCC CGCCCGGACCTCCTATCAGCAGAGGCCTTCTCTGGTAGACGACGGC GGCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACAC TTCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAACA ACCAGCTGTTTCGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAG TTCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGCT GCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCA AGCT
72	MBP-eIF4G3 (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNIPIPSAIAANSYGGGGNSQPPQTRSPFFQRPQIQPPRATI PNSSPSIRPGAQTPTAVYQANQHIMMVNHLMPYPVPQGPQYQYQYR SGPPYVGPPQQYPVQPPGPGPFYPGPGGDFPNAYGTPFYPSQPVYQSAP IIVPTQQQPPPAKREKKTIRIRDPNQGKIDITEIMSGGSRNPTPIGRPT STPTPPQQLPSQVPEHSPVVGTVESAHLAASPVTAAADQKQEEKPKP DPVLKSPSPVLRLLVLSGEKKEQEGQTSETTAIVSIAELPLPSPTTVSSVA RSTIAAPTSSALSSQPIFTTAIDDRCELSSPREDTIPIPSLTSTCTESDPLPTN ENDDDICKKPCSVAPNDIPLVSSTNLINENGVEKLSATSETIVEIVKQEV LPLTLELEILENPPEEMKLECIAPITPSTVPSFPPTPTPASPPHTPVIVPA AATTVSSPSAAITVQRVLEEDSIRTCLSEDAKEIQNKIEVEADGQTEEIL DSQNLNSRRSPVPAQIAITVPKTKWKPKDRTRTTEEMLEAELELKAEEE LSIDKVLSEQDKMSQGFHPERDPSDLKKVKA VEENGEEAEPVRNGAE SVSEGEIDANSGSTDSSGDGVTFFPKPESWKPTDTEGKKQYDREFLLD FQFMPACIQKPEGLPPISDVVLKINQPKLPMRTLDPRLPRGPDFTPAFA DFGRQTPGGRGVPLLNVGSRSSQPGQRREPRKIITVSVKEDVHLKKAEN AWKPSQKRDSQADDPENIKTQELFRKVR SILNKLTPQMFNQLMKQVSG LTVDTTEERLKGVIDLVFEKAIDEPSFSVAYANMCRCLVTLKVPMAADKP GNTVNFRKLLLNRQCQKEFEKDKADDDVFEKKQKELEAASAPEERTRLH DELEEAKDKARRRSIGNIKFIGELFKLKM LTEAMHDCVVKLLKNHDEE SLECLCRLTTTIGKDLDFEKAKPRMDQYFNQMEKIVKERKTSSRIRFML QDVIDLRLCNWVSRRADQGPKTIEQIHKEAKIEEQEEQRKVQQLMKE KRRPGVQRVDEGGWNTVQGAKNRVLDP SKFLKITKPTIDEKIQLVPKA QLGSWGKGSSGGA KASETDALRSSASSLNRFSA LQPPAPSGSTPSTPVEF DSRRTLTSRSGSMGREKNDKPLPSATARPNTFMRGGSSKDLLDNQSQEE QRREMLETVKQLTGGVDVERNSTEAERNKTRESAKPEISAMSAHDKAA LSEELERKSKSIIDEFLHINDFKEAMQCVEELNAQGLLHVFRVGVST LERSQITRDHMGQLLYQLVQSEKLSKQDFFKGFSETLELADDAIDIPHI WLYLAELVTPMLKEGGISMREL TIEFSKPLLPVGRAGVLLSEILHLLCKQ MSHKKVGALWREADLSWKDFLPEGEDVHNFLLEQKLD FIESDSPCSSE ALSKKELSAEELYKRLEKLIHEDKANDEQIFDWVEANLDEIQMSSPTFLR ALMTAVCKAAIIADSSSTRVDTAVIKQRPILLKYLDSDTEKELQALYA LQASIVKLDQPANLLRMFFDCLYDEEVISED AFYKWESSKDP AEQNGK GVALKSVTAFFTWLREAESEEDN
73	MBP-eIF4G3 (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCCAAGACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCAA CAGCCAACCCAGACACGGAGCCCCCTTCTTCAGCGGCCCCAGATCC

SEQ ID NO	Sequence information	Sequence
		AGCCACCACGGGCAACCATCCCCAACAGCAGCCCCAGCATCCGACC CGGCGCACAGACACCCACCGCCGTGTACCAGGCCAACAGCACATC ATGATGGTGAACCACCTGCCCATGCCCTACCCCGTGCCACAAGGCC TCAGTACTGCATCCCACAGTACCGGCACAGCGGTCCTCCCTACGTGG GGCCGCCACAGCAGTACCCCGTTTACGCCTCCTGGACCGGGACCCTTT TATCCGGGACCTGGACCCGAGACTTCCCCAACGCCTACGGCACACC CTTCTACCCAGCCAGCCAGTGTACCAGAGCGCGCCCATCATCGTGC CCACCCAGCAGCAGCCTCCACCCGCCAAGCGGGAGAAGAAGACCAT CCGGATCCGGGACCCCAACCAGGGCGGCAAGGACATCACCGAGGAG ATCATGAGCGGAGGCGGCAGTCGGAACCCACACCGCCCATCGGCA GACCCACCAGCACCCCAACTCCACCACAGCAGCTGCCAGCCAGGT GCCCCAGCACAGCCCCGTGGTGTACGGCACCGTGGAAGCGCCAC CTGGCCGCTAGCACCCCTGTGACCGCCGCCAGCGACCAGAAGCAGG AGGAGAAGCCCAACCCGACCCCGTGCTGAAGTCTCCAGCCCCGT GCTGCGTCTGGTGTGAGCGGCGAGAAGAAGGAGCAGGAGGGCCAG ACCAGCGAGACAACCGCCATCGTGAGCATCGCCGAGCTGCCACTGC CTCCAGTCCCACCACCGTGAGCAGCGTGGCCGAGCACCATCGC AGCGCCACAAAGCAGCGCCCTGAGCAGCCAGCCATCTTACCACC GCCATCGACGACCGGTGCGAGCTGAGCAGTCCCCGGGAGGACACCA TCCCCATCCCCAGCCTGACCAGCTGCACCGAGACCAGCGACCCACTG CCCACCAACGAGAACGACGACGACATCTGCAAGAAGCCCTGCAGCG TGGCACCCAAACGACATCCCGCTGGTGAGCAGCACCAACCTGATCAA CGAGATCAACGGCGTTAGCGAGAACTGTCTGCCACCGAGAGCATC GTGGAGATCGTCAAGCAGGAGGTGCTGCCCTGACCTGGAGCTAG AGATCCTGGAGAACCCTCCAGAGGAGATGAAGCTGGAGTGCATCCC CGCCCCAATCACCCCGAGCACCGTGCCAGCTTTCCTCCAACCCAC CCACCCACCAGCCAGTCCACCTCACACTCCCGTGATCGTGCCTGCC GCCGCTACCACCGTGTCTAGCCCCAGCGCCGCCATCACCGTCCAGCG GGTACTGGAGGAGGACGAGAGCATCCGACCTGCCTGAGCGAAGAC GCCAAGGAGATCCAGAACAAGATCGAGGTGGAGGCCGACGGCCAG ACCGAGGAAATCCTGGACAGCCAGAACCTGAATTCCCGGCGAAGCC CTGTGCCCCGACAGATCGCCATCACAGTGCCCCAAGACCTGGAAGAA GCCCAAGGACCGGACCCGGACCACAGAGGAAATGCTGGAGGCCGA GCTGGAGCTGAAGGCCGAGGAAGAGCTGAGCATCGACAAGGTGCTG GAGAGCGAGCAGGACAAGATGAGCCAGGGCTTCCACCCTGAGCGGG ACCCAGCGACCTGAAGAAGGTGAAGGCCGTGGAGGAGAACGGCG AGGAGGCAGAGCCTGTGCGGAACGGGGCCGAGAGCGTGAGCGAGG GCGAGGGCATCGACGCCAACAGCGGCAGCACTGACAGCAGCGGCGA CGGCGTGACCTTCCCCTTCAAGCCCCGAAAGCTGGAAGCCACCGAC ACCGAAGGAAAGAAGCAGTACGACCGGAGTTCCTGCTGGACTTCC AGTTCATGCCCCGCTGCATCCAGAAACCCGAGGGGCTGCCACCAATC AGCGACGTGGTGTGAGACAAGATCAACCAGCCCAAGCTGCCCATGC GGACCCTGGATCCTCGGATCCTTCCAGAGGCCCGGACTTCACTCCC GCCTTCGCCGACTTCGGCAGGCAGACCCAGGTGGTGGGGGTGTGCC CCTGCTGAACGTGGGCAGCCGTGCGAGCCAGCCTGGCCAACGGCGG GAACCCCGGAAGATCATCACCGTGAGCGTGAAGGAGGACGTGCATC TGAAGAAGGCCGAGAACGCCTGGAAACCCAGCCAGAAGCGGGACA GCCAAGCCGACGACCCCGAGAACATCAAGACCCAGGAGCTGTTCCG GAAGGTTCCGAGCATCCTGAACAAGCTGACACCCAGATGTTCAAC CAGCTGATGAAGCAGGTGAGCGGCCTGACCGTGACACCGAGGAGC GGCTGAAGGGCGTGATCGACCTGGTGTGTTGAGAAAGCCATTGACGA GCCCAGCTTCAGCGTGGCCTACGCCAACATGTGCCGGTGCTGGTGA CCCTGAAGGTGCCCATGGCCGACAAGCCCGGCAACACCGTGAACCTT

SEQ ID NO	Sequence information	Sequence
		CCGGAAGCTGCTGCTGAACCGGTGCCAGAAGGAGTTCGAGAAGGAC AAGGCAGACGACGACGTGTTTCGAGAAGAAGCAGAAAAGAGTTAGAG GCCGCAAGCGCCCCTGAGGAGCGGACCCGGCTGCACGACGAGCTGG AGGAGGCCAAAGACAAGGCCCGGCGGAGGAGCATCGGCAACATCA AGTTCATCGGCGAGCTGTTCAAGCTGAAGATGCTGACCGAGGCCATC ATGCACGACTGCGTGGTGAAGCTGCTGAAGAACCACGACGAGGAGA GCCTGGAGTGCCTGTGCCGGCTGCTGACCACCATCGGCAAGGACCTG GACTTCGAGAAGGCCAAGCCCCGGATGGACCAGTACTTCAACCAGA TGGAGAAGATCGTGAAGGAGCGGAAGACCAGCAGCCGGATCCGGTT CATGCTGCAGGACGTTATCGACCTGCGGCTGTGCAACTGGGTGAGCC GGAGGGCTGACCAGGGCCCCAAGACCATCGAGCAGATCCACAAGGA GGCTAAGATCGAGGAGCAGGAAGAGCAGCGGAAGGTGCAGCAGCT GATGACCAAAGAGAAGCGGAGGCCTGGCGTGCAGCGAGTGGACGA GGGCGGCTGGAACACCGTGCAGGGCGCCAAGAACAGCCGGGTGCTG GACCCAGCAAGTTCCTGAAGATCACCAAGCCCACCATCGACGAGA AGATCCAGCTGGTGCCCAAGGCCAGCTCGGCAGCTGGGGCAAGGG CAGCTCAGGTGGCGCCAAGGCCAGCGAGACCAGCCCTGAGAAGC AGCGCCAGCAGCCTGAACCGGTTCTCTGCCCTGCAGCCTCTGCCCC TAGCGGCTCCACCCCAAGCACACCCGTGGAGTTCGACAGCCGGCGG ACCCTGACCAGCCGGGGCTCCATGGGTGCGGAGAAGAACGACAAGC CCTTACCCAGCGCAACCGCTCGGCCCCAACACCTTCATGCGGGGCGGC AGCTCCAAGGACCTGCTGGACAACCAGAGCCAGGAGGAGCAGCGGC GGGAGATGCTGGAGACCGTGAAGCAGCTGACCGGAGGCGTGGACGT GGAGCGGAACAGCACCGAGGCCGAGCGGAACAAGACCCGGGAGAG CGCCAAGCCGAGATCAGCGCCATGAGCGCCACGACAAGGCCGCC CTGAGCGAGGAGGAGCTGGAACGGAAGAGCAAGAGCATCATCGAC GAGTTCCTGCACATCAACGACTTCAAGGAGGCCATGCAGTGCCTGG AGGAGCTGAACGCCCAGGGCCTGCTGCACGTGTTCTGTCGGGTGGG CGTGGAGAGCACCTGGAGCGGAGCCAGATCACCCGGGACCACATG GGCCAGCTGCTGTACCAACTGGTGCAGAGCGAGAAGCTGAGCAAGC AGGACTTCTTCAAGGGCTTCAGCGAGACTCTGGAGCTGGCCGACGA CATGGCCATCGACATTCCCCACATCTGGCTGTACCTGGCCGAGCTTG TGACCCCAATGCTGAAGGAGGGCGGCATCAGCATGCGGGAGCTGAC CATCGAGTTCAGCAAGCCCCCTTGCCAGTCGGCAGAGCAGGAGTG CTGCTGAGCGAGATCCTGCACCTGCTGTGCAAGCAGATGAGCCACA AGAAGGTGGGCGCCCTGTGGAGGGAGGCCGACCTGAGCTGGAAGGA CTTCCTGCCCCAGGGCGAAGACGTGCACAACTTCTTGCTGGAGCAGA AGCTGGACTTCATCGAGAGCGACAGCCCCTGCAGCAGCGAGGCCCT GAGCAAGAAGGAACTGAGCGCCGAGGAGCTGTACAAGCGGCTGGA GAAGCTGATCATCGAGGACAAGGCCAACGACGAGCAGATCTTCGAC TGGGTGGAGGCCAACCTGGACGAGATCCAGATGAGCAGCCCCACCT TCCTGCGGGCCCTGATGACCGCCGTGTGCAAGGCCGCCATCATCGCC GACAGCAGCACCTTCCGGGTGGACACCGCCGTGATCAAGCAGCGGG TGCCCATCCTGCTGAAGTACCTGGACAGCGACACCGAGAAGGAGCT GCAGGCCCTGTACGCCCTGCAGGCCAGCATCGTGAAGCTGGACCAG CCCGCCAACCTGCTGCGGATGTTCTTCGACTGCCTGTACGACGAAGA GGTGATCAGCGAGGACGCCTTCTACAAGTGGGAGAGCAGCAAGGAC CCCGCCGAGCAGAACGGCAAGGGCGTGGCCCTGAAGAGCGTGACCG CCTTCTTACCTGGCTGCGAGAGGCCGAGGAGGAGAGCGAGGACAA C
74	MBP-eIF4G3-mid (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELIVK AMQGLLDGNPIPSAIAANSIGYGGGSRITLDPRLPRGPDFTPAFADFG

SEQ ID NO	Sequence information	Sequence
		RQTPGGRGVPLLNVGSRSSQPGQRREPRKIITVSVKEDVHLKKAENAW KPSQKRDSQADDPENIKTQELFRKVRSILNKLTPQMFMNQLMKQVSGLT DTEERLKGVIDLVFEKAIDEPSFSVAYANMCRCLVTLKVPMAADKPGNT VNFRKLLLNRCQKEFEKDKADDDVFEKKQKELEAASAPEERTRLHDEL EEAKDKARRRSIGNIKFIGELFKLKMLTEAMHDCVVKLLKNHDEESLE CLCRLLTTIGKDLD FEKAKPRMDQYFNQMEKIVKERKTSSRIRFMLQDV IDLRLCNWVSRRADQGPKTIEQIHKEAKIEEQEEQRKVQQLMTKEKRRP GVQRVDEGGWNTVQGAKNRSRLDPSKFLKITKPTIDEKIQLVPAQQLGS WGKGSSG
75	MBP-eIF4G3-mid (nt)	ATGGCCAGCAACTTCACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCCG GACCTTGATCCTCGGATCCTTCCAGAGGCCCCGACTTCACTCCCG CCTTCGCCGACTTCGGCAGGCAGACCCAGGTGGTGGGCTGTGCC CTGCTGAACGTGGGCAGCCGTGCGAGCCAGCCTGGCCAACGGCGGG AACCCCGGAAGATCATCACCGTGAGCGTGAAGGAGGACGTGCATCT GAAGAAGGCCGAGAACGCCTGGAAACCCAGCCAGAAGCGGGACAG CCAAGCCGACGACCCCGAGAACATCAAGACCCAGGAGCTGTTCCGG AAGGTTCCGAGCATCCTGAACAAGCTGACACCCAGATGTTCAACC AGCTGATGAAGCAGGTGAGCGGCCTGACCGTGGACACCGAGGAGCG GCTGAAGGGCGTGATCGACCTGGTGTGTTGAGAAAGCCATTGACGAG CCCAGCTTCAGCGTGGCCTACGCCAACATGTGCCGGTGCCTGGTAC CCTGAAGGTGCCCATGGCCGACAAGCCCGGCAACACCGTGAACCTC CGGAAGCTGCTGCTGAACCGGTGCCAGAAGGAGTTCGAGAAGGACA AGGCAGACGACGACGTGTTTCGAGAAGAAGCAGAAAGAGTTAGAGG CCGCAAGCGCCCCCTGAGGAGCGGACCCGGCTGCACGACGAGCTGGA GGAGGCCAAAGACAAGGCCCGGCGGAGGAGCATCGGCCAACATCAA GTTTCATCGGCGAGCTGTTCAAGCTGAAGATGCTGACCGAGGCCATCA TGACGACTGCGTGGTGAAGCTGCTGAAGAACCACGACGAGGAGAG CCTGGAGTGCTGTGCCGGCTGCTGACCACCATCGGCCAAGGACCTGG ACTTCGAGAAGGCCAAGCCCCGGATGGACCAGTACTTCAACCAGAT GGAGAAGATCGTGAAGGAGCGGAAGACCAGCAGCCGGATCCGGTTC ATGCTGCAGGACGTTATCGACCTGCGGCTGTGCAACTGGGTGAGCCG GAGGGCTGACCAGGGCCCCAAGACCATCGAGCAGATCCACAAGGAG GCTAAGATCGAGGAGCAGGAAGAGCAGCGGAAGGTGCAGCAGCTG ATGACCAAAGAGAAGCGGAGGCCTGGCGTGCAGCGAGGTGGACGAG GGCGGCTGGAACACCGTGCAGGGCGCCAAGAACAGCCGGGTGCTGG ACCCAGCAAGTTCCTGAAGATCACCAAGCCACCATCGACGAGAA GATCCAGCTGGTGCCCAAGGCCAGCTCGGCAGCTGGGGCAAGGGC AGCTCAGGT
76	Gld2 (aa)	FPNSILGRPPFTPNHQHNNFFTLSPVYSHQQQLIDAQFNFNQADLS RAVSLQQLTYGNVSPQTSASPLFRGRKRLSDEKNLPLDGKRQRFHSPH QEPTVVNQIVPLSGERRYMPPLFHTHYVPDIVRCVPPFREIAFLEPREIT LPEAKDKLSQQILELFETCQQQISDLKKKELCRTQLQREIQLLPQSRLFL VGSSLNGFGTRSSDGLCLVVKEEPCFFQVNQKTEARHILTLVHKHFT RLSGYIERPQLIRAKVPIVKFRDKVSCVEFDLNVNNIVGIRNTFLLRTYA YLENRVRPLVLVIKKWASHHQINDASRGTLSSYSLVLMVLHYLQTLPEP ILPSLQKIYPESFSPAIQLHLVHQAPCNVPPYLSKNESNLGDLGLFLKYY

SEQ ID NO	Sequence information	Sequence
		ATEFDWNSQMISVREAKAIPRPDGIWRNKYICVEEPFDGTNTARAVHE KQKFDMIKDQFLKSWHRLKNKRDLSILPVRAAVLKR
77	Gld2 (nt)	TTCCCTAACAGCATCCTGGGCAGACCTCCTTTACCCCTAACCAC CAGCAGCACAACTTCTTACCCTGAGCCCTACCGTGTACAGCCA TCAGCAGCTGATCGACGCCAGTTCAACTTCCAGAACGCCGACCTGA GCAGAGCCGTGAGCCTGCAGCAACTCACCTACGGCAACGTTAGCCC AATCCAGACCAGCGCCAGCCCTCTGTTTACAGAGGCAGAAAGAGACTG AGCGACGAGAAGAACCTGCCTCTGGACGGCAAGAGACAGAGATTCC ACAGCCCTCACCAGGAGCCAAACAGTGGTGAACCAGATCGTGCCTCT GAGCGGCGAGAGAAGATACAGCATGCCTCCACTGTTCCACACCCAC TACGTGCCTGACATCGTGAGGTGCGTGCCACCTTTCAGAGAGATCGC CTTCTGGAGCCTCGCGAGATTACCCTGCCTGAGGCCAAGGACAAGC TGAGCCAGCAGATTCTGGAAGTGTTCGAGACTTGCCAACAACAAATT AGCGACCTGAAGAAGAAGGAGCTGTGCAGAACCCAGCTGCAGCGCG AGATCCAGCTGCTGTTCCCTCAGAGCAGACTGTTCTGGTGGGCAGC AGCCTGAACGGCTTCGGCACCAGAAGCAGCGACGGCGACCTGTGCC TGGTGGTGAAGGAGGAGCCTTGCTTCTTCCAAGTAAATCAGAAGAC CGAGGCCAGACACATCCTGACCCTGGTGCACAAAGCACTTTCGACCA GACTGTCCGGCTACATCGAGCGCCCTCAGCTTATAAGAGCCAAGGTG CCTATCGTGAAGTTCAGAGACAAGGTGAGCTGCGTGGAGTTCGATCT GAACGTGAACAACATCGTGGGCATCAGAAACACCTTCTGCTGAGA ACCTACGCCTACCTGGAGAACAGAGTCAGACCTCTGGTGTGGTGAT CAAGAAGTGGGCTTCTCATCACCAGATCAACGACGCCTCCCGCGGTA CTCTGTCTAGCTACAGCCTCGTACTGATGGTGTGCTGCACTACCTGCAG ACTCTCCCTGAACCTATCCTGCCTAGTCTGCAGAAGATCTACCCTGA GAGCTTCAGCCCTGCCATCCAATTGCATCTCGTTACCAGGCCCTT GCAACGTTCTCCTTATTTGAGCAAGAACGAGAGCAACCTGGGAGA TCTCCTGCTGGGCTTCTGAAGTACTACGCCACCGAATTCGATTGGA ACAGCCAGATGATCAGCGTGAGAGAAGCAAAGGCCATCCCTAGACC TGACGGCATCGAGTGGAGAAACAAGTACATCTGTGTGGAGGAACCT TTCGACGGCACCAACACCGCCCGCGCAGTACACGAGAAGCAGAAGT TCGACATGATCAAGGACCAGTTCCTCAAGAGCTGGCACAGACTGAA GAACAAGAGAGACTTGAATAGCATCTTGCCGGTGAGAGCCGCCGTG CTGAAGAGA
78	MBP-Gld2 (aa)	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITIPFATNSDCELVK AMQGLLKDGNIPIPSAIAANSYGGGSGFPNSILGRPPFTPNHQHNNFF TLPSTVYSHQQLIDAQFNFNADLSRAVSLQQLTYGNVSPIQTSASPLFR GRKRLSDEKNLPLDGKRQRFHSPHQEPTVVNQIVPLSGERRYMPPLFH THYVPDIVRCVPPFREIAFLEPREITLPEAKDKLSQQILELFETCQQQISDL KKKELCRTQLQREIQLLFPQSRLFLVGSSLNGFGTRSSDGDCLLVKEEP CFFQVNQKTEARHILTLVHKHFCTRLSGYIERPQLIRAKVPIVKFRDKVS CVEFDLNVNNIVGIRNTFLLRTYAYLENRVRPLVLVIKKWASHHQINDA SRGTLSSYSLVLMVLHYLQTLPEPILPSLQKIYPESFSPAIQLHLVHQAPC NVPPYLSKNESNLGDLGLFLKYAYATEFDWNSQMISVREAKAIPRPDGI EWRNKYICVEEPFDGTNTARAVHEKQKFDMIKDQFLKSWHRLKNKRD LNSILPVRAAVLKR
79	MBP-Gld2 (nt)	ATGGCCAGCAACTTACCCAGTTCGTGCTGGTGGACAACGGCGG TACCGGAGACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCG CCGAGTGGATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGAC CTGCAGCGTGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATC AAGGTGGAGGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGG AGCTGACCATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATC

SEQ ID NO	Sequence information	Sequence
		GTGAAGGCCATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTA GCGCCATCGCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCTT CCCTAACAGCATCCTGGGCAGACCTCCTTTCACCCCTAACACCAGC AGCACAACAACCTTCTTCACCCTGAGCCCTACCGTGTACAGCCATCAG CAGCTGATCGACGCCCAGTTCAACTTCCAGAACGCCGACCTGAGCA GAGCCGTGAGCCTGCAGCAACTCACCTACGGCAACGTTAGCCCAAT CCAGACCAGCGCCAGCCCTCTGTTTCAGAGGCAGAAAGAGACTGAGC GACGAGAAGAACCTGCCTCTGGACGGCAAGAGACAGAGATTCCACA GCCCTCACCAGGAGCCAACAGTGGTGAACCAGATCGTGCCTCTGAG CGGCGAGAGAAGATACAGCATGCCTCCACTGTTCCACACCCACTAC GTGCCTGACATCGTGAGGTGCGTGCCACCTTTCAGAGAGATCGCCTT CCTGGAGCCTCGCGAGATTACCCTGCCTGAGGCCAAGGACAAGCTG AGCCAGCAGATTCTGGAAGTGTTCGAGACTTGCCAACAACAAATTA GCGACCTGAAGAAGAAGGAGCTGTGCAGAACCCAGCTGCAGCGCA GATCCAGCTGCTGTTCCCTCAGAGCAGACTGTTCTTGGTGGGCAGCA GCCTGAACGGCTTCGGCACCAGAAGCAGCAGCGGACCTGTGCCT GGTGGTGAAGGAGGAGCCTTGCTTCTTCCAAGTAAATCAGAAGACC GAGGCCAGACACATCCTGACCCTGGTGCACAAGCACTTCTGCACCA GACTGTCCGGCTACATCGAGCGCCCTCAGCTTATAAGAGCCAAGGTG CCTATCGTGAAGTTCAGAGACAAGGTGAGCTGCGTGAGTTCGATCT GAACGTGAACAACATCGTGGGCATCAGAAACACCTTCTGCTGAGA ACCTACGCCTACCTGGAGAACAGAGTCAGACCTCTGGTGCTGGTGAT CAAGAAGTGGGCTTCTCATCACCAGATCAACGACGCCTCCCGCGGTA CTCTGTCTAGCTACAGCCTCGTACTGATGGTGCTGCACTACCTGCAG ACTCTCCCTGAACCTATCCTGCCTAGTCTGCAGAAGATCTACCCTGA GAGCTTCAGCCCTGCCATCCAATTGCATCTCGTTCACCAGGCCCTT GCAACGTTCCCTCCTTATTTGAGCAAGAACGAGAGCAACCTGGGAGA TCTCCTGCTGGGCTTCTTGAAGTACTACGCCACCGAATTCGATTGGA ACAGCCAGATGATCAGCGTGAGAGAAGCAAAGGCCATCCCTAGACC TGACGGCATCGAGTGGAGAAACAAGTACATCTGTGTGGAGGAACCT TTCGACGGCACCAACACCGCCCGCGCAGTACACGAGAAGCAGAAGT TCGACATGATCAAGGACCAGTTCTCAAGAGCTGGCACAGACTGAA GAACAAGAGAGACTTGAATAGCATCTTGCCGGTGAGAGCCGCCGTG CTGAAGAGA
155	eIF4G-mid2mut (aa)	DPTRLQGINCGPDFTPSFANLGRITTLSTRGPPRGGPGGELPRGPAGL GPRRSQQGPRKEPRKIIATVLMTEDIKLNKAEKAWKPSSKRTAADKDR GEEDADGSKTQDLFRRVRSILNKLTPQMFQQLMKQVTQLAIDTEERLK GVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTVTNFRKLLL NRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGRLEELEEARDIAR RRLGNIAIGELFKLKMLTEAIMHDCVVKLLKNHDEESLECLRLLTTI GKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDRGSNW VPRRGDQGPKTIDQIHKEAEMEEHREHIKVQQLMAKGSDKRRGGPPGP PISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKPGSIDSNNQLFAPGG RLSWGKGSSGSGAKPSDAASEAARPATSTLNRFSALQQAV
156	eIF4G-mid2mut (nt)	GACCCTACCAGACTGCAGGGCATCAACTGCGGCCCTGACTTCAC TCCTTCTTTTCGAAACCTGGGCAGAACACCCTGAGCACCAGAGGCC CTCCTAGAGGTGGTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGG CCTGGGCCCTAGACGCTCCCAGCAAGGTCTAGAAAGGAGCCAAGG AAGATCATCGCCACCGTGCTGATGACCGAGGACATCAAGCTGAACA AAGCTGAGAAGGCCTGGAAGCCTAGCAGCAAGAGAACCGCCGCCGA CAAGGACAGAGGCGAGGAGGACGCCGACGGATCCAAGACCCAGGA CCTGTTTCAAGAGTGAGAAGCATCCTCAACAAGCTGACCCCTCAG ATGTTCCAGCAGCTGATGAAGCAGGTGACGCAGCTCGCCATCGACA

SEQ ID NO	Sequence information	Sequence
		CCGAGGAGAGACTGAAGGGCGTGATCGACCTGATCTTTGAGAAGGC TATCTCAGAGCCTAACTTCAGCGTGGCCTACGCCAACATGTGCCGTT GCCTGATGGCATTGAAGGTGCCAACCACCGAGAAGCCTACTGTGAC CGTCAATTTCCGTAACTGCTGCTGAACCGGTGCCAGAAAGAGTTTCG AGAAGGATAAGGACGACGACGAGGTCTTCGAGAAGAAACAGAAAG AAATGGACGAGGCCGCCACCGCAGAGGAAAGGGGCCGATTAAAGG AGGAGCTGGAGGAGGCCAGAGACATCGCCAGACGGCGTTCTCTGGG CAACATCAAGGCGATAGGTGAGCTGTTCAAGCTAAAGATGCTCACC GAGGCCATAATGCACGACTGCGTGGTGAAGCTACTGAAGAACCACG ACGAAGAAAGCCTGGAGTGCCTGTGCAGACTGCTGACCACCATCGG CAAGGACCTGGACTTCGAGAAGGCCAAAGCCTCGAATGGACCAGTAC TTCAACCAGATGGAGAAGATTATCAAGGAGAAGAAGACCAGCAGCA GAATCAGATTATGCTGCAGGACGTACTGGACCTGCGCGGAAGCAA CTGGGTGCCAAGGAGAGGGGACCAAGGACCAAGACCATCGACCA GATCCACAAGGAAGCGGAGATGGAGGAGCACAGAGAGCACATAAA GGTGCAGCAGCTTATGGCCAAGGGCAGCGACAAGCGAAGAGGGCGGC CCGCCCGGACCTCCTATCAGCAGAGGCCTTCCTCTGGTAGACGACGG CGGCTGGAACACCGTGCCTATCTCTAAGGGCTCCAGACCTATCGACA CTTCCCGTCTTACCAAGATCACCAAGCCAGGATCTATTGACAGCAAC AACCAGCTGTTGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAA GTTCCGGCGGATCCGGCGCCAAGCCTTCCGACGCCGCCAGCGAGGC TGCCAGACCTGCCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGC AAGCTGTG

Tether Molecule

Disclosed herein, *inter alia*, is a system or LNP comprising a polynucleotide, e.g., a second polynucleotide, e.g., mRNA, encoding an RNA-binding protein, e.g., a tether molecule. In some embodiments, the second polynucleotide encodes a effector molecule which further comprises a tether molecule.

In some embodiments, a system or LNP disclosed herein comprises a first polynucleotide comprising a binding element. In some embodiments, the tether molecule, e.g., effector molecule further comprising a tether molecule, binds to a binding element in the first polynucleotide. In some embodiments, the tether molecule, e.g., effector molecule further comprising a tether molecule, binds to a sequence of the binding element or to a structure comprising the sequence of the binding element. In some embodiments, a tether molecule comprises an RNA-binding protein or a variant or a fragment thereof. Exemplary RNA-binding proteins are provided in Tables 1 and 4.

In some embodiments, the tether molecule comprises a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, 15.5kd, LARP7 or a variant or fragment thereof. In some embodiments, the tether molecule is MBP. In some embodiments, the tether molecule is PCP. In some embodiments, the tether molecule is Lambda N. In some embodiments, the tether molecule is U1A. In some embodiments, the tether molecule is PUF. In some embodiments, the tether molecule is 15.5 kd. In some embodiments, the tether molecule is LARP7.

In some embodiments, when the tether molecule is MBP (e.g., wildtype MBP, a variant or fragment thereof) the binding element is MS2 (e.g., wildtype MS2, or a variant or fragment thereof).

In some embodiments, when the tether molecule is PCP (e.g., wildtype PCP, or a variant or fragment thereof) the binding element is PP7 (e.g., wildtype PP7, or a variant or fragment thereof).

In some embodiments, when the tether molecule is Lambda N (e.g., wildtype Lambda N, or a variant or fragment thereof) the binding element is BoxB (e.g., wildtype BoxB, or a variant or fragment thereof).

In some embodiments, when the tether molecule is U1A (e.g., wildtype U1A, or a variant or fragment thereof) the binding element is U1A hairpin (e.g., wildtype U1A hairpin, or a variant or fragment thereof).

In some embodiments, when the tether molecule is 15.5kd (e.g., wildtype 15.5kd, or a variant or fragment thereof) the binding element is a kink-turn forming sequence (e.g., wildtype U1A hairpin, or a variant or fragment thereof).

In some embodiments, when the tether molecule is PUF (e.g., wildtype PUF, or a variant or fragment thereof) the binding element is PRE (e.g., wildtype PRE, or a variant or fragment thereof).

In some embodiments, when the tether molecule is LARP7 (e.g., wildtype LARP7, or a variant or fragment thereof) the binding element is 7SK (e.g., wildtype 7SK, or a variant or fragment thereof).

Additional exemplary RNA-binding proteins or RNA-binding domains which can be used as tether molecules are disclosed in Corley et al, *Molecular Cell* 78:1 pp. 9-29, the entire contents of which are hereby incorporated by reference. For example,

Table 3 provides additional exemplary RNA-binding proteins or domains which can be used as tether molecules. In an embodiment, a tether molecule disclosed herein comprises a domain (or a variant, or a fragment thereof) or a protein (or a variant or a fragment thereof) listed in Table 3.

5

Table 3: Exemplary RNA-binding proteins and domains.

Domain name	Protein family containing domain
Cold shock domain	Cold shock proteins, Y-box proteins
Double stranded RNA binding domain	RNases, ADARs, DICER
Helicase	DExH/D-box, Ski2-like, RIGI-like, NS3, UPF1-like RNA binding helicases
Intrinsically disordered regions (IDR)	Most RBPs
K homology	hnRNPs, translation regulation proteins,
La motif (LAM)	La proteins, La-related proteins (LARPs)
Piwi-Argonaute-Zwille (PAZ)	Argonaute proteins, Dicer
P-element Induced Wimpy Testis (PIWI)	Argonaute proteins,
Pentatricopeptide repeat (PPR)	RNA editing proteins
Pseudouridine synthase and archaeosine transglycosylase (PUA)	RNA modifying enzymes, metabolic enzymes
Pumillo-like repeat (PUM)	PUF proteins
Ribosomal S1-like	Ribosomal proteins, Translation initiation factors, RNase II, PNPase
RNA recognition motif (RRM)	hnRNPs, splicing factors,

Sm and Like-Sm (Sm / Lsm)	U1 spliceosomal proteins, Hfq
thiouridine synthases, RNA methylases and pseudouridine synthases (THUMP)	tRNA modifying enzyme
YT521-B homology (YTH)	YTH family m6A readers
Zn finger	Transcription factors, METTL enzymes,

In some embodiments, the tether molecule comprises MBP. In some embodiments, the tether molecule comprises an amino acid sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof. In some embodiments, the tether molecule comprises the amino acid sequence of SEQ ID NO: 6, or an amino acid sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the tether molecule comprises is encoded by a nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof. In some embodiments, the tether molecule comprises is encoded by the nucleotide sequence of SEQ ID NO: 7, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

Effector molecule

Disclosed herein, *inter alia*, is a system or LNP comprising a polynucleotide, e.g., a second polynucleotide, e.g., mRNA, encoding an effector molecule. In some embodiments, the effector molecule is chosen from a factor provided in Table 4, e.g., a translation factor, a splicing factor, an RNA stabilizing factor, an RNA editing factor, an RNA-binding factor (e.g., PABP that binds the polyA tail of an mRNA), an RNA localizing factor, or an RNA modulating factor (such as Gld2, TENT 4A and TENT4B which are known to add more As and or A/G nucleotides to the polyA tail of an mRNA) or a combination thereof.

In some embodiments, the effector molecule is a translation factor, e.g., a translation factor provided in Table 4, e.g., eIF4G; Poly A binding protein (PABP); eIF3d or a component thereof; Dazl, or a fragment, or variant or combination thereof. Additional exemplary translation factors are provided in Pelletier and Soneneberg.

5 Annu. Rev. Biochem. 2019. 88:307–35, the entire contents of which are hereby incorporated by reference

In some embodiments, the effector molecule is a splicing factor, e.g., a splicing factor provided in Table 4, e.g., Rnps1, Magoh, Y14, or a fragment or variant, or combination thereof. Additional splicing factors are provided in Nott et al, (2004) *Genes*
10 & *Dev*, 2004. 18, 210-222, the entire contents of which are hereby incorporated by reference.

In some embodiments, the effector molecule is an RNA stabilizing factor, e.g., a stabilizing factor provided in Table 4, e.g., HuR or a fragment, or variant thereof. Exemplary RNA stabilizing factors are provided in Goldberg et al., (2002) *J Biol Chem*.
15 2002 Apr 19;277(16):13635-40 , the entire contents of which are hereby incorporated by reference.

In some embodiments, the effector molecule is an RNA editing factor. Exemplary RNA editing factors are provided in Kim D. et al (2019), *Ann Rev Biochem*, 88, 191-200, the entire contents of which are hereby incorporated by reference.

20 In some embodiments, the effector molecule is an RNA binding factor. Exemplary RNA binding factors are provided in Singh G. et al (2015) *Annu Rev. Biochem*; 84: 325–354, the entire contents of which are hereby incorporated by reference.

In some embodiments, the effector molecule is an RNA localizing factor. Exemplary RNA localizing factors are provided in Blower M.D. (2013) *Int Rev Cell Mol Biol*.; 302: 1–39, the entire contents of which are hereby incorporated by reference.

In some embodiments, the effector molecule is an RNA modulating factor. Exemplary RNA factors are provided in Philos Trans R Soc Lond B Biol Sci. 2018 Dec 19; 373(1762), the entire contents of which are hereby incorporated by reference.

30

Table 4: Exemplary effector molecules

Effector molecule	Examples
Translation factor	eIF4G; Poly A binding protein (PABP); eIF3d; Dazl
Splicing factor	Rnps1, Magoh, Y14
RNA stabilizing factor	HuR Exemplary RNA stabilizing factors provided in Goldberg et al., (2002) J Biol Chem. 2002 Apr 19;277(16):13635-40
RNA editing factor	Exemplary RNA editing factors are provided in Kim D. et al (2019), Ann Rev Biochem, 88, 191-200
RNA binding factor	Exemplary RNA binding factors provided in Singh G. et al (2015) Annu Rev. Biochem; 84: 325–354
RNA localizing factor	Blower M.D. (2013) Int Rev Cell Mol Biol.; 302: 1–39
RNA modulating factor	Philos Trans R Soc Lond B Biol Sci. 2018 Dec 19; 373(1762)

In some embodiments, the effector molecule binds directly to the binding
5 element. The effector molecule may have a specific target sequence to which it can bind. Effector molecules include, but are not limited to eIF4G, eIF4d, PABPC, TENT4A, TENT4B, and Gld2.

In some embodiments, the effector molecule further comprises a polypeptide that binds to, e.g., recognizes, the binding element (a tether molecule).

10 In an embodiment, the effector molecule polypeptide comprising a tether molecule comprises a first domain which modulates a parameter of, e.g., level and/or activity of: an RNA (e.g., an mRNA); or a protein encoded by the RNA. In an embodiment, the parameter comprises one, two, three or all of: (1) mRNA level and/or activity and/or subcellular localization (e.g., half-life and/or expression); (2) protein
15 level and/or activity (e.g., half-life and/or expression); (3) protein translation rate or (4)

protein localization, e.g., location. In an embodiment the effector molecule polypeptide comprising a tether molecule comprises a second domain which binds to, e.g., recognizes, the binding element (a tether molecule).

In an embodiment the effector molecule comprising the tether molecule
 5 comprises a polypeptide comprising the first domain and the second domain. In an embodiment, the first and second domains are operatively linked.

In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the effector molecule is upstream of the nucleotide sequence encoding the tether molecule.

10 In an embodiment, the nucleotide sequence encoding the effector molecule is downstream of the nucleotide sequence encoding the tether molecule. In an embodiment, the nucleotide sequence encoding the effector molecule is separated from the nucleotide sequence encoding the tether molecule by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A) site) or an internal ribosomal entry site.

15 In an embodiment, in the second polynucleotide encoding the effector molecule which further comprises a tether molecule, the nucleotide sequence encoding the effector molecule is adjacent to the nucleotide sequence encoding the tether molecule.

In some embodiments, the effector molecule is a translation factor which
 20 modulates, e.g., facilitates, ribosome binding, e.g., recruitment, pre-initiation complex formation, or RNA unwinding. In some embodiments, the effector molecule comprises eIF4G, e.g., wildtype eIF4G, a variant of eIF4G, or a fragment thereof.

In some embodiments, the effector molecule comprises wildtype eIF4G. In some embodiments, wildtype eIF4G comprises a sequence of about 1600 amino acids.

25 In some embodiments, the effector molecule comprises a fragment of eIF4G, e.g., as disclosed herein. In some embodiments, the eIF4G fragment retains ribosome binding, e.g., recruitment.

In some embodiments, the eIF4G fragment is about 1,500-200 amino acids, about 1,400-300 amino acids, about 1,300-350 amino acids, about 1,200-400 amino
 30 acids, about 1,100-450 amino acids, about 1,000-500 amino acids, about 900-550 amino acids, about 800-600 amino acids, about 1,500-300 amino acids, 1,500-400 amino

acids, 1,500-500 amino acids, about 1,500-600 amino acids, amino acids, about 1,500-700 amino acids, about 1,500-800 amino acids, about 1,500-900 amino acids, about 1,500-1000 amino acids, about 1,500-1,100 amino acids, about 1,500-1,200 amino acids, about 1,500-1,300 amino acids, about 1,500-1,400 amino acids, about 1,400-200 amino acids, about 1,300-200 amino acids, about 1,200-200 amino acids, about 1,100-200 amino acids, about 1,000-200 amino acids, about 900-200 amino acids, about 800-200 amino acids, about 700-200 amino acids, about 600-200 amino acids, or about 500-200 amino acids in length.

In some embodiments, the eIF4G fragment is about 500 amino acids in length.

10 In some embodiments, the eIF4G fragment is about 600 amino acids in length. In some embodiments, the eIF4G fragment is about 700 amino acids in length. In some embodiments, the eIF4G fragment is about 800 amino acids in length. In some embodiments, the eIF4G fragment is about 900 amino acids in length. In some embodiments, the eIF4G fragment is about 1000 amino acids in length. In some
15 embodiments, the eIF4G fragment is about 1100 amino acids in length. In some embodiments, the eIF4G fragment is about 1200 amino acids in length. In some embodiments, the eIF4G fragment is about 1300 amino acids in length. In some embodiments, the eIF4G fragment is about 1400 amino acids in length. In some embodiments, the eIF4G fragment is about 1500 amino acids in length.

20 In some embodiments, the effector molecule comprises a variant of eIF4G, e.g., as disclosed herein. In some embodiments, the eIF4G variant retains ribosome binding, e.g., recruitment. In some embodiments, the eIF4G variant comprises a mutation (*e.g.*, substitution) in the eIF4G polypeptide sequence at any one, two, all or a combination of the following positions: amino acid 768, amino acid 771, or amino acid 776. In some
25 embodiments, the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 768 of the eIF4G polypeptide sequence, *e.g.*, a Leucine to Alanine substitution at position 768. In some embodiments, the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, a Leucine to Alanine substitution at position 771. In some embodiments, the eIF4G variant comprises a
30 mutation, *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, a Phenylalanine to Alanine at position 776. In some embodiments, the eIF4G variant

comprises a mutation, *e.g.*, substitution, at position 768 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 768; and a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771. In some embodiments, the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 768 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 768; and a mutation, *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 776. In some embodiments, the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771; and a mutation, *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 776. In some embodiments, the eIF4G variant comprises a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771; a mutation, *e.g.*, substitution, at position 771 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 771; and a mutation, *e.g.*, substitution, at position 776 of the eIF4G polypeptide sequence, *e.g.*, an Alanine at position 776.

In some embodiments, the effector molecule is a part of the eIF3 complex, *e.g.*, which can recruit the ribosome. In some embodiments, the eIF3 complex comprises eIF3d, eIF3c, eIF3e, or eIF3i, or a fragment thereof, or any combination thereof.

In some embodiments, the effector molecule, *e.g.*, eIF4G, PABP, TENT4A, TENT4B, or Gld2 comprises an amino acid sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, *e.g.*, eIF4G, comprises SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 60, SEQ ID NO: 62, SEQ ID NO: 64, SEQ ID NO: 66, SEQ ID NO: 68, SEQ ID NO: 70, SEQ ID NO: 72, SEQ ID NO: 74, or SEQ ID NO: 155, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., PABP, comprises SEQ ID NO: 48, or SEQ ID NO: 50, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., TENT4A, comprises SEQ ID NO: 52, or SEQ ID NO: 54, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., TENT4B, comprises SEQ ID NO: 56, or SEQ ID NO: 58, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., Gld2, comprises SEQ ID NO: 76, or SEQ ID NO: 78, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule e.g., eIF4G, PABP, TENT4A, TENT4B, or Gld2 is encoded by a nucleotide sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule e.g., eIF4G, is encoded by the nucleotide sequence of SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 73, SEQ ID NO: 75, or SEQ ID NO: 156 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., PABP, is encoded by the nucleotide sequence of SEQ ID NO: 49, or SEQ ID NO: 51, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., TENT4A, comprises SEQ ID NO: 53, or SEQ ID NO: 55, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., TENT4B, comprises SEQ ID NO: 57, or SEQ ID NO: 59, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

In some embodiments, the effector molecule, e.g., Gld2, comprises SEQ ID NO: 77, or SEQ ID NO: 79, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

Lipid content of LNPs

As set forth above, with respect to lipids, LNPs disclosed herein comprise an (i) ionizable lipid; (ii) sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and, optionally a (iv) PEG lipid. These categories of lipids are set forth in more detail below. In some embodiments, nucleic acids of the invention are formulated as lipid nanoparticle (LNP) compositions.

Lipid nanoparticles typically comprise amino lipid, phospholipid, structural lipid and PEG lipid components along with the nucleic acid cargo of interest. The lipid nanoparticles of the invention can be generated using components, compositions, and methods as are generally known in the art, see for example PCT/US2016/052352; PCT/US2016/068300; PCT/US2017/037551; PCT/US2015/027400; PCT/US2016/047406; PCT/US2016000129; PCT/US2016/014280; PCT/US2016/014280; PCT/US2017/038426; PCT/US2014/027077; PCT/US2014/055394; PCT/US2016/52117; PCT/US2012/069610; PCT/US2017/027492; PCT/US2016/059575; PCT/US2016/069491; PCT/US2016/069493; and PCT/US2014/66242, all of which are incorporated by reference herein in their entirety.

In some embodiments, the lipid nanoparticle comprises a molar ratio of 20-60% amino lipid relative to the other lipid components. For example, the lipid nanoparticle may comprise a molar ratio of 20-50%, 20-40%, 20-30%, 30-60%, 30-50%, 30-40%, 40-60%, 40-50%, or 50-60% amino lipid. In some embodiments, the lipid nanoparticle comprises a molar ratio of 20%, 30%, 40%, 50, or 60% amino lipid.

In some embodiments, the lipid nanoparticle comprises a molar ratio of 5-25% phospholipid relative to the other lipid components. For example, the lipid nanoparticle may comprise a molar ratio of 5-30%, 5-15%, 5-10%, 10-25%, 10-20%, 10-25%, 15-25%, 15-20%, 20-25%, or 25-30% phospholipid. In some embodiments, the lipid nanoparticle comprises a molar ratio of 5%, 10%, 15%, 20%, 25%, or 30% non-cationic lipid.

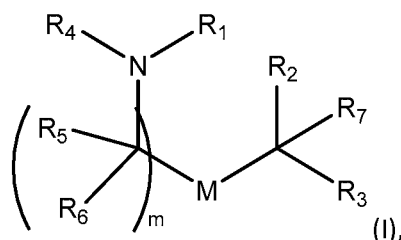
In some embodiments, the lipid nanoparticle comprises a molar ratio of 25-55% structural lipid relative to the other lipid components. For example, the lipid nanoparticle may comprise a molar ratio of 10-55%, 25-50%, 25-45%, 25-40%, 25-35%, 25-30%, 30-55%, 30-50%, 30-45%, 30-40%, 30-35%, 35-55%, 35-50%, 35-45%, 35-40%, 40-55%, 40-50%, 40-45%, 45-55%, 45-50%, or 50-55% structural lipid. In some embodiments, the lipid nanoparticle comprises a molar ratio of 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, or 55% structural lipid.

In some embodiments, the lipid nanoparticle comprises a molar ratio of 0.5-15% PEG lipid relative to the other lipid components. For example, the lipid nanoparticle may comprise a molar ratio of 0.5-10%, 0.5-5%, 1-15%, 1-10%, 1-5%, 2-15%, 2-10%, 2-5%, 5-15%, 5-10%, or 10-15% PEG lipid. In some embodiments, the lipid nanoparticle comprises a molar ratio of 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, or 15% PEG-lipid.

In some embodiments, the lipid nanoparticle comprises a molar ratio of 20-60% amino lipid, 5-25% phospholipid, 25-55% structural lipid, and 0.5-15% PEG lipid.

In some embodiments, the lipid nanoparticle comprises a molar ratio of 20-60% amino lipid, 5-30% phospholipid, 10-55% structural lipid, and 0.5-15% PEG lipid.

Amino lipids In some aspects, the amino lipids of the present disclosure may be one or more of compounds of Formula (I):



or their N-oxides, or salts or isomers thereof, wherein:

R₁ is selected from the group consisting of C₅₋₃₀ alkyl, C₅₋₂₀ alkenyl, -R*YR'', -YR'', and -R''M'R';

R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, C₂₋₁₄ alkenyl, -R*YR'', -YR'', and -R*OR'', or R₂ and R₃, together with the atom to which they are attached, form a heterocycle or carbocycle;

R₄ is selected from the group consisting of hydrogen, a C₃₋₆ carbocycle, -(CH₂)_nQ, -(CH₂)_nCHQR,

-CHQR, -CQ(R)₂, and unsubstituted C₁₋₆ alkyl, where Q is selected from a carbocycle, heterocycle, -OR, -O(CH₂)_nN(R)₂, -C(O)OR, -OC(O)R, -CX₃, -CX₂H, -CXH₂, -CN, -N(R)₂, -C(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)C(O)N(R)₂, -N(R)C(S)N(R)₂, -N(R)R₈, -N(R)S(O)₂R₈, -O(CH₂)_nOR, -N(R)C(=NR₉)N(R)₂, -N(R)C(=CHR₉)N(R)₂, -OC(O)N(R)₂, -N(R)C(O)OR, -N(OR)C(O)R, -N(OR)S(O)₂R, -N(OR)C(O)OR, -N(OR)C(O)N(R)₂, -N(OR)C(S)N(R)₂, -N(OR)C(=NR₉)N(R)₂, -N(OR)C(=CHR₉)N(R)₂, -C(=NR₉)N(R)₂, -C(=NR₉)R, -C(O)N(R)OR, and -C(R)N(R)₂C(O)OR, and each n is independently selected from 1, 2, 3, 4, and 5;

each R₅ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R₆ is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

M and M' are independently selected

from -C(O)O-, -OC(O)-, -OC(O)-M''-C(O)O-, -C(O)N(R')-,

-N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)₂-, -S-S-, an aryl group, and a heteroaryl group, in which M'' is a bond, C₁₋₁₃ alkyl or C₂₋₁₃ alkenyl;

R₇ is selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

R₈ is selected from the group consisting of C₃₋₆ carbocycle and heterocycle;

R₉ is selected from the group consisting of H, CN, NO₂, C₁₋₆ alkyl, -OR, -S(O)₂R, -S(O)₂N(R)₂, C₂₋₆ alkenyl, C₃₋₆ carbocycle and heterocycle;

5 each R is independently selected from the group consisting of C₁₋₃ alkyl, C₂₋₃ alkenyl, and H;

each R' is independently selected from the group consisting of C₁₋₁₈ alkyl, C₂₋₁₈ alkenyl, -R*YR'', -YR'', and H;

each R'' is independently selected from the group consisting of C₃₋₁₅ alkyl and

10 C₃₋₁₅ alkenyl;

each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and

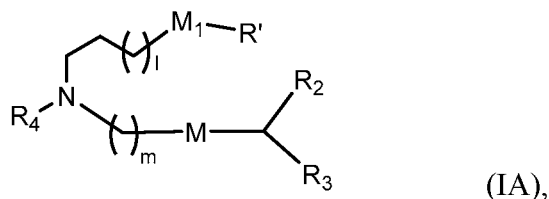
C₂₋₁₂ alkenyl;

each Y is independently a C₃₋₆ carbocycle;

each X is independently selected from the group consisting of F, Cl, Br, and I; and

15 m is selected from 5, 6, 7, 8, 9, 10, 11, 12, and 13; and wherein when R₄ is -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, or -CQ(R)₂, then (i) Q is not -N(R)₂ when n is 1, 2, 3, 4 or 5, or (ii) Q is not 5, 6, or 7-membered heterocycloalkyl when n is 1 or 2.

In certain embodiments, a subset of compounds of Formula (I) includes those of Formula (IA):

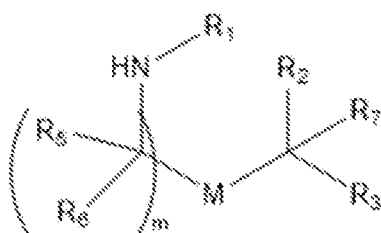


20

or its N-oxide, or a salt or isomer thereof, wherein l is selected from 1, 2, 3, 4, and 5; m is selected from 5, 6, 7, 8, and 9; M₁ is a bond or M'; R₄ is hydrogen, unsubstituted C₁₋₃ alkyl, or -(CH₂)_nQ, in which Q is

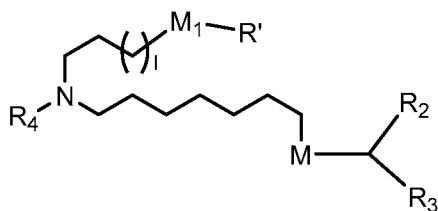
OH, $-\text{NHC}(\text{S})\text{N}(\text{R})_2$, $-\text{NHC}(\text{O})\text{N}(\text{R})_2$, $-\text{N}(\text{R})\text{C}(\text{O})\text{R}$, $-\text{N}(\text{R})\text{S}(\text{O})_2\text{R}$, $-\text{N}(\text{R})\text{R}_8$, $-\text{NHC}(=\text{NR}_9)\text{N}(\text{R})_2$, $-\text{NHC}(=\text{CHR}_9)\text{N}(\text{R})_2$, $-\text{OC}(\text{O})\text{N}(\text{R})_2$, $-\text{N}(\text{R})\text{C}(\text{O})\text{OR}$, heteroaryl or heterocycloalkyl; M and M' are independently selected from $-\text{C}(\text{O})\text{O}-$, $-\text{OC}(\text{O})-$, $-\text{OC}(\text{O})-\text{M}''-\text{C}(\text{O})\text{O}-$, $-\text{C}(\text{O})\text{N}(\text{R}')-$, $-\text{P}(\text{O})(\text{OR}')\text{O}-$, $-\text{S}-\text{S}-$, an aryl group, and a heteroaryl group; and R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, and C₂₋₁₄ alkenyl. For example, m is 5, 7, or 9. For example, Q is OH, $-\text{NHC}(\text{S})\text{N}(\text{R})_2$, or $-\text{NHC}(\text{O})\text{N}(\text{R})_2$. For example, Q is $-\text{N}(\text{R})\text{C}(\text{O})\text{R}$, or $-\text{N}(\text{R})\text{S}(\text{O})_2\text{R}$.

10 In certain embodiments, a subset of compounds of Formula (I) includes those of Formula (IB):



(IB), or its N-oxide, or a salt or isomer thereof in which all variables are as defined herein. For example, m is selected from 5, 6, 7, 8, and 9; R₄ is hydrogen, unsubstituted C₁₋₃ alkyl, or $-(\text{CH}_2)_n\text{Q}$, in which Q is OH, $-\text{NHC}(\text{S})\text{N}(\text{R})_2$, $-\text{NHC}(\text{O})\text{N}(\text{R})_2$, $-\text{N}(\text{R})\text{C}(\text{O})\text{R}$, $-\text{N}(\text{R})\text{S}(\text{O})_2\text{R}$, $-\text{N}(\text{R})\text{R}_8$, $-\text{NHC}(=\text{NR}_9)\text{N}(\text{R})_2$, $-\text{NHC}(=\text{CHR}_9)\text{N}(\text{R})_2$, $-\text{OC}(\text{O})\text{N}(\text{R})_2$, $-\text{N}(\text{R})\text{C}(\text{O})\text{OR}$, heteroaryl or heterocycloalkyl; M and M' are independently selected from $-\text{C}(\text{O})\text{O}-$, $-\text{OC}(\text{O})-$, $-\text{OC}(\text{O})-\text{M}''-\text{C}(\text{O})\text{O}-$, $-\text{C}(\text{O})\text{N}(\text{R}')-$, $-\text{P}(\text{O})(\text{OR}')\text{O}-$, $-\text{S}-\text{S}-$, an aryl group, and a heteroaryl group; and R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, and C₂₋₁₄ alkenyl. For example, m is 5, 7, or 9. For example, Q is OH, $-\text{NHC}(\text{S})\text{N}(\text{R})_2$, or $-\text{NHC}(\text{O})\text{N}(\text{R})_2$. For example, Q is $-\text{N}(\text{R})\text{C}(\text{O})\text{R}$, or $-\text{N}(\text{R})\text{S}(\text{O})_2\text{R}$.

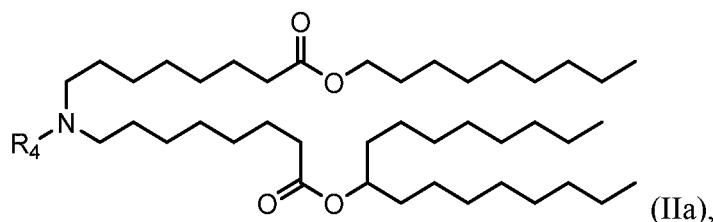
In certain embodiments, a subset of compounds of Formula (I) includes those of Formula (II):



(II), or its N-oxide, or a salt or isomer

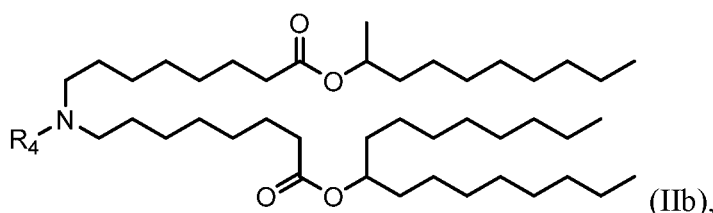
thereof, wherein l is selected from 1, 2, 3, 4, and 5; M₁ is a bond or M'; R₄ is hydrogen, unsubstituted C₁₋₃ alkyl, or -(CH₂)_nQ, in which n is 2, 3, or 4, and Q is OH, -NHC(S)N(R)₂, -NHC(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)R₈,
5 -NHC(=NR₉)N(R)₂, -NHC(=CHR₉)N(R)₂, -OC(O)N(R)₂, -N(R)C(O)OR, heteroaryl or heterocycloalkyl; M and M' are independently selected from -C(O)O-, -OC(O)-, -OC(O)-M''-C(O)O-, -C(O)N(R')-, -P(O)(OR')O-, -S-S-, an aryl group, and a heteroaryl group; and R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, and C₂₋₁₄ alkenyl.

10 In one embodiment, the compounds of Formula (I) are of Formula (IIa),



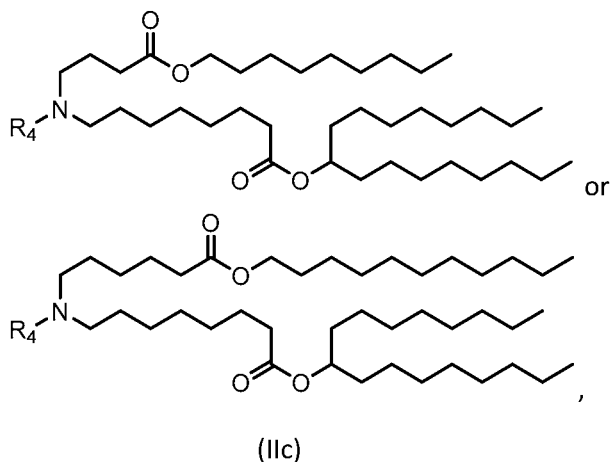
or their N-oxides, or salts or isomers thereof, wherein R₄ is as described herein.

In another embodiment, the compounds of Formula (I) are of Formula (IIb),



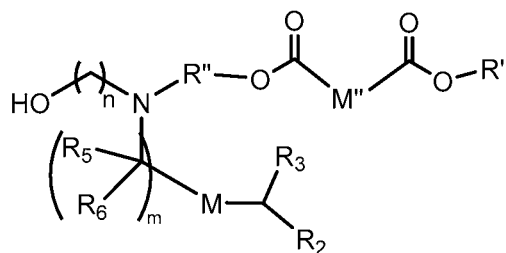
15 or their N-oxides, or salts or isomers thereof, wherein R₄ is as described herein.

In another embodiment, the compounds of Formula (I) are of Formula (IIc) or (IIe):



or their N-oxides, or salts or isomers thereof, wherein R₄ is as described herein.

5 In another embodiment, the compounds of Formula (I) are of Formula (IIf):



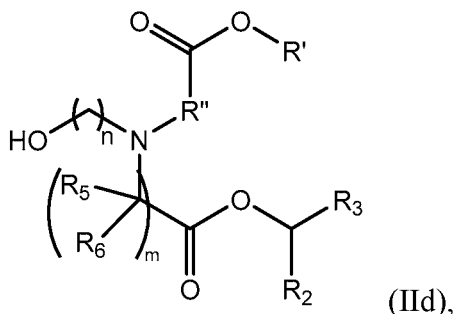
(IIIf) or their N-oxides, or salts or isomers

thereof,

wherein M is -C(O)O- or -OC(O)-, M' is C₁₋₆ alkyl or C₂₋₆ alkenyl, R₂ and R₃ are independently selected from the group consisting of C₅₋₁₄ alkyl and C₅₋₁₄ alkenyl, and n

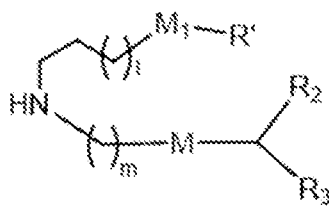
10 is selected from 2, 3, and 4.

In a further embodiment, the compounds of Formula (I) are of Formula (II_d),



or their N-oxides, or salts or isomers thereof, wherein n is 2, 3, or 4; and m, R', R'', and R₂ through R₆ are as described herein. For example, each of R₂ and R₃ may be independently selected from the group consisting of C₅₋₁₄ alkyl and C₅₋₁₄ alkenyl.

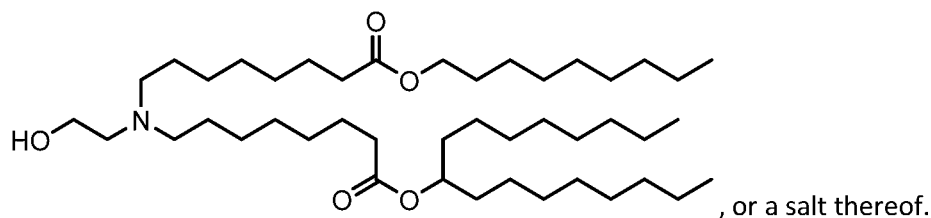
In a further embodiment, the compounds of Formula (I) are of Formula (IIg),



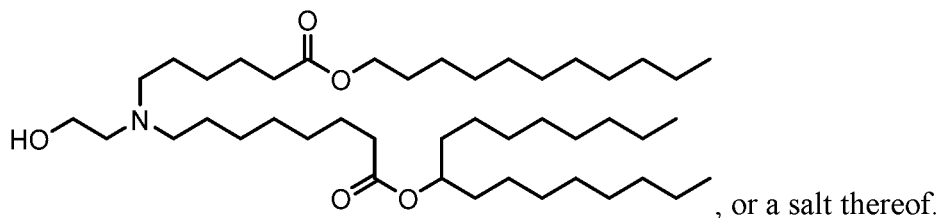
(IIg), or their N-oxides, or salts or isomers thereof, wherein l is selected from 1, 2, 3, 4, and 5; m is selected from 5, 6, 7, 8, and 9; M₁ is a bond or M'; M and M' are independently selected from -C(O)O-, -OC(O)-, -OC(O)-M''-C(O)O-, -C(O)N(R')-, -P(O)(OR')O-, -S-S-, an aryl group, and a heteroaryl group; and R₂ and R₃ are independently selected from the group consisting of H, C₁₋₁₄ alkyl, and C₂₋₁₄ alkenyl. For example, M'' is C₁₋₆ alkyl (e.g., C₁₋₄ alkyl) or C₂₋₆ alkenyl (e.g. C₂₋₄ alkenyl). For example, R₂ and R₃ are independently selected from the group consisting of C₅₋₁₄ alkyl and C₅₋₁₄ alkenyl.

In some embodiments, the amino lipids are one or more of the compounds described in U.S. Application Nos. 62/220,091, 62/252,316, 62/253,433, 62/266,460, 62/333,557, 62/382,740, 62/393,940, 62/471,937, 62/471,949, 62/475,140, and 62/475,166, and PCT Application No. PCT/US2016/052352.

In some embodiments, the amino lipid is



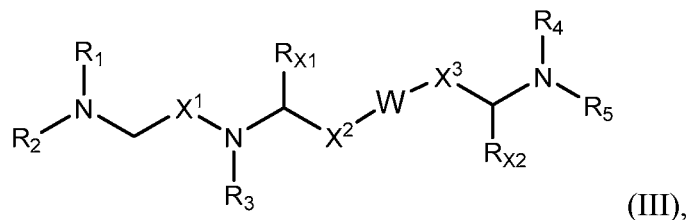
In some embodiments, the amino lipid is



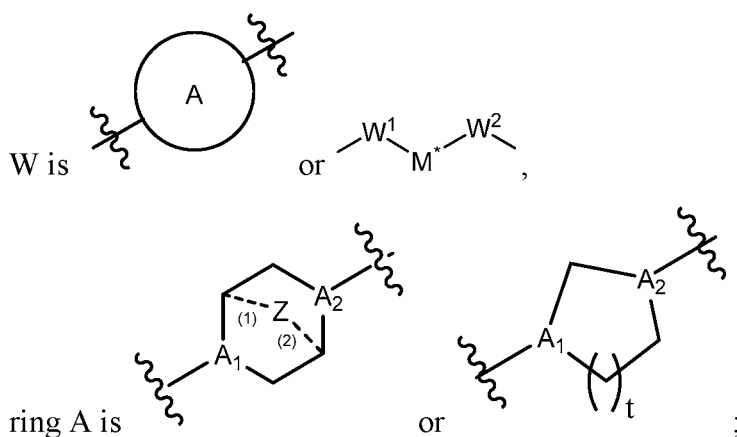
The central amine moiety of a lipid according to Formula (I), (IA), (IB), (II), (IIa), (IIb), (IIc), (IId), (IIe), (IIf), or (IIg) may be protonated at a physiological pH. Thus, a lipid may have a positive or partial positive charge at physiological pH. Such amino lipids may be referred to as cationic lipids, ionizable lipids, cationic amino lipids,

or ionizable amino lipids. Amino lipids may also be zwitterionic, i.e., neutral molecules having both a positive and a negative charge.

In some aspects, the amino lipids of the present disclosure may be one or more of compounds of formula (III),



or salts or isomers thereof, wherein



t is 1 or 2;

A₁ and A₂ are each independently selected from CH or N;

Z is CH₂ or absent wherein when Z is CH₂, the dashed lines (1) and (2) each represent a single bond; and when Z is absent, the dashed lines (1) and (2) are both absent;

R₁, R₂, R₃, R₄, and R₅ are independently selected from the group consisting of C₅₋₂₀ alkyl, C₅₋₂₀ alkenyl, -R''MR', -R''YR'', -YR'', and -R''OR'';

R_{X1} and R_{X2} are each independently H or C₁₋₃ alkyl;

each M is independently selected from the group consisting of -C(O)O-, -OC(O)-, -OC(O)O-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-,

-CH(OH)-, -P(O)(OR')O-, -S(O)₂-, -C(O)S-, -SC(O)-, an aryl group, and a heteroaryl group;

M* is C₁-C₆ alkyl,

W¹ and W² are each independently selected from the group consisting of -O- and -N(R₆)-;

each R₆ is independently selected from the group consisting of H and C₁₋₅ alkyl;

5 X¹, X², and X³ are independently selected from the group consisting of a bond, -CH₂-,

-(CH₂)₂-, -CHR-, -CHY-, -C(O)-, -C(O)O-, -OC(O)-, -(CH₂)_n-C(O)-, -C(O)-(CH₂)_n-,
-(CH₂)_n-C(O)O-, -OC(O)-(CH₂)_n-, -(CH₂)_n-OC(O)-, -C(O)O-
(CH₂)_n-, -CH(OH)-, -C(S)-, and -CH(SH)-;

10 each Y is independently a C₃₋₆ carbocycle;

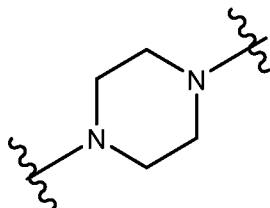
each R* is independently selected from the group consisting of C₁₋₁₂ alkyl and C₂₋₁₂ alkenyl;

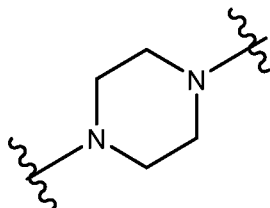
each R is independently selected from the group consisting of C₁₋₃ alkyl and a C₃₋₆ carbocycle;

15 each R' is independently selected from the group consisting of C₁₋₁₂ alkyl, C₂₋₁₂ alkenyl, and H;

each R'' is independently selected from the group consisting of C₃₋₁₂ alkyl, C₃₋₁₂ alkenyl and -R*MR'; and

n is an integer from 1-6;

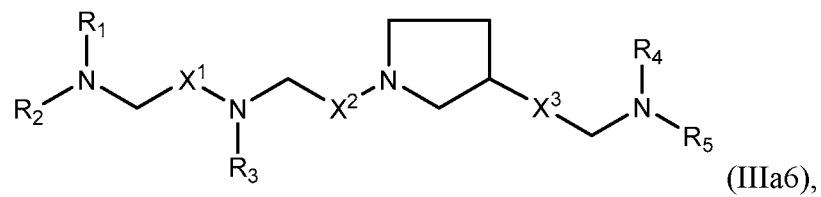
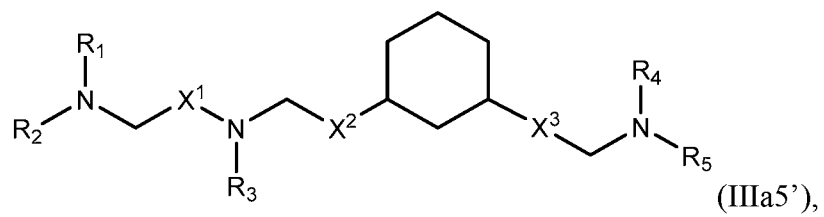
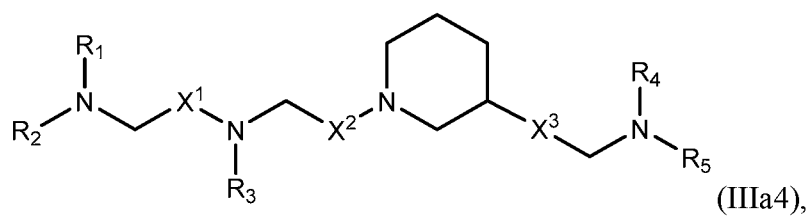
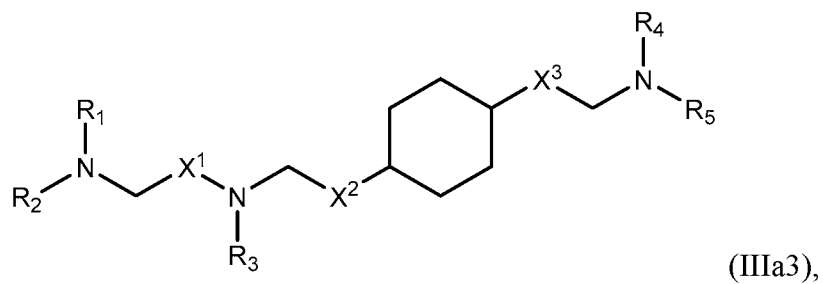
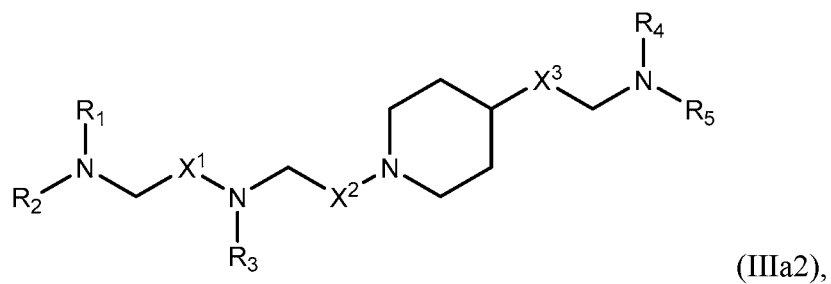
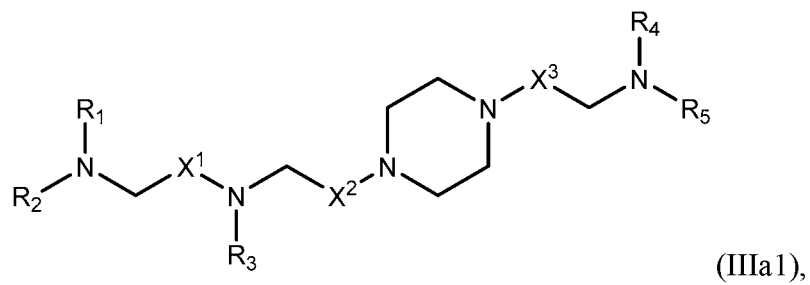


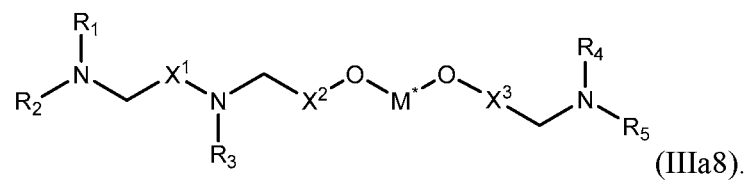
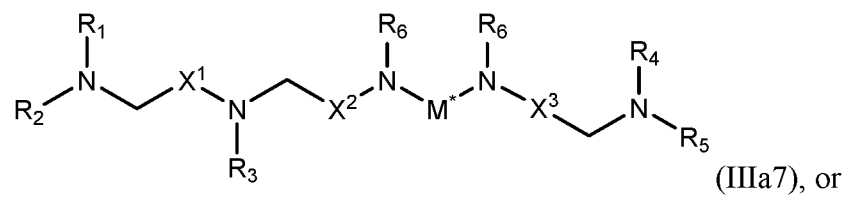
20 wherein when ring A is , then

i) at least one of X¹, X², and X³ is not -CH₂-; and/or

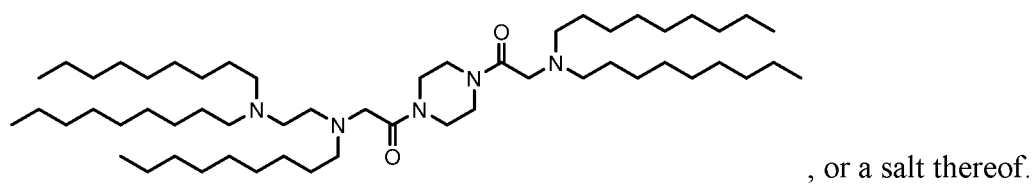
ii) at least one of R₁, R₂, R₃, R₄, and R₅ is -R''MR'.

In some embodiments, the compound is of any of formulae (IIIa1)-(IIIa8):





In some embodiments, the amino lipid is



The central amine moiety of a lipid according to Formula (III), (IIIa1), (IIIa2), (IIIa3), (IIIa4), (IIIa5), (IIIa6), (IIIa7), or (IIIa8) may be protonated at a physiological pH. Thus, a lipid may have a positive or partial positive charge at physiological pH.

Phospholipids

5 The lipid composition of the lipid nanoparticle composition disclosed herein can comprise one or more phospholipids, for example, one or more saturated or (poly)unsaturated phospholipids or a combination thereof. In general, phospholipids comprise a phospholipid moiety and one or more fatty acid moieties.

10 A phospholipid moiety can be selected, for example, from the non-limiting group consisting of phosphatidyl choline, phosphatidyl ethanolamine, phosphatidyl glycerol, phosphatidyl serine, phosphatidic acid, 2-lysophosphatidyl choline, and a sphingomyelin.

15 A fatty acid moiety can be selected, for example, from the non-limiting group consisting of lauric acid, myristic acid, myristoleic acid, palmitic acid, palmitoleic acid, stearic acid, oleic acid, linoleic acid, alpha-linolenic acid, erucic acid, phytanoic acid, arachidic acid, arachidonic acid, eicosapentaenoic acid, behenic acid, docosapentaenoic acid, and docosahexaenoic acid.

20 Particular phospholipids can facilitate fusion to a membrane. For example, a cationic phospholipid can interact with one or more negatively charged phospholipids of a membrane (e.g., a cellular or intracellular membrane). Fusion of a phospholipid to a membrane can allow one or more elements (e.g., a therapeutic agent) of a lipid-containing composition (e.g., LNPs) to pass through the membrane permitting, e.g., delivery of the one or more elements to a target tissue.

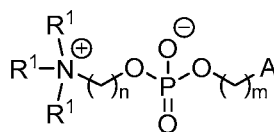
25 Non-natural phospholipid species including natural species with modifications and substitutions including branching, oxidation, cyclization, and alkynes are also contemplated. For example, a phospholipid can be functionalized with or cross-linked to one or more alkynes (e.g., an alkenyl group in which one or more double bonds is replaced with a triple bond). Under appropriate reaction conditions, an alkyne group can undergo a copper-catalyzed cycloaddition upon exposure to an azide. Such reactions
30 can be useful in functionalizing a lipid bilayer of a nanoparticle composition to facilitate

membrane permeation or cellular recognition or in conjugating a nanoparticle composition to a useful component such as a targeting or imaging moiety (e.g., a dye).

Phospholipids include, but are not limited to, glycerophospholipids such as phosphatidylcholines, phosphatidylethanolamines, phosphatidylserines,
 5 phosphatidylinositols, phosphatidyl glycerols, and phosphatidic acids. Phospholipids also include phosphosphingolipid, such as sphingomyelin.

In some embodiments, a phospholipid of the invention comprises 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-Distearoyl-sn-glycero-3-phosphoethanolamine (DSPE), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine
 10 (DOPE), 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-diundecanoyl-sn-glycero-3-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2
 15 cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemSPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine, 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16.0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine,
 20 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), sphingomyelin, and mixtures thereof.

In certain embodiments, a phospholipid useful or potentially useful in the
 25 present invention is an analog or variant of DSPC. In certain embodiments, a phospholipid useful or potentially useful in the present invention is a compound of Formula (IV):



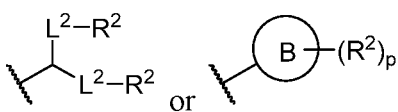
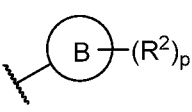
(IV),

or a salt thereof, wherein:

each R^1 is independently optionally substituted alkyl; or optionally two R^1 are joined together with the intervening atoms to form optionally substituted monocyclic carbocyclyl or optionally substituted monocyclic heterocyclyl; or optionally three R^1 are joined together with the intervening atoms to form optionally substituted bicyclic carbocyclyl or optionally substituted bicyclic heterocyclyl;

n is 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

m is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

A is of the formula:  or  ;

each instance of L^2 is independently a bond or optionally substituted C_{1-6} alkylene, wherein one methylene unit of the optionally substituted C_{1-6} alkylene is optionally replaced with O, $N(R^N)$, S, C(O), $C(O)N(R^N)$, $NR^NC(O)$, $C(O)O$, $OC(O)$, $-OC(O)O$, $OC(O)N(R^N)$, $NR^NC(O)O$, or $NR^NC(O)N(R^N)$;

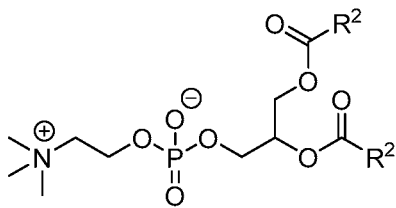
each instance of R^2 is independently optionally substituted C_{1-30} alkyl, optionally substituted C_{1-30} alkenyl, or optionally substituted C_{1-30} alkynyl; optionally wherein one or more methylene units of R^2 are independently replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, $N(R^N)$, O, S, C(O), $C(O)N(R^N)$, $NR^NC(O)$, $-NR^NC(O)N(R^N)$, $C(O)O$, $OC(O)$, $OC(O)O$, $OC(O)N(R^N)$, $NR^NC(O)O$, $C(O)S$, $SC(O)$, $C(=NR^N)$, $C(=NR^N)N(R^N)$, $NR^NC(=NR^N)$, $NR^NC(=NR^N)N(R^N)$, $C(S)$, $C(S)N(R^N)$, $-NR^NC(S)$, $NR^NC(S)N(R^N)$, $S(O)$, $OS(O)$, $S(O)O$, $OS(O)O$, $OS(O)_2$, $S(O)_2O$, $OS(O)_2O$, $N(R^N)S(O)$, $S(O)N(R^N)$, $N(R^N)S(O)N(R^N)$, $OS(O)N(R^N)$, $N(R^N)S(O)O$, $S(O)_2$, $-N(R^N)S(O)_2$, $S(O)_2N(R^N)$, $N(R^N)S(O)_2N(R^N)$, $OS(O)_2N(R^N)$, or $N(R^N)S(O)_2O$;

each instance of R^N is independently hydrogen, optionally substituted alkyl, or a nitrogen protecting group;

Ring B is optionally substituted carbocyclyl, optionally substituted heterocyclyl, optionally substituted aryl, or optionally substituted heteroaryl; and

p is 1 or 2;

provided that the compound is not of the formula:



wherein each instance of R^2 is independently unsubstituted alkyl, unsubstituted alkenyl, or unsubstituted alkynyl.

In some embodiments, the phospholipids may be one or more of the
 5 phospholipids described in U.S. Application No. 62/520,530, or in International
 Application PCT/US2018/037922 filed on 15 June 2018, the entire contents of each of
 which is hereby incorporated by reference in its entirety.

Structural Lipids

10 The lipid composition of a pharmaceutical composition disclosed herein can
 comprise one or more structural lipids. As used herein, the term “structural lipid” refers
 to sterols and also to lipids containing sterol moieties.

Incorporation of structural lipids in the lipid nanoparticle may help mitigate
 aggregation of other lipids in the particle. Structural lipids can be selected from the
 15 group including but not limited to, cholesterol, fecosterol, sitosterol, ergosterol,
 campesterol, stigmasterol, brassicasterol, tomatidine, tomatine, ursolic acid, alpha-
 tocopherol, hopanoids, phytosterols, steroids, and mixtures thereof. In some
 embodiments, the structural lipid is a sterol. As defined herein, “sterols” are a subgroup
 of steroids consisting of steroid alcohols. In certain embodiments, the structural lipid is
 20 a steroid. In certain embodiments, the structural lipid is cholesterol. In certain
 embodiments, the structural lipid is an analog of cholesterol. In certain embodiments,
 the structural lipid is alpha-tocopherol.

In some embodiments, the structural lipids may be one or more of the structural
 lipids described in U.S. Application No. 16/493,814.

25

Polyethylene Glycol (PEG)-Lipids

The lipid composition of a pharmaceutical composition disclosed herein can comprise one or more polyethylene glycol (PEG) lipids.

As used herein, the term “PEG-lipid” refers to polyethylene glycol (PEG)-modified lipids. Non-limiting examples of PEG-lipids include PEG-modified phosphatidylethanolamine and phosphatidic acid, PEG-ceramide conjugates (e.g., PEG-CerC14 or PEG-CerC20), PEG-modified dialkylamines and PEG-modified 1,2-diacyloxypropan-3-amines. Such lipids are also referred to as PEGylated lipids. For example, a PEG lipid can be PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, or a PEG-DSPE lipid.

In some embodiments, the PEG-lipid includes, but not limited to 1,2-dimyristoyl-sn-glycerol methoxypolyethylene glycol (PEG-DMG), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[amino(polyethylene glycol)] (PEG-DSPE), PEG-disteryl glycerol (PEG-DSG), PEG-dipalmitoleyl, PEG-dioleoyl, PEG-distearyl, PEG-diacylglyceramide (PEG-DAG), PEG-dipalmitoyl phosphatidylethanolamine (PEG-DPPE), or PEG-1,2-dimyristyloxypropyl-3-amine (PEG-c-DMA).

In one embodiment, the PEG-lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof. In some embodiments, the PEG-modified lipid is PEG-DMG, PEG-c-DOMG (also referred to as PEG-DOMG), PEG-DSG and/or PEG-DPG.

In some embodiments, the lipid moiety of the PEG-lipids includes those having lengths of from about C₁₄ to about C₂₂, preferably from about C₁₄ to about C₁₆. In some embodiments, a PEG moiety, for example an mPEG-NH₂, has a size of about 1000, 2000, 5000, 10,000, 15,000 or 20,000 daltons. In one embodiment, the PEG-lipid is PEG_{2k}-DMG.

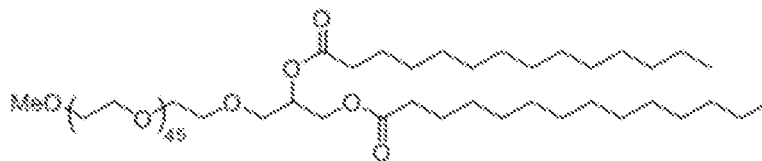
In one embodiment, the lipid nanoparticles described herein can comprise a PEG lipid which is a non-diffusible PEG. Non-limiting examples of non-diffusible PEGs include PEG-DSG and PEG-DSPE.

PEG-lipids are known in the art, such as those described in U.S. Patent No. 8158601 and International Publ. No. WO 2015/130584 A2, which are incorporated herein by reference in their entirety.

In general, some of the other lipid components (e.g., PEG lipids) of various formulae, described herein may be synthesized as described International Patent Application No. PCT/US2016/000129, filed December 10, 2016, entitled “Compositions and Methods for Delivery of Therapeutic Agents,” which is incorporated by reference in its entirety.

The lipid component of a lipid nanoparticle composition may include one or more molecules comprising polyethylene glycol, such as PEG or PEG-modified lipids. Such species may be alternately referred to as PEGylated lipids. A PEG lipid is a lipid modified with polyethylene glycol. A PEG lipid may be selected from the non-limiting group including PEG-modified phosphatidylethanolamines, PEG-modified phosphatidic acids, PEG-modified ceramides, PEG-modified dialkylamines, PEG-modified diacylglycerols, PEG-modified dialkylglycerols, and mixtures thereof. For example, a PEG lipid may be PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, or a PEG-DSPE lipid.

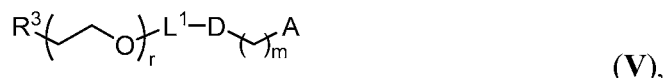
In some embodiments the PEG-modified lipids are a modified form of PEG DMG. PEG-DMG has the following structure:



In one embodiment, PEG lipids useful in the present invention can be PEGylated lipids described in International Publication No. WO2012099755, the contents of which is herein incorporated by reference in its entirety. Any of these exemplary PEG lipids described herein may be modified to comprise a hydroxyl group on the PEG chain. In certain embodiments, the PEG lipid is a PEG-OH lipid. As generally defined herein, a “PEG-OH lipid” (also referred to herein as “hydroxy-PEGylated lipid”) is a PEGylated lipid having one or more hydroxyl (–OH) groups on the lipid. In certain embodiments, the PEG-OH lipid includes one or more hydroxyl

groups on the PEG chain. In certain embodiments, a PEG-OH or hydroxy-PEGylated lipid comprises an -OH group at the terminus of the PEG chain. Each possibility represents a separate embodiment of the present invention.

In certain embodiments, a PEG lipid useful in the present invention is a compound of Formula (V). Provided herein are compounds of Formula (V):



or salts thereof, wherein:

R^3 is ---OR^O ;

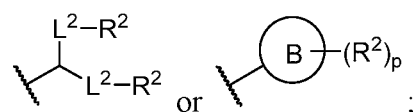
R^O is hydrogen, optionally substituted alkyl, or an oxygen protecting group;

r is an integer between 1 and 100, inclusive;

L^1 is optionally substituted C_{1-10} alkylene, wherein at least one methylene of the optionally substituted C_{1-10} alkylene is independently replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, O, $N(R^N)$, S, $C(O)$, -
 $C(O)N(R^N)$, $NR^NC(O)$, $C(O)O$, $OC(O)$, $OC(O)O$, $OC(O)N(R^N)$, $NR^NC(O)O$, or -
 $NR^NC(O)N(R^N)$;

D is a moiety obtained by click chemistry or a moiety cleavable under physiological conditions;

m is 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10;

A is of the formula:  ;

each instance of L^2 is independently a bond or optionally substituted C_{1-6} alkylene, wherein one methylene unit of the optionally substituted C_{1-6} alkylene is optionally replaced with O, $N(R^N)$, S, $C(O)$, $C(O)N(R^N)$, $NR^NC(O)$, $C(O)O$, $OC(O)$, -
 $OC(O)O$, $OC(O)N(R^N)$, $NR^NC(O)O$, or $NR^NC(O)N(R^N)$;

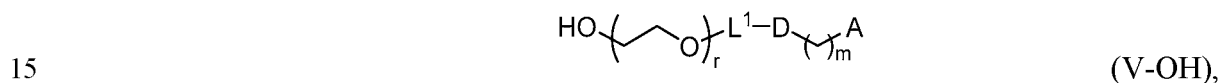
each instance of R^2 is independently optionally substituted C_{1-30} alkyl, optionally substituted C_{1-30} alkenyl, or optionally substituted C_{1-30} alkynyl; optionally wherein one or more methylene units of R^2 are independently replaced with optionally substituted carbocyclylene, optionally substituted heterocyclylene, optionally substituted arylene,

optionally substituted heteroarylene, $N(R^N)$, O, S, C(O), C(O) $N(R^N)$, $NR^N C(O)$, -
 $NR^N C(O)N(R^N)$, C(O)O, OC(O), OC(O)O, OC(O) $N(R^N)$, $NR^N C(O)O$, C(O)S, SC(O),
 C(=NR^N), C(=NR^N) $N(R^N)$, $NR^N C(=NR^N)$, $NR^N C(=NR^N)N(R^N)$, C(S), C(S) $N(R^N)$, -
 $NR^N C(S)$, $NR^N C(S)N(R^N)$, S(O), OS(O), S(O)O, OS(O)O, OS(O)₂, S(O)₂O, OS(O)₂O,
 5 $N(R^N)S(O)$, S(O) $N(R^N)$, $N(R^N)S(O)N(R^N)$, OS(O) $N(R^N)$, $N(R^N)S(O)O$, S(O)₂, -
 $N(R^N)S(O)_2$, S(O)₂ $N(R^N)$, $N(R^N)S(O)_2N(R^N)$, OS(O)₂ $N(R^N)$, or $N(R^N)S(O)_2O$;

each instance of R^N is independently hydrogen, optionally substituted alkyl, or a nitrogen protecting group;

Ring B is optionally substituted carbocyclyl, optionally substituted heterocyclyl,
 10 optionally substituted aryl, or optionally substituted heteroaryl; and
 p is 1 or 2.

In certain embodiments, the compound of Formula (V) is a PEG-OH lipid (*i.e.*, R^3 is $-OR^O$, and R^O is hydrogen). In certain embodiments, the compound of Formula (V) is of Formula (V-OH):



or a salt thereof.

In certain embodiments, a PEG lipid useful in the present invention is a PEGylated fatty acid. In certain embodiments, a PEG lipid useful in the present invention is a compound of Formula (VI). Provided herein are compounds of Formula
 20 (VI-A):



or a salts thereof, wherein:

R^3 is $-OR^O$;

R^O is hydrogen, optionally substituted alkyl or an oxygen protecting group;

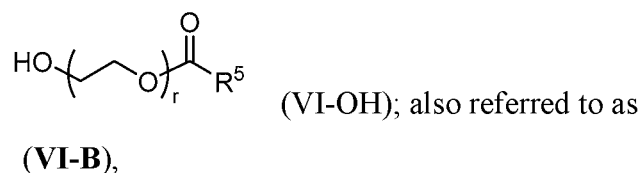
25 r is an integer between 1 and 100, inclusive;

R^5 is optionally substituted C₁₀₋₄₀ alkyl, optionally substituted C₁₀₋₄₀ alkenyl, or optionally substituted C₁₀₋₄₀ alkynyl; and optionally one or more methylene groups of R^5 are replaced with optionally substituted carbocyclylene, optionally substituted

heterocyclylene, optionally substituted arylene, optionally substituted heteroarylene, -
 $N(R^N)$, O, S, C(O), C(O) $N(R^N)$, NR^N C(O), NR^N C(O) $N(R^N)$, C(O)O, OC(O), OC(O)O,
 OC(O) $N(R^N)$, NR^N C(O)O, C(O)S, SC(O), C(=NR^N), C(=NR^N) $N(R^N)$, NR^N C(=NR^N),
 NR^N C(=NR^N) $N(R^N)$, C(S), C(S) $N(R^N)$, NR^N C(S), NR^N C(S) $N(R^N)$, S(O), OS(O), -
 5 S(O)O, OS(O)O, OS(O)₂, S(O)₂O, OS(O)₂O, $N(R^N)$ S(O), S(O) $N(R^N)$, -
 $N(R^N)$ S(O) $N(R^N)$, OS(O) $N(R^N)$, $N(R^N)$ S(O)O, S(O)₂, $N(R^N)$ S(O)₂, S(O)₂ $N(R^N)$, -
 $N(R^N)$ S(O)₂ $N(R^N)$, OS(O)₂ $N(R^N)$, or $N(R^N)$ S(O)₂O; and

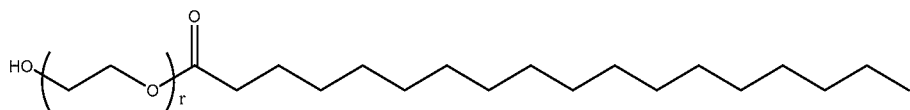
each instance of R^N is independently hydrogen, optionally substituted alkyl, or a
 nitrogen protecting group.

10 In certain embodiments, the compound of Formula (VI) is of Formula (VI-OH):



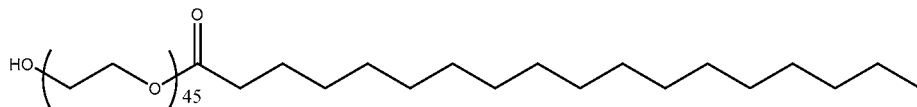
or a salt thereof. In some embodiments, r is 40-50.

15 In yet other embodiments the compound of Formula (VI-C) is:



or a salt thereof.

In one embodiment, the compound of Formula (VI-D) is



20 In some aspects, the lipid composition of the pharmaceutical compositions disclosed herein does not comprise a PEG-lipid.

In some embodiments, the PEG-lipids may be one or more of the PEG lipids described in U.S. Application No. US15/674,872.

In some embodiments, a LNP of the invention comprises an amino lipid of any
 25 of Formula I, II or III, a phospholipid comprising DSPC, a structural lipid, and a PEG lipid comprising PEG-DMG.

In some embodiments, a LNP of the invention comprises an amino lipid of any of Formula I, II or III, a phospholipid comprising DSPC, a structural lipid, and a PEG lipid comprising a compound having Formula VI.

5 In some embodiments, a LNP of the invention comprises an amino lipid of Formula I, II or III, a phospholipid comprising a compound having Formula IV, a structural lipid, and the PEG lipid comprising a compound having Formula V or VI.

In some embodiments, a LNP of the invention comprises an amino lipid of Formula I, II or III, a phospholipid comprising a compound having Formula IV, a structural lipid, and the PEG lipid comprising a compound having Formula V or VI.

10 In some embodiments, a LNP of the invention comprises an amino lipid of Formula I, II or III, a phospholipid having Formula IV, a structural lipid, and a PEG lipid comprising a compound having Formula VI.

In some embodiments, a LNP of the invention comprises an N:P ratio of from about 2:1 to about 30:1.

15 In some embodiments, a LNP of the invention comprises an N:P ratio of about 6:1.

In some embodiments, a LNP of the invention comprises an N:P ratio of about 3:1, 4:1, or 5:1.

20 In some embodiments, a LNP of the invention comprises a wt/wt ratio of the amino lipid component to the RNA of from about 10:1 to about 100:1.

In some embodiments, a LNP of the invention comprises a wt/wt ratio of the amino lipid component to the RNA of about 20:1.

In some embodiments, a LNP of the invention comprises a wt/wt ratio of the amino lipid component to the RNA of about 10:1.

25 In some embodiments, a LNP of the invention has a mean diameter from about 30nm to about 150nm.

In some embodiments, a LNP of the invention has a mean diameter from about 60nm to about 120nm.

Exemplary Additional LNP Components

Surfactants

5 In certain embodiments, the lipid nanoparticles of the disclosure optionally includes one or more surfactants.

 In certain embodiments, the surfactant is an amphiphilic polymer. As used herein, an amphiphilic “polymer” is an amphiphilic compound that comprises an oligomer or a polymer. For example, an amphiphilic polymer can comprise an
10 oligomer fragment, such as two or more PEG monomer units. For example, an amphiphilic polymer described herein can be PS 20.

 For example, the amphiphilic polymer is a block copolymer.

 For example, the amphiphilic polymer is a lyoprotectant.

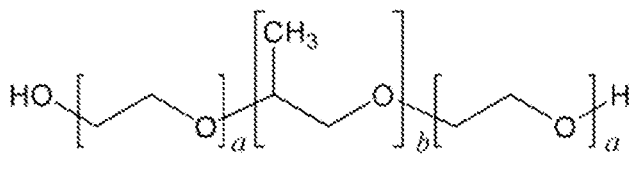
 For example, amphiphilic polymer has a critical micelle concentration (CMC) of
15 less than 2×10^{-4} M in water at about 30 °C and atmospheric pressure.

 For example, amphiphilic polymer has a critical micelle concentration (CMC) ranging between about 0.1×10^{-4} M and about 1.3×10^{-4} M in water at about 30 °C and atmospheric pressure.

 For example, the concentration of the amphiphilic polymer ranges between
20 about its CMC and about 30 times of CMC (e.g., up to about 25 times, about 20 times, about 15 times, about 10 times, about 5 times, or about 3 times of its CMC) in the formulation, e.g., prior to freezing or lyophilization.

 For example, the amphiphilic polymer is selected from poloxamers (Pluronic®), poloxamines (Tetronic®), polyoxyethylene glycol sorbitan alkyl esters (polysorbates)
25 and polyvinyl pyrrolidones (PVPs).

 For example, the amphiphilic polymer is a poloxamer. For example, the amphiphilic polymer is of the following structure:



wherein a is an integer between 10 and 150 and b is an integer between 20 and 60. For example, a is about 12 and b is about 20, or a is about 80 and b is about 27, or a is about 64 and b is about 37, or a is about 141 and b is about 44, or a is about 101 and b is about

5 56.

For example, the amphiphilic polymer is P124, P188, P237, P338, or P407.

For example, the amphiphilic polymer is P188 (e.g., Poloxamer 188, CAS Number 9003-11-6, also known as Kolliphor P188).

For example, the amphiphilic polymer is a poloxamine, e.g., tetronic 304 or
10 tetronic 904.

For example, the amphiphilic polymer is a polyvinylpyrrolidone (PVP), such as PVP with molecular weight of 3 kDa, 10 kDa, or 29 kDa.

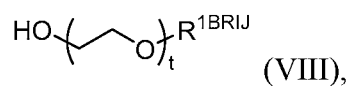
For example, the amphiphilic polymer is a polysorbate, such as PS 20.

In certain embodiments, the surfactant is a non-ionic surfactant.

15 In some embodiments, the lipid nanoparticle comprises a surfactant. In some embodiments, the surfactant is an amphiphilic polymer. In some embodiments, the surfactant is a non-ionic surfactant.

For example, the non-ionic surfactant is selected from the group consisting of polyethylene glycol ether (Brij), poloxamer, polysorbate, sorbitan, and derivatives
20 thereof.

For example, the polyethylene glycol ether is a compound of Formula (VIII):



or a salt or isomer thereof, wherein:

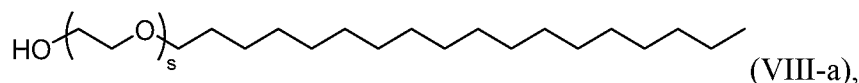
t is an integer between 1 and 100;

25 R^{1BRIJ} independently is C₁₀-40 alkyl, C₁₀-40 alkenyl, or C₁₀-40 alkynyl; and optionally one or more methylene groups of R^{5PEG} are independently replaced with C₃-10 carbocyclylene, 4 to 10 membered heterocyclylene, C₆-10 arylene, 4 to 10

membered heteroarylene, $-N(RN)-$, $-O-$, $-S-$, $-C(O)-$, $-C(O)N(RN)-$, $-NRNC(O)-$, $-NRNC(O)N(RN)-$, $-C(O)O-$, $-OC(O)-$, $-OC(O)O-$, $-OC(O)N(RN)-$, $-NRNC(O)O-$, $-C(O)S-$, $-SC(O)-$, $-C(=NRN)-$, $-C(=NRN)N(RN)-$, $-NRNC(=NRN)-$, $-NRNC(=NRN)N(RN)-$, $-C(S)-$, $-C(S)N(RN)-$, $-NRNC(S)-$, $-NRNC(S)N(RN)-$, $-S(O)-$, $-OS(O)-$, $-S(O)O-$, $-OS(O)O-$, $-OS(O)2-$, $-S(O)2O-$, $-OS(O)2O-$, $-N(RN)S(O)-$, $-S(O)N(RN)-$, $-N(RN)S(O)N(RN)-$, $-OS(O)N(RN)-$, $-N(RN)S(O)O-$, $-S(O)2-$, $-N(RN)S(O)2-$, $-S(O)2N(RN)-$, $-N(RN)S(O)2N(RN)-$, $-OS(O)2N(RN)-$, or $-N(RN)S(O)2O-$; and

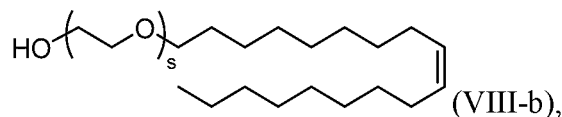
each instance of RN is independently hydrogen, C1-6 alkyl, or a nitrogen protecting group

In some embodiment, R1BRIJ is C18 alkyl. For example, the polyethylene glycol ether is a compound of Formula (VIII-a):



or a salt or isomer thereof.

In some embodiments, R1BRIJ is C18 alkenyl. For example, the polyethylene glycol ether is a compound of Formula (VIII-b):



or a salt or isomer thereof

In some embodiments, the poloxamer is selected from the group consisting of poloxamer 101, poloxamer 105, poloxamer 108, poloxamer 122, poloxamer 123, poloxamer 124, poloxamer 181, poloxamer 182, poloxamer 183, poloxamer 184, poloxamer 185, poloxamer 188, poloxamer 212, poloxamer 215, poloxamer 217, poloxamer 231, poloxamer 234, poloxamer 235, poloxamer 237, poloxamer 238, poloxamer 282, poloxamer 284, poloxamer 288, poloxamer 331, poloxamer 333, poloxamer 334, poloxamer 335, poloxamer 338, poloxamer 401, poloxamer 402, poloxamer 403, and poloxamer 407.

In some embodiments, the polysorbate is Tween® 20, Tween® 40, Tween®, 60, or Tween® 80.

In some embodiments, the derivative of sorbitan is Span® 20, Span® 60, Span® 65, Span® 80, or Span® 85.

In some embodiments, the concentration of the non-ionic surfactant in the lipid nanoparticle ranges from about 0.00001 % w/v to about 1 % w/v, e.g., from about
5 0.00005 % w/v to about 0.5 % w/v, or from about 0.0001 % w/v to about 0.1 % w/v.

In some embodiments, the concentration of the non-ionic surfactant in lipid nanoparticle ranges from about 0.000001 wt% to about 1 wt%, e.g., from about 0.000002 wt% to about 0.8 wt%, or from about 0.000005 wt% to about 0.5 wt%.

In some embodiments, the concentration of the PEG lipid in the lipid
10 nanoparticle ranges from about 0.01 % by molar to about 50 % by molar, e.g., from about 0.05 % by molar to about 20 % by molar, from about 0.07 % by molar to about 10 % by molar, from about 0.1 % by molar to about 8 % by molar, from about 0.2 % by molar to about 5 % by molar, or from about 0.25 % by molar to about 3 % by molar.

15 **Adjuvants**

In some embodiments, an LNP of the invention optionally includes one or more adjuvants, e.g., Glucopyranosyl Lipid Adjuvant (GLA), CpG oligodeoxynucleotides (e.g., Class A or B), poly(I:C), aluminum hydroxide, and Pam3CSK4.

20 **Other components**

An LNP of the invention may optionally include one or more components in addition to those described in the preceding sections. For example, a lipid nanoparticle may include one or more small hydrophobic molecules such as a vitamin (e.g., vitamin A or vitamin E) or a sterol.

25 Lipid nanoparticles may also include one or more permeability enhancer molecules, carbohydrates, polymers, surface altering agents, or other components. A permeability enhancer molecule may be a molecule described by U.S. patent application publication No. 2005/0222064, for example. Carbohydrates may include simple sugars (e.g., glucose) and polysaccharides (e.g., glycogen and derivatives and analogs thereof).

A polymer may be included in and/or used to encapsulate or partially encapsulate a lipid nanoparticle. A polymer may be biodegradable and/or biocompatible. A polymer may be selected from, but is not limited to, polyamines, polyethers, polyamides, polyesters, polycarbamates, polyureas, polycarbonates, polystyrenes, polyimides, polysulfones, polyurethanes, polyacetylenes, polyethylenes, polyethyleneimines, polyisocyanates, polyacrylates, polymethacrylates, polyacrylonitriles, and polyarylates. For example, a polymer may include poly(caprolactone) (PCL), ethylene vinyl acetate polymer (EVA), poly(lactic acid) (PLA), poly(L-lactic acid) (PLLA), poly(glycolic acid) (PGA), poly(lactic acid-co-glycolic acid) (PLGA), poly(L-lactic acid-co-glycolic acid) (PLLGA), poly(D,L-lactide) (PDLA), poly(L-lactide) (PLLA), poly(D,L-lactide-co-caprolactone), poly(D,L-lactide-co-caprolactone-co-glycolide), poly(D,L-lactide-co-PEO-co-D,L-lactide), poly(D,L-lactide-co-PPO-co-D,L-lactide), polyalkyl cyanoacrylate, polyurethane, poly-L-lysine (PLL), hydroxypropyl methacrylate (HPMA), polyethyleneglycol, poly-L-glutamic acid, poly(hydroxy acids), polyanhydrides, polyorthoesters, poly(ester amides), polyamides, poly(ester ethers), polycarbonates, polyalkylenes such as polyethylene and polypropylene, polyalkylene glycols such as poly(ethylene glycol) (PEG), polyalkylene oxides (PEO), polyalkylene terephthalates such as poly(ethylene terephthalate), polyvinyl alcohols (PVA), polyvinyl ethers, polyvinyl esters such as poly(vinyl acetate), polyvinyl halides such as poly(vinyl chloride) (PVC), polyvinylpyrrolidone (PVP), polysiloxanes, polystyrene, polyurethanes, derivatized celluloses such as alkyl celluloses, hydroxyalkyl celluloses, cellulose ethers, cellulose esters, nitro celluloses, hydroxypropylcellulose, carboxymethylcellulose, polymers of acrylic acids, such as poly(methyl(meth)acrylate) (PMMA), poly(ethyl(meth)acrylate), poly(butyl(meth)acrylate), poly(isobutyl(meth)acrylate), poly(hexyl(meth)acrylate), poly(isodecyl(meth)acrylate), poly(lauryl(meth)acrylate), poly(phenyl(meth)acrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly(isobutyl acrylate), poly(octadecyl acrylate) and copolymers and mixtures thereof, polydioxanone and its copolymers, polyhydroxyalkanoates, polypropylene fumarate, polyoxymethylene, poloxamers, poloxamines, poly(ortho)esters, poly(butyric acid), poly(valeric acid), poly(lactide-co-

caprolactone), trimethylene carbonate, poly(N-acryloylmorpholine) (PACM), poly(2-methyl-2-oxazoline) (PMOX), poly(2-ethyl-2-oxazoline) (PEOZ), and polyglycerol.

Surface altering agents may include, but are not limited to, anionic proteins (e.g., bovine serum albumin), surfactants (e.g., cationic surfactants such as
5 dimethyldioctadecyl-ammonium bromide), sugars or sugar derivatives (e.g., cyclodextrin), nucleic acids, polymers (e.g., heparin, polyethylene glycol, and poloxamer), mucolytic agents (e.g., acetylcysteine, mugwort, bromelain, papain, clerodendrum, bromhexine, carbocysteine, eprazinone, mesna, ambroxol, sobrerol, domiodol, letosteine, stepronin, tiopronin, gelsolin, thymosin β 4, dornase alfa,
10 neltenexine, and erdosteine), and DNases (e.g., rhDNase). A surface altering agent may be disposed within a nanoparticle and/or on the surface of a LNP (e.g., by coating, adsorption, covalent linkage, or other process).

A lipid nanoparticle may also comprise one or more functionalized lipids. For example, a lipid may be functionalized with an alkyne group that, when exposed to an
15 azide under appropriate reaction conditions, may undergo a cycloaddition reaction. In particular, a lipid bilayer may be functionalized in this fashion with one or more groups useful in facilitating membrane permeation, cellular recognition, or imaging. The surface of a LNP may also be conjugated with one or more useful antibodies. Functional groups and conjugates useful in targeted cell delivery, imaging, and
20 membrane permeation are well known in the art.

In addition to these components, lipid nanoparticles may include any substance useful in pharmaceutical compositions. For example, the lipid nanoparticle may include one or more pharmaceutically acceptable excipients or accessory ingredients such as, but not limited to, one or more solvents, dispersion media, diluents, dispersion aids,
25 suspension aids, granulating aids, disintegrants, fillers, glidants, liquid vehicles, binders, surface active agents, isotonic agents, thickening or emulsifying agents, buffering agents, lubricating agents, oils, preservatives, and other species. Excipients such as waxes, butters, coloring agents, coating agents, flavorings, and perfuming agents may also be included. Pharmaceutically acceptable excipients are well known in the art (see
30 for example Remington's The Science and Practice of Pharmacy, 21st Edition, A. R. Gennaro; Lippincott, Williams & Wilkins, Baltimore, MD, 2006).

Examples of diluents may include, but are not limited to, calcium carbonate, sodium carbonate, calcium phosphate, dicalcium phosphate, calcium sulfate, calcium hydrogen phosphate, sodium phosphate lactose, sucrose, cellulose, microcrystalline cellulose, kaolin, mannitol, sorbitol, inositol, sodium chloride, dry starch, cornstarch, powdered sugar, and/or combinations thereof. Granulating and dispersing agents may be selected from the non-limiting list consisting of potato starch, corn starch, tapioca starch, sodium starch glycolate, clays, alginic acid, guar gum, citrus pulp, agar, bentonite, cellulose and wood products, natural sponge, cation-exchange resins, calcium carbonate, silicates, sodium carbonate, cross-linked poly(vinyl-pyrrolidone) (crospovidone), sodium carboxymethyl starch (sodium starch glycolate), carboxymethyl cellulose, cross-linked sodium carboxymethyl cellulose (croscarmellose), methylcellulose, pregelatinized starch (starch 1500), microcrystalline starch, water insoluble starch, calcium carboxymethyl cellulose, magnesium aluminum silicate (VEEGUM®), sodium lauryl sulfate, quaternary ammonium compounds, and/or combinations thereof.

Surface active agents and/or emulsifiers may include, but are not limited to, natural emulsifiers (e.g., acacia, agar, alginic acid, sodium alginate, tragacanth, chondrux, cholesterol, xanthan, pectin, gelatin, egg yolk, casein, wool fat, cholesterol, wax, and lecithin), colloidal clays (e.g., bentonite [aluminum silicate] and VEEGUM® [magnesium aluminum silicate]), long chain amino acid derivatives, high molecular weight alcohols (e.g., stearyl alcohol, cetyl alcohol, oleyl alcohol, triacetin monostearate, ethylene glycol distearate, glyceryl monostearate, and propylene glycol monostearate, polyvinyl alcohol), carbomers (e.g., carboxy polymethylene, polyacrylic acid, acrylic acid polymer, and carboxyvinyl polymer), carrageenan, cellulosic derivatives (e.g., carboxymethylcellulose sodium, powdered cellulose, hydroxymethyl cellulose, hydroxypropyl cellulose, hydroxypropyl methylcellulose, methylcellulose), sorbitan fatty acid esters (e.g., polyoxyethylene sorbitan monolaurate [TWEEN®20], polyoxyethylene sorbitan [TWEEN® 60], polyoxyethylene sorbitan monooleate [TWEEN®80], sorbitan monopalmitate [SPAN®40], sorbitan monostearate [SPAN®60], sorbitan tristearate [SPAN®65], glyceryl monooleate, sorbitan monooleate [SPAN®80]), polyoxyethylene esters (e.g., polyoxyethylene monostearate

[MYRJ® 45], polyoxyethylene hydrogenated castor oil, polyethoxylated castor oil, polyoxymethylene stearate, and SOLUTOL®), sucrose fatty acid esters, polyethylene glycol fatty acid esters (e.g., CREMOPHOR®), polyoxyethylene ethers, (e.g., polyoxyethylene lauryl ether [BRIJ® 30]), poly(vinyl-pyrrolidone), diethylene glycol monolaurate, triethanolamine oleate, sodium oleate, potassium oleate, ethyl oleate, oleic acid, ethyl laurate, sodium lauryl sulfate, PLURONIC®F 68, POLOXAMER® 188, cetrimonium bromide, cetylpyridinium chloride, benzalkonium chloride, docusate sodium, and/or combinations thereof.

A binding agent may be starch (e.g., cornstarch and starch paste); gelatin; sugars (e.g., sucrose, glucose, dextrose, dextrin, molasses, lactose, lactitol, mannitol); natural and synthetic gums (e.g., acacia, sodium alginate, extract of Irish moss, panwar gum, ghatti gum, mucilage of isapol husks, carboxymethylcellulose, methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropyl cellulose, hydroxypropyl methylcellulose, microcrystalline cellulose, cellulose acetate, poly(vinyl-pyrrolidone), magnesium aluminum silicate (VEEGUM®), and larch arabogalactan); alginates; polyethylene oxide; polyethylene glycol; inorganic calcium salts; silicic acid; polymethacrylates; waxes; water; alcohol; and combinations thereof, or any other suitable binding agent.

Examples of preservatives may include, but are not limited to, antioxidants, chelating agents, antimicrobial preservatives, antifungal preservatives, alcohol preservatives, acidic preservatives, and/or other preservatives. Examples of antioxidants include, but are not limited to, alpha tocopherol, ascorbic acid, ascorbyl palmitate, butylated hydroxyanisole, butylated hydroxytoluene, monothioglycerol, potassium metabisulfite, propionic acid, propyl gallate, sodium ascorbate, sodium bisulfite, sodium metabisulfite, and/or sodium sulfite. Examples of chelating agents include ethylenediaminetetraacetic acid (EDTA), citric acid monohydrate, disodium edetate, dipotassium edetate, edetic acid, fumaric acid, malic acid, phosphoric acid, sodium edetate, tartaric acid, and/or trisodium edetate. Examples of antimicrobial preservatives include, but are not limited to, benzalkonium chloride, benzethonium chloride, benzyl alcohol, bronopol, cetrimide, cetylpyridinium chloride, chlorhexidine, chlorobutanol, chlorocresol, chloroxylenol, cresol, ethyl alcohol, glycerin, hexetidine, imidurea, phenol, phenoxyethanol, phenylethyl alcohol, phenylmercuric nitrate,

propylene glycol, and/or thimerosal. Examples of antifungal preservatives include, but are not limited to, butyl paraben, methyl paraben, ethyl paraben, propyl paraben, benzoic acid, hydroxybenzoic acid, potassium benzoate, potassium sorbate, sodium benzoate, sodium propionate, and/or sorbic acid. Examples of alcohol preservatives include, but are not limited to, ethanol, polyethylene glycol, benzyl alcohol, phenol, phenolic compounds, bisphenol, chlorobutanol, hydroxybenzoate, and/or phenylethyl alcohol. Examples of acidic preservatives include, but are not limited to, vitamin A, vitamin C, vitamin E, beta-carotene, citric acid, acetic acid, dehydroascorbic acid, ascorbic acid, sorbic acid, and/or phytic acid. Other preservatives include, but are not limited to, tocopherol, tocopherol acetate, deteroxime mesylate, cetrimide, butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), ethylenediamine, sodium lauryl sulfate (SLS), sodium lauryl ether sulfate (SLES), sodium bisulfite, sodium metabisulfite, potassium sulfite, potassium metabisulfite, GLYDANT PLUS®, PHENONIP®, methylparaben, GERMALL® 115, GERMABEN®II, NEOLONE™, KATHON™, and/or EUXYL®.

Examples of buffering agents include, but are not limited to, citrate buffer solutions, acetate buffer solutions, phosphate buffer solutions, ammonium chloride, calcium carbonate, calcium chloride, calcium citrate, calcium gluconate, calcium gluceptate, calcium gluconate, d-gluconic acid, calcium glycerophosphate, calcium lactate, calcium lactobionate, propanoic acid, calcium levulinate, pentanoic acid, dibasic calcium phosphate, phosphoric acid, tribasic calcium phosphate, calcium hydroxide phosphate, potassium acetate, potassium chloride, potassium gluconate, potassium mixtures, dibasic potassium phosphate, monobasic potassium phosphate, potassium phosphate mixtures, sodium acetate, sodium bicarbonate, sodium chloride, sodium citrate, sodium lactate, dibasic sodium phosphate, monobasic sodium phosphate, sodium phosphate mixtures, tromethamine, amino-sulfonate buffers (e.g., HEPES), magnesium hydroxide, aluminum hydroxide, alginic acid, pyrogen-free water, isotonic saline, Ringer's solution, ethyl alcohol, and/or combinations thereof. Lubricating agents may be selected from the non-limiting group consisting of magnesium stearate, calcium stearate, stearic acid, silica, talc, malt, glyceryl behenate, hydrogenated vegetable oils, polyethylene

glycol, sodium benzoate, sodium acetate, sodium chloride, leucine, magnesium lauryl sulfate, sodium lauryl sulfate, and combinations thereof.

Examples of oils include, but are not limited to, almond, apricot kernel, avocado, babassu, bergamot, black current seed, borage, cade, camomile, canola, caraway, carnauba, castor, cinnamon, cocoa butter, coconut, cod liver, coffee, corn, cotton seed, emu, eucalyptus, evening primrose, fish, flaxseed, geraniol, gourd, grape seed, hazel nut, hyssop, isopropyl myristate, jojoba, kukui nut, lavandin, lavender, lemon, litsea cubeba, macademia nut, mallow, mango seed, meadowfoam seed, mink, nutmeg, olive, orange, orange roughy, palm, palm kernel, peach kernel, peanut, poppy seed, pumpkin seed, rapeseed, rice bran, rosemary, safflower, sandalwood, sasquana, savoury, sea buckthorn, sesame, shea butter, silicone, soybean, sunflower, tea tree, thistle, tsubaki, vetiver, walnut, and wheat germ oils as well as butyl stearate, caprylic triglyceride, capric triglyceride, cyclomethicone, diethyl sebacate, dimethicone 360, simethicone, isopropyl myristate, mineral oil, octyldodecanol, oleyl alcohol, silicone oil, and/or combinations thereof.

Methods of using the systems or LNP compositions

The present disclosure provides LNP compositions, which can be delivered to cells, e.g., target cells, e.g., in vitro or in vivo. For in vitro protein expression, the cell is contacted with the LNP by incubating the LNP and the cell ex vivo. Such cells may subsequently be introduced in vivo. For in vivo protein expression, the cell is contacted with the LNP by administering the LNP to a subject to thereby increase or induce protein expression in or on the cells within the subject. For example, in one embodiment, the LNP is administered intravenously. In another embodiment, the LNP is administered intramuscularly. In yet other embodiment, the LNP is administered by a route selected from the group consisting of subcutaneously, intranodally and intratumorally.

For in vitro delivery, in one embodiment the cell is contacted with the LNP by incubating the LNP and the target cell ex vivo. In one embodiment, the cell is a human cell. Various types of cells have been demonstrated to be transfectable by the LNP.

In another embodiment, the cell is contacted with the LNP for, e.g., at least 30 minutes, at least 1 hour, at least 2 hours, at least 3 hours, at least 4 hours, at least 5 hours, at least 6 hours, at least 12 hours or at least 24 hours.

In one embodiment, the cell is contacted with the LNP for a single
5 treatment/transfection. In another embodiment, the cell is contacted with the LNP for multiple treatments/transfections (e.g., two, three, four or more treatments/transfections of the same cells).

In another embodiment, for in vivo delivery, the cell is contacted with the LNP by administering the LNP to a subject to thereby deliver the nucleic acid to cells within
10 the subject. For example, in one embodiment, the LNP is administered intravenously. In another embodiment, the LNP is administered intramuscularly. In yet other embodiment, the LNP is administered by a route selected from the group consisting of subcutaneously, intranodally and intratumorally.

In an aspect, provided herein is a method of increasing expression of a
15 therapeutic payload or prophylactic payload in a cell, comprising administering to the cell a system, or LNP composition disclosed herein.

In a related aspect, provided herein is a system or LNP composition for use in a method of increasing expression of a therapeutic payload or prophylactic payload in a cell.

20 In another aspect, the disclosure provides a method of increasing expression of a therapeutic payload or prophylactic payload, in a subject, comprising administering to the subject an effective amount of a system or LNP composition disclosed herein.

In a related aspect, provided herein is a system or LNP composition for use in a method of increasing expression of a therapeutic payload or prophylactic payload in a
25 subject.

In yet another aspect, provided herein is a method of delivering a system, or LNP composition disclosed herein.

In a related aspect, provided herein is a system or LNP composition for use in a
30 method of delivering the system or LNP composition to a cell.

In an embodiment, the method or use, comprises contacting the cell in vitro, in vivo or ex vivo with the system or LNP composition.

In an embodiment, the LNP compositions or systems formulated as LNPs of the present disclosure are contacted with cells, e.g., ex vivo or in vivo and can be used to
5 deliver a secreted polypeptide, an intracellular polypeptide or a transmembrane polypeptide to a subject.

In an aspect, the disclosure provides a method of delivering a system or LNP composition disclosed herein to a subject having a disease or disorder, e.g., as described
10 herein.

In a related aspect, provided herein is a system or LNP composition for use in a method of delivering the system or LNP composition to a subject having a disease or disorder, e.g., as described herein.

15 In another aspect, provided herein is a method of modulating an immune response in a subject, comprising administering to the subject in need thereof an effective amount of a system, or LNP composition disclosed herein.

In a related aspect, provided herein is a system or LNP composition for use in a method of modulating an immune response in a subject, comprising administering to
20 the subject an effective amount of the system, or LNP composition.

In another aspect, provided herein is a method of delivering a secreted polypeptide, an intracellular polypeptide or a transmembrane polypeptide to a subject.

25 In an aspect, provided herein is a method of treating, preventing, or preventing a symptom of, a disease or disorder comprising administering to a subject in need thereof an effective amount of a system, or LNP composition disclosed herein.

In a related aspect, provided herein is a system or LNP composition for use in a method of treating, preventing, or preventing a symptom of, a disease or disorder in a
30 subject, comprising administering to the subject in need thereof an effective amount of the system, or LNP composition.

In an embodiment, the first polynucleotide and/or the second polynucleotide of the system is formulated as an LNP. In an embodiment, the first polynucleotide of the system is formulated as an LNP. In an embodiment, the second polynucleotide of the system is formulated as an LNP. In an embodiment, both the first and the second
5 polynucleotides of the system are formulated as LNPs.

In an embodiment, the LNP comprising the first polynucleotide is the same as the LNP comprising the second polynucleotide. In an embodiment, the LNP comprising the first polynucleotide is different from the LNP comprising the second polynucleotide.

In an embodiment, the LNP comprising the first polynucleotide is in a
10 composition. In an embodiment, the LNP comprising the second polynucleotide is in a separate composition. In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are in the same composition.

In an embodiment, the LNP comprising the first polynucleotide and the LNP
15 comprising the second polynucleotide are administered simultaneously, e.g., substantially simultaneously. In some embodiments, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are co-delivered.

In an embodiment, the LNP comprising the first polynucleotide and the LNP comprising the second polynucleotide are administered sequentially.

20 In an embodiment, the LNP comprising the first polynucleotide is administered first.

In an embodiment, the LNP comprising the first polynucleotide is administered first followed by administration of the LNP comprising the second polynucleotide.

In an embodiment, the LNP comprising the second polynucleotide is
25 administered first.

In an embodiment, the LNP comprising the second polynucleotide is administered first followed by administration of the LNP comprising the first polynucleotide.

In some embodiments, any of the methods or composition for use disclosed herein, results in one, two, three, four, five, six or all, or any combination thereof, of the following in a cell (e.g., in a cell contacted with the system or LNP composition):

- 5 (i) increased expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (ii) sustained expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (iii) increased expression and/or level of therapeutic payload or prophylactic payload;
- 10 (iv) sustained expression and/or level of therapeutic payload or prophylactic payload;
- (v) increased stability of mRNA encoding the therapeutic payload or prophylactic payload;
- (vi) increased resistance of translation of therapeutic payload or prophylactic payload to cellular environment, e.g., stress or nutrient deprivation or translation factor availability;
- 15 (vii) reduced dosing of the therapeutic payload or prophylactic payload; or
- (viii) reduced toxicity, e.g., reduced modulation of a protein translated from endogenous mRNA in a cell.

20 In some embodiments, the methods or composition for use result in an increased expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload.

In some embodiments, the methods or composition for use result in sustained expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload.

25 In some embodiments, the methods or composition for use result in increased expression and/or level of therapeutic payload or prophylactic payload.

In some embodiments, the methods or composition for use result in sustained expression and/or level of therapeutic payload or prophylactic payload.

30 In some embodiments, the methods or composition for use result in increased stability of mRNA encoding the therapeutic payload or prophylactic payload;

In some embodiments, the methods or composition for use result in increased resistance of translation of therapeutic payload or prophylactic payload to cellular environment, e.g., stress or nutrient deprivation or translation factor availability.

In some embodiments, the methods or composition for use result in reduced dosing of the therapeutic payload or prophylactic payload.

In some embodiments, the methods or composition for use result in reduced toxicity, e.g., reduced modulation of a protein translated from endogenous mRNA in a cell.

In some embodiments, any one, or all of (i)-(vii) is compared to a cell which:

- (a) has not been contacted with the system disclosed herein;
- (b) has not been contacted with the LNP composition disclosed herein;
- (c) has not been contacted with an LNP comprising the first polynucleotide; or
- (d) has been contacted with an LNP comprising the first polynucleotide but has not been contacted with the second polynucleotide, e.g., an LNP comprising the second polynucleotide.

Combination therapies

In some embodiments, the methods of treatment or compositions for use disclosed herein, comprise administering an LNP disclosed herein in combination with an additional agent. In an embodiment, the additional agent is a standard of care for the disease or disorder, e.g., autoimmune disease. In an embodiment, the additional agent is an mRNA.

In some aspects, the subject for the present methods or compositions has been treated with one or more standard of care therapies. In other aspects, the subject for the present methods or compositions has not been responsive to one or more standard of care therapies or anti-cancer therapies.

Sequence optimization and methods thereof

In some embodiments, a polynucleotide of the disclosure comprises a sequence-optimized nucleotide sequence encoding a polypeptide disclosed herein, e.g., a polynucleotide encoding a therapeutic payload or prophylactic payload, an effector

molecule and/or a tether molecule. In some embodiments, the polynucleotide of the disclosure comprises an open reading frame (ORF) encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule, wherein the ORF has been sequence optimized.

5 The sequence-optimized nucleotide sequences disclosed herein are distinct from the corresponding wild type nucleotide acid sequences and from other known sequence-optimized nucleotide sequences, e.g., these sequence-optimized nucleic acids have unique compositional characteristics.

10 In some embodiments, the percentage of uracil or thymine nucleobases in a sequence-optimized nucleotide sequence (e.g., encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule, a functional fragment, or a variant thereof) is modified (e.g., reduced) with respect to the percentage of uracil or thymine nucleobases in the reference wild-type nucleotide sequence. Such a sequence is referred to as a uracil-modified or thymine-modified sequence. The
15 percentage of uracil or thymine content in a nucleotide sequence can be determined by dividing the number of uracils or thymines in a sequence by the total number of nucleotides and multiplying by 100. In some embodiments, the sequence-optimized nucleotide sequence has a lower uracil or thymine content than the uracil or thymine content in the reference wild-type sequence. In some embodiments, the uracil or
20 thymine content in a sequence-optimized nucleotide sequence of the disclosure is greater than the uracil or thymine content in the reference wild-type sequence and still maintain beneficial effects, e.g., increased expression and/or signaling response when compared to the reference wild-type sequence.

25 In some embodiments, the optimized sequences of the present disclosure contain unique ranges of uracils or thymine (if DNA) in the sequence. The uracil or thymine content of the optimized sequences can be expressed in various ways, e.g., uracil or thymine content of optimized sequences relative to the theoretical minimum (%UTM or %TTM), relative to the wild-type (%UWT or %TWT), and relative to the total nucleotide content (%UTL or %TTL). For DNA it is recognized that thymine (T) is
30 present instead of uracil (U), and one would substitute T where U appears. For RNA it is recognized that uracil (U) is present instead of thymine (T). One of skill in the art

could readily obtain an RNA sequence when the DNA sequence is provided by substituting thymine in the DNA sequence to uracil. Thus, all the disclosures related to, e.g., %UTM, %UWT, or %UTL, with respect to RNA are equally applicable to %TTM, %TWT, or %TTL with respect to DNA.

5 Uracil- or thymine- content relative to the uracil or thymine theoretical minimum, refers to a parameter determined by dividing the number of uracils or thymines in a sequence-optimized nucleotide sequence by the total number of uracils or thymines in a hypothetical nucleotide sequence in which all the codons in the hypothetical sequence are replaced with synonymous codons having the lowest possible
10 uracil or thymine content and multiplying by 100. This parameter is abbreviated herein as %UTM or %TTM.

 In some embodiments, a uracil-modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule of the disclosure has a reduced number of consecutive uracils with respect to the
15 corresponding wild-type nucleic acid sequence. For example, two consecutive leucines can be encoded by the sequence CUUUUG, which includes a four uracil cluster. Such a subsequence can be substituted, e.g., with CUGCUC, which removes the uracil cluster. Phenylalanine can be encoded by UUC or UUU. Thus, even if phenylalanines encoded by UUU are replaced by UUC, the synonymous codon still contains a uracil pair (UU).
20 Accordingly, the number of phenylalanines in a sequence establishes a minimum number of uracil pairs (UU) that cannot be eliminated without altering the number of phenylalanines in the encoded polypeptide.

 In some embodiments, a uracil-modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule of the
25 disclosure has a reduced number of uracil triplets (UUU) with respect to the wild-type nucleic acid sequence. In some embodiments, a uracil-modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule has a reduced number of uracil pairs (UU) with respect to the number of uracil pairs (UU) in the wild-type nucleic acid sequence. In some embodiments, a uracil-
30 modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule of the disclosure has a number of uracil pairs (UU)

corresponding to the minimum possible number of uracil pairs (UU) in the wild-type nucleic acid sequence.

The phrase "uracil pairs (UU) relative to the uracil pairs (UU) in the wild type nucleic acid sequence," refers to a parameter determined by dividing the number of uracil pairs (UU) in a sequence-optimized nucleotide sequence by the total number of uracil pairs (UU) in the corresponding wild-type nucleotide sequence and multiplying by 100. This parameter is abbreviated herein as %UUwt. In some embodiments, a uracil-modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule has a %UUwt between below 100%.

In some embodiments, the polynucleotide of the disclosure comprises a uracil-modified sequence encoding an encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule disclosed herein. In some embodiments, the uracil-modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule comprises at least one chemically modified nucleobase, e.g., 5-methoxyuracil. In some embodiments, at least 95% of a nucleobase (e.g., uracil) in a uracil-modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule of the disclosure are modified nucleobases. In some embodiments, at least 95% of uracil in a uracil-modified sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule is 5-methoxyuracil.

In some embodiments, a polynucleotide of the disclosure (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule (e.g., the wild-type sequence, functional fragment, or variant thereof) is sequence optimized.

A sequence optimized nucleotide sequence (nucleotide sequence is also referred to as "nucleic acid" herein) comprises at least one codon modification with respect to a reference sequence (e.g., a wild-type sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule). Thus, in a sequence optimized nucleic acid, at least one codon is different from a corresponding codon in a reference sequence (e.g., a wild-type sequence).

In general, sequence optimized nucleic acids are generated by at least a step comprising substituting codons in a reference sequence with synonymous codons (i.e., codons that encode the same amino acid). Such substitutions can be effected, for example, by applying a codon substitution map (i.e., a table providing the codons that will encode each amino acid in the codon optimized sequence), or by applying a set of rules (e.g., if glycine is next to neutral amino acid, glycine would be encoded by a certain codon, but if it is next to a polar amino acid, it would be encoded by another codon). In addition to codon substitutions (i.e., "codon optimization") the sequence optimization methods disclosed herein comprise additional optimization steps which are not strictly directed to codon optimization such as the removal of deleterious motifs (destabilizing motif substitution). Compositions and formulations comprising these sequence optimized nucleic acids (e.g., a RNA, e.g., an mRNA) can be administered to a subject in need thereof to facilitate in vivo expression of functionally active encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule.

Additional and exemplary methods of sequence optimization are disclosed in International PCT application WO 2017/201325, filed on 18 May 2017, the entire contents of which are hereby incorporated by reference.

Micro RNA (miRNA) binding sites

Nucleic acid molecules (e.g., RNA, e.g., mRNA) of the disclosure can include regulatory elements, for example, microRNA (miRNA) binding sites, transcription factor binding sites, structured mRNA sequences and/or motifs, artificial binding sites engineered to act as pseudo-receptors for endogenous nucleic acid binding molecules, and combinations thereof.

In some embodiments, a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure comprises an open reading frame (ORF) encoding a polypeptide of interest and further comprises one or more miRNA binding site(s). Inclusion or incorporation of miRNA binding site(s) provides for regulation of nucleic acid molecules (e.g., RNA, e.g., mRNA) of the disclosure, and in turn, of the polypeptides encoded therefrom,

based on tissue-specific and/or cell-type specific expression of naturally-occurring miRNAs.

A miRNA, e.g., a natural-occurring miRNA, is a 19-25 nucleotide long noncoding RNA that binds to a nucleic acid molecule (e.g., RNA, e.g., mRNA) and down-regulates gene expression either by reducing stability or by inhibiting translation of the polynucleotide. A miRNA sequence comprises a “seed” region, i.e., a sequence in the region of positions 2-8 of the mature miRNA. A miRNA seed can comprise positions 2-8 or 2-7 of the mature miRNA. In some embodiments, a miRNA seed can comprise 7 nucleotides (e.g., nucleotides 2-8 of the mature miRNA), wherein the seed-complementary site in the corresponding miRNA binding site is flanked by an adenosine (A) opposed to miRNA position 1. In some embodiments, a miRNA seed can comprise 6 nucleotides (e.g., nucleotides 2-7 of the mature miRNA), wherein the seed-complementary site in the corresponding miRNA binding site is flanked by an adenosine (A) opposed to miRNA position 1. See, for example, Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP; Mol Cell. 2007 Jul 6;27(1):91-105. miRNA profiling of the target cells or tissues can be conducted to determine the presence or absence of miRNA in the cells or tissues. In some embodiments, a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure comprises one or more microRNA binding sites, microRNA target sequences, microRNA complementary sequences, or microRNA seed complementary sequences. Such sequences can correspond to, e.g., have complementarity to, any known microRNA such as those taught in US Publication US2005/0261218 and US Publication US2005/0059005, the contents of each of which are incorporated herein by reference in their entirety.

As used herein, the term “microRNA (miRNA or miR) binding site” refers to a sequence within a nucleic acid molecule, e.g., within a DNA or within an RNA transcript, including in the 5'UTR and/or 3'UTR, that has sufficient complementarity to all or a region of a miRNA to interact with, associate with or bind to the miRNA. In some embodiments, a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure comprising an ORF encoding a polypeptide of interest and further comprises one or more miRNA binding site(s). In exemplary embodiments, a 5'UTR and/or 3'UTR of the

nucleic acid molecule (e.g., RNA, e.g., mRNA) comprises the one or more miRNA binding site(s).

A miRNA binding site having sufficient complementarity to a miRNA refers to a degree of complementarity sufficient to facilitate miRNA-mediated regulation of a nucleic acid molecule (e.g., RNA, e.g., mRNA), e.g., miRNA-mediated translational repression or degradation of the nucleic acid molecule (e.g., RNA, e.g., mRNA). In exemplary aspects of the disclosure, a miRNA binding site having sufficient complementarity to the miRNA refers to a degree of complementarity sufficient to facilitate miRNA-mediated degradation of the nucleic acid molecule (e.g., RNA, e.g., mRNA), e.g., miRNA-guided RNA-induced silencing complex (RISC)-mediated cleavage of mRNA. The miRNA binding site can have complementarity to, for example, a 19-25 nucleotide miRNA sequence, to a 19-23 nucleotide miRNA sequence, or to a 22 nucleotide miRNA sequence. A miRNA binding site can be complementary to only a portion of a miRNA, e.g., to a portion less than 1, 2, 3, or 4 nucleotides of the full length of a naturally-occurring miRNA sequence. Full or complete complementarity (e.g., full complementarity or complete complementarity over all or a significant portion of the length of a naturally-occurring miRNA) is preferred when the desired regulation is mRNA degradation.

In some embodiments, a miRNA binding site includes a sequence that has complementarity (e.g., partial or complete complementarity) with a miRNA seed sequence. In some embodiments, the miRNA binding site includes a sequence that has complete complementarity with a miRNA seed sequence. In some embodiments, a miRNA binding site includes a sequence that has complementarity (e.g., partial or complete complementarity) with an miRNA sequence. In some embodiments, the miRNA binding site includes a sequence that has complete complementarity with a miRNA sequence. In some embodiments, a miRNA binding site has complete complementarity with a miRNA sequence but for 1, 2, or 3 nucleotide substitutions, terminal additions, and/or truncations.

In some embodiments, the miRNA binding site is the same length as the corresponding miRNA. In other embodiments, the miRNA binding site is one, two, three, four, five, six, seven, eight, nine, ten, eleven or twelve nucleotide(s) shorter than the corresponding miRNA at the 5' terminus, the 3' terminus, or both. In still other
5 embodiments, the microRNA binding site is two nucleotides shorter than the corresponding microRNA at the 5' terminus, the 3' terminus, or both. The miRNA binding sites that are shorter than the corresponding miRNAs are still capable of degrading the mRNA incorporating one or more of the miRNA binding sites or preventing the mRNA from translation.

10 In some embodiments, the miRNA binding site binds the corresponding mature miRNA that is part of an active RISC containing Dicer. In another embodiment, binding of the miRNA binding site to the corresponding miRNA in RISC degrades the mRNA containing the miRNA binding site or prevents the mRNA from being translated. In some embodiments, the miRNA binding site has sufficient complementarity to miRNA
15 so that a RISC complex comprising the miRNA cleaves the nucleic acid molecule (e.g., RNA, e.g., mRNA) comprising the miRNA binding site. In other embodiments, the miRNA binding site has imperfect complementarity so that a RISC complex comprising the miRNA induces instability in the nucleic acid molecule (e.g., RNA, e.g., mRNA) comprising the miRNA binding site. In another embodiment, the miRNA binding site
20 has imperfect complementarity so that a RISC complex comprising the miRNA represses transcription of the nucleic acid molecule (e.g., RNA, e.g., mRNA) comprising the miRNA binding site.

25 In some embodiments, the miRNA binding site has one, two, three, four, five, six, seven, eight, nine, ten, eleven or twelve mismatch(es) from the corresponding miRNA.

In some embodiments, the miRNA binding site has at least about ten, at least about eleven, at least about twelve, at least about thirteen, at least about fourteen, at least about fifteen, at least about sixteen, at least about seventeen, at least about eighteen, at least about nineteen, at least about twenty, or at least about twenty-one contiguous

nucleotides complementary to at least about ten, at least about eleven, at least about twelve, at least about thirteen, at least about fourteen, at least about fifteen, at least about sixteen, at least about seventeen, at least about eighteen, at least about nineteen, at least about twenty, or at least about twenty-one, respectively, contiguous nucleotides of the corresponding miRNA.

By engineering one or more miRNA binding sites into a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure, the nucleic acid molecule (e.g., RNA, e.g., mRNA) can be targeted for degradation or reduced translation, provided the miRNA in question is available. This can reduce off-target effects upon delivery of the nucleic acid molecule (e.g., RNA, e.g., mRNA). For example, if a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure is not intended to be delivered to a tissue or cell but ends up in said tissue or cell, then a miRNA abundant in the tissue or cell can inhibit the expression of the gene of interest if one or multiple binding sites of the miRNA are engineered into the 5'UTR and/or 3'UTR of the nucleic acid molecule (e.g., RNA, e.g., mRNA).

For example, one of skill in the art would understand that one or more miRNA binding sites can be included in a nucleic acid molecule (e.g., an RNA, e.g., mRNA) to minimize expression in cell types other than lymphoid cells. In one embodiment, a miR122 binding site can be used. In another embodiment, a miR126 binding site can be used. In still another embodiment, multiple copies of these miRNA binding sites or combinations may be used.

Conversely, miRNA binding sites can be removed from nucleic acid molecule (e.g., RNA, e.g., mRNA) sequences in which they naturally occur in order to increase protein expression in specific tissues. For example, a binding site for a specific miRNA can be removed from a nucleic acid molecule (e.g., RNA, e.g., mRNA) to improve protein expression in tissues or cells containing the miRNA.

Regulation of expression in multiple tissues can be accomplished through introduction or removal of one or more miRNA binding sites, e.g., one or more distinct miRNA binding sites. The decision whether to remove or insert a miRNA binding site

can be made based on miRNA expression patterns and/or their profilings in tissues and/or cells in development and/or disease. Identification of miRNAs, miRNA binding sites, and their expression patterns and role in biology have been reported (e.g., Bonauer et al., Curr Drug Targets 2010 11:943-949; Anand and Cheresh Curr Opin Hematol 5 2011 18:171-176; Contreras and Rao Leukemia 2012 26:404-413 (2011 Dec 20. doi: 10.1038/leu.2011.356); Bartel Cell 2009 136:215-233; Landgraf et al, Cell, 2007 129:1401-1414; Gentner and Naldini, Tissue Antigens. 2012 80:393-403 and all references therein; each of which is incorporated herein by reference in its entirety).

miRNAs and miRNA binding sites can correspond to any known sequence, 10 including non-limiting examples described in U.S. Publication Nos. 2014/0200261, 2005/0261218, and 2005/0059005, each of which are incorporated herein by reference in their entirety.

Examples of tissues where miRNA are known to regulate mRNA, and thereby protein expression, include, but are not limited to, liver (miR-122), muscle (miR-133, 15 miR-206, miR-208), endothelial cells (miR-17-92, miR-126), myeloid cells (miR-142-3p, miR-142-5p, miR-16, miR-21, miR-223, miR-24, miR-27), adipose tissue (let-7, miR-30c), heart (miR-1d, miR-149), kidney (miR-192, miR-194, miR-204), and lung epithelial cells (let-7, miR-133, miR-126).

Specifically, miRNAs are known to be differentially expressed in immune cells 20 (also called hematopoietic cells), such as antigen presenting cells (APCs) (e.g., dendritic cells and monocytes), monocytes, monocytes, B lymphocytes, T lymphocytes, granulocytes, natural killer cells, etc. Immune cell specific miRNAs are involved in immunogenicity, autoimmunity, the immune response to infection, inflammation, as well as unwanted immune response after gene therapy and tissue/organ transplantation. 25 Immune cell specific miRNAs also regulate many aspects of development, proliferation, differentiation and apoptosis of hematopoietic cells (immune cells). For example, miR-142 and miR-146 are exclusively expressed in immune cells, particularly abundant in myeloid dendritic cells. It has been demonstrated that the immune response to a nucleic acid molecule (e.g., RNA, e.g., mRNA) can be shut-off by adding miR-142 binding

sites to the 3'-UTR of the polynucleotide, enabling more stable gene transfer in tissues and cells. miR-142 efficiently degrades exogenous nucleic acid molecules (e.g., RNA, e.g., mRNA) in antigen presenting cells and suppresses cytotoxic elimination of transduced cells (e.g., Annoni A et al., blood, 2009, 114, 5152-5161; Brown BD, et al.,
5 Nat med. 2006, 12(5), 585-591; Brown BD, et al., blood, 2007, 110(13): 4144-4152, each of which is incorporated herein by reference in its entirety).

An antigen-mediated immune response can refer to an immune response triggered by foreign antigens, which, when entering an organism, are processed by the antigen presenting cells and displayed on the surface of the antigen presenting cells. T
10 cells can recognize the presented antigen and induce a cytotoxic elimination of cells that express the antigen.

Introducing a miR-142 binding site into the 5'UTR and/or 3'UTR of a nucleic acid molecule of the disclosure can selectively repress gene expression in antigen presenting cells through miR-142 mediated degradation, limiting antigen presentation in
15 antigen presenting cells (e.g., dendritic cells) and thereby preventing antigen-mediated immune response after the delivery of the nucleic acid molecule (e.g., RNA, e.g., mRNA). The nucleic acid molecule (e.g., RNA, e.g., mRNA) is then stably expressed in target tissues or cells without triggering cytotoxic elimination.

In one embodiment, binding sites for miRNAs that are known to be expressed in
20 immune cells, in particular, antigen presenting cells, can be engineered into a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure to suppress the expression of the nucleic acid molecule (e.g., RNA, e.g., mRNA) in antigen presenting cells through miRNA mediated RNA degradation, subduing the antigen-mediated immune response. Expression of the nucleic acid molecule (e.g., RNA, e.g., mRNA) is maintained in non-
25 immune cells where the immune cell specific miRNAs are not expressed. For example, in some embodiments, to prevent an immunogenic reaction against a liver specific protein, any miR-122 binding site can be removed and a miR-142 (and/or mirR-146) binding site can be engineered into the 5'UTR and/or 3'UTR of a nucleic acid molecule of the disclosure.

To further drive the selective degradation and suppression in APCs and macrophage, a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure can include a further negative regulatory element in the 5'UTR and/or 3'UTR, either alone or in combination with miR-142 and/or miR-146 binding sites. As a non-limiting
 5 example, the further negative regulatory element is a Constitutive Decay Element (CDE).

Immune cell specific miRNAs include, but are not limited to, hsa-let-7a-2-3p, hsa-let-7a-3p, hsa-7a-5p, hsa-let-7c, hsa-let-7e-3p, hsa-let-7e-5p, hsa-let-7g-3p, hsa-let-7g-5p, hsa-let-7i-3p, hsa-let-7i-5p, miR-10a-3p, miR-10a-5p, miR-1184, hsa-let-7f-1--
 10 3p, hsa-let-7f-2--5p, hsa-let-7f-5p, miR-125b-1-3p, miR-125b-2-3p, miR-125b-5p, miR-1279, miR-130a-3p, miR-130a-5p, miR-132-3p, miR-132-5p, miR-142-3p, miR-142-5p, miR-143-3p, miR-143-5p, miR-146a-3p, miR-146a-5p, miR-146b-3p, miR-146b-5p, miR-147a, miR-147b, miR-148a-5p, miR-148a-3p, miR-150-3p, miR-150-5p, miR-151b, miR-155-3p, miR-155-5p, miR-15a-3p, miR-15a-5p, miR-15b-5p, miR-15b-3p,
 15 miR-16-1-3p, miR-16-2-3p, miR-16-5p, miR-17-5p, miR-181a-3p, miR-181a-5p, miR-181a-2-3p, miR-182-3p, miR-182-5p, miR-197-3p, miR-197-5p, miR-21-5p, miR-21-3p, miR-214-3p, miR-214-5p, miR-223-3p, miR-223-5p, miR-221-3p, miR-221-5p, miR-23b-3p, miR-23b-5p, miR-24-1-5p, miR-24-2-5p, miR-24-3p, miR-26a-1-3p, miR-26a-2-3p, miR-26a-5p, miR-26b-3p, miR-26b-5p, miR-27a-3p, miR-27a-5p, miR-27b-3p, miR-27b-5p, miR-28-3p, miR-28-5p, miR-2909, miR-29a-3p, miR-29a-5p, miR-29b-1-5p, miR-29b-2-5p, miR-29c-3p, miR-29c-5p, miR-30e-3p, miR-30e-5p, miR-331-5p, miR-339-3p, miR-339-5p, miR-345-3p, miR-345-5p, miR-346, miR-34a-3p, miR-34a-5p, , miR-363-3p, miR-363-5p, miR-372, miR-377-3p, miR-377-5p, miR-493-3p, miR-493-5p, miR-542, miR-548b-5p, miR-548c-5p, miR-548i, miR-548j, miR-548n,
 20 miR-574-3p, miR-598, miR-718, miR-935, miR-99a-3p, miR-99a-5p, miR-99b-3p, and miR-99b-5p. Furthermore, novel miRNAs can be identified in immune cell through micro-array hybridization and microtome analysis (e.g., Jima DD et al, Blood, 2010, 116:e118-e127; Vaz C et al., BMC Genomics, 2010, 11,288, the content of each of which is incorporated herein by reference in its entirety.)

In some embodiments, a miRNA binding site is inserted in the nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure in any position of the nucleic acid molecule (e.g., RNA, e.g., mRNA) (e.g., the 5'UTR and/or 3'UTR). In some embodiments, the 5'UTR comprises a miRNA binding site. In some embodiments, the 3'UTR comprises a miRNA binding site. In some embodiments, the 5'UTR and the 3'UTR comprise a miRNA binding site. The insertion site in the nucleic acid molecule (e.g., RNA, e.g., mRNA) can be anywhere in the nucleic acid molecule (e.g., RNA, e.g., mRNA) as long as the insertion of the miRNA binding site in the nucleic acid molecule (e.g., RNA, e.g., mRNA) does not interfere with the translation of a functional polypeptide in the absence of the corresponding miRNA; and in the presence of the miRNA, the insertion of the miRNA binding site in the nucleic acid molecule (e.g., RNA, e.g., mRNA) and the binding of the miRNA binding site to the corresponding miRNA are capable of degrading the polynucleotide or preventing the translation of the nucleic acid molecule (e.g., RNA, e.g., mRNA).

In some embodiments, a miRNA binding site is inserted in at least about 30 nucleotides downstream from the stop codon of an ORF in a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure comprising the ORF. In some embodiments, a miRNA binding site is inserted in at least about 10 nucleotides, at least about 15 nucleotides, at least about 20 nucleotides, at least about 25 nucleotides, at least about 30 nucleotides, at least about 35 nucleotides, at least about 40 nucleotides, at least about 45 nucleotides, at least about 50 nucleotides, at least about 55 nucleotides, at least about 60 nucleotides, at least about 65 nucleotides, at least about 70 nucleotides, at least about 75 nucleotides, at least about 80 nucleotides, at least about 85 nucleotides, at least about 90 nucleotides, at least about 95 nucleotides, or at least about 100 nucleotides downstream from the stop codon of an ORF in a polynucleotide of the disclosure. In some embodiments, a miRNA binding site is inserted in about 10 nucleotides to about 100 nucleotides, about 20 nucleotides to about 90 nucleotides, about 30 nucleotides to about 80 nucleotides, about 40 nucleotides to about 70 nucleotides, about 50 nucleotides to about 60 nucleotides, about 45 nucleotides to about 65 nucleotides downstream from

the stop codon of an ORF in a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure.

miRNA gene regulation can be influenced by the sequence surrounding the miRNA such as, but not limited to, the species of the surrounding sequence, the type of
5 sequence (e.g., heterologous, homologous, exogenous, endogenous, or artificial), regulatory elements in the surrounding sequence and/or structural elements in the surrounding sequence. The miRNA can be influenced by the 5'UTR and/or 3'UTR. As a non-limiting example, a non-human 3'UTR can increase the regulatory effect of the miRNA sequence on the expression of a polypeptide of interest compared to a human
10 3'UTR of the same sequence type.

In one embodiment, other regulatory elements and/or structural elements of the 5'UTR can influence miRNA mediated gene regulation. One example of a regulatory element and/or structural element is a structured IRES (Internal Ribosome Entry Site) in the 5'UTR, which is necessary for the binding of translational elongation factors to
15 initiate protein translation. EIF4A2 binding to this secondarily structured element in the 5'-UTR is necessary for miRNA mediated gene expression (Meijer HA et al., Science, 2013, 340, 82-85, herein incorporated by reference in its entirety). The nucleic acid molecules (e.g., RNA, e.g., mRNA) of the disclosure can further include this structured 5'UTR in order to enhance microRNA mediated gene regulation.

20 At least one miRNA binding site can be engineered into the 3'UTR of a polynucleotide of the disclosure. In this context, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten, or more miRNA binding sites can be engineered into a 3'UTR of a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure. For example, 1 to 10, 1 to 9, 1 to 8, 1 to 7, 1 to 6,
25 1 to 5, 1 to 4, 1 to 3, 2, or 1 miRNA binding sites can be engineered into the 3'UTR of a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure. In one embodiment, miRNA binding sites incorporated into a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure can be the same or can be different miRNA sites. A combination of different miRNA binding sites incorporated into a nucleic acid molecule

(e.g., RNA, e.g., mRNA) of the disclosure can include combinations in which more than one copy of any of the different miRNA sites are incorporated. In another embodiment, miRNA binding sites incorporated into a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure can target the same or different tissues in the body. As a non-limiting example, through the introduction of tissue-, cell-type-, or disease-specific miRNA binding sites in the 3'-UTR of a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure, the degree of expression in specific cell types (e.g., hepatocytes, myeloid cells, endothelial cells, cancer cells, etc.) can be reduced.

In one embodiment, a miRNA binding site can be engineered near the 5' terminus of the 3'UTR, about halfway between the 5' terminus and 3' terminus of the 3'UTR and/or near the 3' terminus of the 3'UTR in a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure. As a non-limiting example, a miRNA binding site can be engineered near the 5' terminus of the 3'UTR and about halfway between the 5' terminus and 3' terminus of the 3'UTR. As another non-limiting example, a miRNA binding site can be engineered near the 3' terminus of the 3'UTR and about halfway between the 5' terminus and 3' terminus of the 3'UTR. As yet another non-limiting example, a miRNA binding site can be engineered near the 5' terminus of the 3'UTR and near the 3' terminus of the 3'UTR.

In another embodiment, a 3'UTR can comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 miRNA binding sites. The miRNA binding sites can be complementary to a miRNA, miRNA seed sequence, and/or miRNA sequences flanking the seed sequence.

A nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure can be engineered for more targeted expression in specific tissues, cell types, or biological conditions based on the expression patterns of miRNAs in the different tissues, cell types, or biological conditions. Through introduction of tissue-specific miRNA binding sites, a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure can be designed for optimal protein expression in a tissue or cell, or in the context of a biological condition.

In some embodiments, a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure can comprise at least one miRNA binding site in the 3'UTR in order to selectively degrade mRNA therapeutics in the immune cells to subdue unwanted immunogenic reactions caused by therapeutic delivery. As a non-limiting example, the miRNA binding site can make a nucleic acid molecule (e.g., RNA, e.g., mRNA) of the disclosure more unstable in antigen presenting cells. Non-limiting examples of these miRNAs include mir-142-5p, mir-142-3p, mir-146a-5p, and mir-146-3p

IVT polynucleotide architecture

In some embodiments, the polynucleotide of the present disclosure comprising an mRNA encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule is an IVT polynucleotide. Traditionally, the basic components of an mRNA molecule include at least a coding region, a 5'UTR, a 3'UTR, a 5' cap and a poly-A tail. The IVT polynucleotides of the present disclosure can function as mRNA but are distinguished from wild-type mRNA in their functional and/or structural design features which serve, e.g., to overcome existing problems of effective polypeptide production using nucleic-acid based therapeutics.

The primary construct of an IVT polynucleotide comprises a first region of linked nucleotides that is flanked by a first flanking region and a second flanking region. This first region can include, but is not limited to, the encoded therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule. The first flanking region can include a sequence of linked nucleosides which function as a 5' untranslated region (UTR) such as the 5' UTR of any of the nucleic acids encoding the native 5' UTR of the polypeptide or a non-native 5'UTR such as, but not limited to, a heterologous 5' UTR or a synthetic 5' UTR. The IVT encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule can comprise at its 5 terminus a signal sequence region encoding one or more signal sequences. The flanking region can comprise a region of linked nucleotides comprising one or more complete or incomplete 5' UTRs sequences. The flanking region can also comprise a 5' terminal cap. The second flanking region can comprise a region of linked nucleotides

comprising one or more complete or incomplete 3' UTRs which can encode the native 3' UTR of a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule or a non-native 3' UTR such as, but not limited to, a heterologous 3' UTR or a synthetic 3' UTR. The flanking region can also comprise a 3' tailing
5 sequence. The 3' tailing sequence can be, but is not limited to, a polyA tail, a polyA-G quartet and/or a stem loop sequence.

Additional and exemplary features of IVT polynucleotide architecture are disclosed in International PCT application WO 2017/201325, filed on 18 May 2017, the entire contents of which are hereby incorporated by reference.

10

5'UTR and 3' UTR

A UTR can be homologous or heterologous to the coding region in a polynucleotide. In some embodiments, the UTR is homologous to the ORF encoding the therapeutic payload or prophylactic payload, an effector molecule and/or a tether
15 molecule. In some embodiments, the UTR is heterologous to the ORF encoding the therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule.

In some embodiments, the polynucleotide comprises two or more 5' UTRs or functional fragments thereof, each of which has the same or different nucleotide
20 sequences. In some embodiments, the polynucleotide comprises two or more 3' UTRs or functional fragments thereof, each of which has the same or different nucleotide sequences.

In some embodiments, the 5' UTR or functional fragment thereof, 3' UTR or functional fragment thereof, or any combination thereof is sequence optimized.

25 In some embodiments, the 5'UTR or functional fragment thereof, 3' UTR or functional fragment thereof, or any combination thereof comprises at least one chemically modified nucleobase, e.g., N1-methylpseudouracil or 5-methoxyuracil.

UTRs can have features that provide a regulatory role, e.g., increased or decreased stability, localization and/or translation efficiency. A polynucleotide
30 comprising a UTR can be administered to a cell, tissue, or organism, and one or more regulatory features can be measured using routine methods. In some embodiments, a

functional fragment of a 5' UTR or 3' UTR comprises one or more regulatory features of a full length 5' or 3' UTR, respectively.

Natural 5'UTRs bear features that play roles in translation initiation. They harbor signatures like Kozak sequences that are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus CCR(A/G)CCAUGG (SEQ ID NO:157), where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G'. 5' UTRs also have been known to form secondary structures that are involved in elongation factor binding.

By engineering the features typically found in abundantly expressed genes of specific target organs, one can enhance the stability and protein production of a polynucleotide. For example, introduction of 5' UTR of liver-expressed mRNA, such as albumin, serum amyloid A, Apolipoprotein A/B/E, transferrin, alpha fetoprotein, erythropoietin, or Factor VIII, can enhance expression of polynucleotides in hepatic cell lines or liver. Likewise, use of 5'UTR from other tissue-specific mRNA to improve expression in that tissue is possible for muscle (e.g., MyoD, Myosin, Myoglobin, Myogenin, Herculin), for endothelial cells (e.g., Tie-1, CD36), for myeloid cells (e.g., C/EBP, AML1, G-CSF, GM-CSF, CD11b, MSR, Fr-1, i-NOS), for leukocytes (e.g., CD45, CD18), for adipose tissue (e.g., CD36, GLUT4, ACRP30, adiponectin) and for lung epithelial cells (e.g., SP-A/B/C/D).

In some embodiments, UTRs are selected from a family of transcripts whose proteins share a common function, structure, feature or property. For example, an encoded polypeptide can belong to a family of proteins (i.e., that share at least one function, structure, feature, localization, origin, or expression pattern), which are expressed in a particular cell, tissue or at some time during development. The UTRs from any of the genes or mRNA can be swapped for any other UTR of the same or different family of proteins to create a new polynucleotide.

In some embodiments, the 5' UTR and the 3' UTR can be heterologous. In some embodiments, the 5' UTR can be derived from a different species than the 3' UTR. In some embodiments, the 3' UTR can be derived from a different species than the 5' UTR.

Co-owned International Patent Application No. PCT/US2014/021522 (Publ. No. WO/2014/164253, incorporated herein by reference in its entirety) provides a listing of exemplary UTRs that can be utilized in the polynucleotide of the present invention as flanking regions to an ORF.

5 Exemplary UTRs of the application include, but are not limited to, one or more 5'UTR and/or 3'UTR derived from the nucleic acid sequence of: a globin, such as an α - or β -globin (e.g., a *Xenopus*, mouse, rabbit, or human globin); a strong Kozak translational initiation signal; a CYBA (e.g., human cytochrome b-245 α polypeptide); an albumin (e.g., human albumin7); a HSD17B4 (hydroxysteroid (17- β)
10 dehydrogenase); a virus (e.g., a tobacco etch virus (TEV), a Venezuelan equine encephalitis virus (VEEV), a Dengue virus, a cytomegalovirus (CMV) (e.g., CMV immediate early 1 (IE1)), a hepatitis virus (e.g., hepatitis B virus), a sindbis virus, or a PAV barley yellow dwarf virus); a heat shock protein (e.g., hsp70); a translation initiation factor (e.g., eIF4G); a glucose transporter (e.g., hGLUT1 (human glucose
15 transporter 1)); an actin (e.g., human α or β actin); a GAPDH; a tubulin; a histone; a citric acid cycle enzyme; a topoisomerase (e.g., a 5'UTR of a TOP gene lacking the 5' TOP motif (the oligopyrimidine tract)); a ribosomal protein Large 32 (L32); a ribosomal protein (e.g., human or mouse ribosomal protein, such as, for example, rps9); an ATP synthase (e.g., ATP5A1 or the β subunit of mitochondrial H^+ -ATP synthase); a growth
20 hormone e (e.g., bovine (bGH) or human (hGH)); an elongation factor (e.g., elongation factor 1 α 1 (EEF1A1)); a manganese superoxide dismutase (MnSOD); a myocyte enhancer factor 2A (MEF2A); a β -F1-ATPase, a creatine kinase, a myoglobin, a granulocyte-colony stimulating factor (G-CSF); a collagen (e.g., collagen type I, alpha 2 (Col1A2), collagen type I, alpha 1 (Col1A1), collagen type VI, alpha 2 (Col6A2),
25 collagen type VI, alpha 1 (Col6A1)); a ribophorin (e.g., ribophorin I (RPNI)); a low density lipoprotein receptor-related protein (e.g., LRP1); a cardiotrophin-like cytokine factor (e.g., Nnt1); calreticulin (Calr); a procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1 (Plod1); and a nucleobindin (e.g., Nucb1).

In some embodiments, the 5' UTR is selected from the group consisting of a β -globin 5'
30 UTR; a 5'UTR containing a strong Kozak translational initiation signal; a cytochrome b-245 α polypeptide (CYBA) 5' UTR; a hydroxysteroid (17- β) dehydrogenase

(HSD17B4) 5' UTR; a Tobacco etch virus (TEV) 5' UTR; a Venezuelan equine encephalitis virus (TEEV) 5' UTR; a 5' proximal open reading frame of rubella virus (RV) RNA encoding nonstructural proteins; a Dengue virus (DEN) 5' UTR; a heat shock protein 70 (Hsp70) 5' UTR; a eIF4G 5' UTR; a GLUT1 5' UTR; functional
 5 fragments thereof and any combination thereof.

In some embodiments, the 3' UTR is selected from the group consisting of a β -globin 3' UTR; a CYBA 3' UTR; an albumin 3' UTR; a growth hormone (GH) 3' UTR; a VEEV 3' UTR; a hepatitis B virus (HBV) 3' UTR; α -globin 3' UTR; a DEN 3' UTR; a PAV barley yellow dwarf virus (BYDV-PAV) 3' UTR; an elongation factor 1 α 1 (EEF1A1) 3' UTR; a manganese superoxide dismutase (MnSOD) 3' UTR; a β subunit of
 10 mitochondrial H(+)-ATP synthase (β -mRNA) 3' UTR; a GLUT1 3' UTR; a MEF2A 3' UTR; a β -F1-ATPase 3' UTR; functional fragments thereof and combinations thereof.

Wild-type UTRs derived from any gene or mRNA can be incorporated into the polynucleotides of the invention. In some embodiments, a UTR can be altered relative
 15 to a wild type or native UTR to produce a variant UTR, e.g., by changing the orientation or location of the UTR relative to the ORF; or by inclusion of additional nucleotides, deletion of nucleotides, swapping or transposition of nucleotides. In some embodiments, variants of 5' or 3' UTRs can be utilized, for example, mutants of wild type UTRs, or variants wherein one or more nucleotides are added to or removed from a
 20 terminus of the UTR.

Additionally, one or more synthetic UTRs can be used in combination with one or more non-synthetic UTRs. *See*, e.g., Mandal and Rossi, Nat. Protoc. 2013 8(3):568-82, the contents of which are incorporated herein by reference in their entirety.

UTRs or portions thereof can be placed in the same orientation as in the
 25 transcript from which they were selected or can be altered in orientation or location. Hence, a 5' and/or 3' UTR can be inverted, shortened, lengthened, or combined with one or more other 5' UTRs or 3' UTRs.

In some embodiments, the polynucleotide comprises multiple UTRs, e.g., a double, a triple or a quadruple 5' UTR or 3' UTR. For example, a double UTR comprises two
 30 copies of the same UTR either in series or substantially in series. For example, a double

beta-globin 3'UTR can be used (see US2010/0129877, the contents of which are incorporated herein by reference in its entirety).

In certain embodiments, the polynucleotides of the invention comprise a 5' UTR and/or a 3' UTR selected from any of the UTRs disclosed herein, e.g., in Table 5.

5

Table 5: Exemplary 5' UTR and 3' UTR sequences

SEQ ID NO:	Sequence
<u>3' UTR sequences</u>	
99	GCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGC CCCUCCUCCCCUCCUGCACCCGUACCCCUCCAUAAGUAGGAAACAC <u>UACAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC</u> (3' UTR with miR 142-3p binding site)
100	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCU <u>CGCAUUAUU</u> <u>ACUCACGGUACGAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC</u> (3' UTR with miR 126-3p binding site)
101	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCUUGGUCUUU GAAUAAAGUCUGAGUGGGCGGC (3' UTR, no miR binding sites)
102	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCU <u>CCAUAAG</u> <u>UAGGAAACACUACAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC</u> (3' UTR with miR 142-3p binding site)
103	UGAUAAUAGUCCAUAAGUAGGAAACACUACAGCUGGAGCCUCGGUGG CCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCU <u>CGCAUUAUUACUCACGGUACGAGUGGUCUUUGAA</u> UAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p and miR 126-3p binding sites variant 1)
104	UGAUAAUAGUCCAUAAGUAGGAAACACUACAGCUGGAGCCUCGGUGG CCAUGCUUCUUGCCCCUUGGGCC <u>CCAUAAGUAGGAAACACUACAUC</u> CCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCU <u>CCAUAAGUA</u> <u>GGAAACACUACAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC</u> (3' UTR with 3 miR 142-3p binding sites)

105	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCC<u>AGUAGUGCUU</u> <u>UCUACUUUAUG</u>GUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-5p binding site)
106	UGAUAAUAG<u>AGUAGUGCUUUCUACUUUAUG</u>GCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCC<u>AGUAGUGCUUUCUACUUUAUG</u>CCCCC AGCCCCUCCUCCCCUCCUGCACCCGUACCCCC<u>AGUAGUGCUUUCUACU</u> <u>UUAUG</u>GUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with 3 miR 142-5p binding sites)
107	UGAUAAUAG<u>AGUAGUGCUUUCUACUUUAUG</u>GCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCCUCCAUAAGUAGGAAACACUACA<u>UCCC</u> CCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCC<u>AGUAGUGCUUUCUA</u> <u>CUUUAUG</u>GUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with 2 miR 142-5p binding sites and 1 miR 142-3p binding site)
108	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCC<u>ACCCCUAUCA</u> <u>CAAUAGCAUUAAG</u>GUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 155-5p binding site)
109	UGAUAAUAG<u>ACCCCUAUCACAAUUAAGCAUUAAG</u>CUGGAGCCUCGGUGGC CAUGCUUCUUGCCCCUUGGGCC<u>ACCCCUAUCACAAUUAAGCAUUA</u>AUCCCC CCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCC<u>ACCCCUAUCACAAU</u> <u>AGCAUUAAG</u>GUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with 3 miR 155-5p binding sites)
110	UGAUAAUAG<u>ACCCCUAUCACAAUUAAGCAUUAAG</u>CUGGAGCCUCGGUGGC CAUGCUUCUUGCCCCUUGGGCCUCCAUAAGUAGGAAACACUACA<u>UCC</u> CCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCC<u>ACCCCUAUCACAA</u> <u>UUAAGCAUUAAG</u>GUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with 2 miR 155-5p binding sites and 1 miR 142-3p binding site)
111	UGAUAAUAG<u>UCCAUAAGUAGGAAACACUACAG</u>CUGGAGCCUCGGUGG CCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p binding site, P1 insertion)
112	UGAUAAUAGGCUGGAGCCUCGGUGGCCUCCAUAAGUAGGAAACACUAC <u>ACAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCU</u> GCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p binding site, P2 insertion)

113	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCAUAAGUAGGAAACACUACA<u>UCCCCCAGCCCCUCCUCCCCUCCU</u> GCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p binding site, P3 insertion)
114	UGAUAAUAG<u>AGUAGUGCUUUCUACUUA</u>AUGGCUGGAGCCUCGGUGGCC AUGCUUCUUGCCCCUUGGGCC<u>AGUAGUGCUUUCUACUUA</u>AUGUCCCCC AGCCCCUCUCCCCUCCUGCACCCGUACCCCGAGUAGUGCUUUCUACUU <u>UAUGGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC</u> (3' UTR with 3 miR 142-5p binding sites)
115	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUCCAUA AAGUAGGAAACACUACA<u>UGGGCCUCCCCCAGCCCCUCCUCCCCUCCU</u> GCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3'UTR including miR142-3p binding site)
116	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGUCCAUAAGUAGGAAACACUAC<u>ACCCCUCCUCCCCUCCU</u> GCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3'UTR including miR142-3p binding site)
117	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUUCUCCAUAAGUAGGAAACACUAC<u>ACU</u> GCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3'UTR including miR142-3p binding site)
118	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUU GAAUAAAGU<u>UCCAUAAGUAGGAAACACUAC</u>ACUGAGUGGGCGGC (3'UTR including miR142-3p binding site)
119	UGAUAAUAG<u>UCCAUAAGUAGGAAACACUAC</u><u>AGCUGGAGCCUCGGUGG</u> CCUAGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCGCAUUAUUACUCACGGUACGAGUGGUCUUUGAA UAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p and miR 126-3p binding sites variant 2)
120	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGC (3' UTR, no miR binding sites variant 2)

121	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCUCCAUAAG UAGGAAACACUACAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p binding site variant 3)
122	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCGCAUUAUU ACUCACGGUACGAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 126-3p binding site variant 3)
123	UGAUAAUAGUCCAUAAGUAGGAAACACUACAGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCCAUAAGUAGGAAACACUACAUC CCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCUCCAUAAGUA GGAAACACUACAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with 3 miR 142-3p binding sites variant 2)
124	UGAUAAUAGUCCAUAAGUAGGAAACACUACAGCUGGAGCCUCGGUGG CCUAGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p binding site, P1 insertion variant 2)
125	UGAUAAUAGGCUGGAGCCUCGGUGGCCUCCAUAAGUAGGAAACACUAC ACUAGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p binding site, P2 insertion variant 2)
126	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCC UCCAUAAGUAGGAAACACUACAUCCCCCAGCCCCUCCUCCCCUCCU GCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 142-3p binding site, P3 insertion variant 2)
127	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCACCCCUAUCA CAUUUAGCAUUAAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with miR 155-5p binding site variant 2)
128	UGAUAAUAGACCCCUAUCACAAUUAGCAUUAAGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCACCCCUAUCACAAUUAGCAUUAUCCCC CCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCACCCCUAUCACAAU AGCAUUAAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3' UTR with 3 miR 155-5p binding sites variant 2)
129	UGAUAAUAGACCCCUAUCACAAUUAGCAUUAAGCUGGAGCCUCGGUGGC CUAGCUUCUUGCCCCUUGGGCCUCCAUAAGUAGGAAACACUACAUC

	CCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCACCCCUAUCACAA UUAGCAUUAAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3'UTR with 2 miR 155-5p binding sites and 1 miR 142-3p binding site variant 2)
130	UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCAAACACCAU UGUCACACUCCAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3'UTR (MIR122))
131	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCAAACACCAU UGUCACACUCCAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC 3' UTR (MIR122-TAG)
132	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCC UCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCCGCAUUAU ACUCACGGUACGAGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGC (3'UTR V1.1 (MIR126-3P))
5' UTR SEQUENCES	
133	GGGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAGCCACC (5' UTR)
134	GGGAAAUAAAGAGUCCAUAAGUAGGAAACACUACAAGAAAAGAAGAGU AAGAAGAAAUUAAGAGCCACC (5' UTR with miR142-3p binding site at position p1)
135	GGGAAAUAAAGAGAGAAAAGAAGAGUAAUCCAUAAGUAGGAAACACU ACAGAAGAAAUUAAGAGCCACC (5' UTR with miR142-3p binding site at position p2)
136	GGGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAUCCAUA GUAGGAAACACUACAGAGCCACC (5' UTR with miR142-3p binding site at position p3)
137	AGGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAGCCACC 5'-UTR (v1 A-start)
138	GGGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGACCCCGGC GCCGCCACC (5' UTR v1.1)

139	AGGAAAUAAAGAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGACCCCGGC GCCGCCACC (5'UTR v1.1 A-start)
140	GGGAGAUCAAGAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAGCCACC 5' UTR 002 (upstream UTR)
141	GGGAGACAAGCUUGGCAUUCCGGUACUGUUGGUAAAGCCACC 5' UTR-004 (Upstream UTR)
142	GGGAAUUAACAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAGCCACC 5' UTR-008 (Upstream UTR)
143	GGGAAAUUAGACAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAGCCACC 5' UTR-009 (Upstream UTR)
144	GGGAAAUAAAGAGAGUAAAGAACAGUAAGAAGAAAUUAAGAGCCACC 5' UTR-010, Upstream
145	GGGAAAAAAGAGAGAGAAAAGAAGACUAAGAAGAAAUUAAGAGCCACC 5' UTR-011 (Upstream UTR)
146	GGGAAAUAAAGAGAGAGAAAAGAAGAGUAAGAAGAUUAUAAGAGCCACC 5' UTR-012 (Upstream UTR)
147	GGGAAAUAAAGAGACAAAACAAGAGUAAGAAGAAAUUAAGAGCCACC 5' UTR-013 (Upstream UTR)
148	GGGAAAUUAGAGAGUAAAGAACAGUAAGUAGAAUUAAAAGAGCCACC 5' UTR-014 (Upstream UTR)
149	GGGAAAUAAAGAGAGAAUAGAAGAGUAAGAAGAAAUUAAGAGCCACC 5' UTR-015 (Upstream UTR)
150	GGGAAAUAAAGAGAGAGAAAAGAAGAGUAAGAAGAAAAUUAAGAGCCACC 5' UTR-016 (Upstream UTR)
151	GGGAAAUAAAGAGAGAGAAAAGAAGAGUAAGAAGAAAUUUAAGAGCCACC 5' UTR-017 (Upstream UTR)
152	UCAAGCUUUUGGACCCUCGUACAGAAGCUAAUACGACUCACUAUAGGG AAAUAAAGAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAGCCACC

	5' UTR-018 (Upstream UTR)

Stop codon = bold

miR 142-3p binding site = underline

miR 126-3p binding site = bold underline

5 miR 155-5p binding site = italicized

miR 142-5p binding site = italicized and bold underline

In certain embodiments, the 5' UTR and/or 3' UTR sequence of the invention comprises a nucleotide sequence at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identical to a sequence provided in Table 5.

In some embodiments, the 5' UTR comprises a nucleotide sequence at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identical to a sequence provided in Table 5.

In some embodiments, the 3' UTR comprises a nucleotide sequence at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identical to a sequence provided in Table 5.

In some embodiments, the polynucleotide disclosed herein, e.g., the polynucleotide encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule, comprises a 5' UTR having the sequence of a 5' UTR provided in Table 5, or a sequence with at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity thereto. In some embodiments, the polynucleotide comprises a 5' UTR comprising the sequence of any one of SEQ ID NOs: 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, or 152, or a sequence with at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least

about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity thereto.

In some embodiments, the polynucleotide disclosed herein, e.g., the
5 polynucleotide encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule, comprises a 3' UTR having the sequence of a 3' UTR provided in Table 5, or a sequence with at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity thereto. In
10 some embodiments, the polynucleotide comprises a 3' UTR comprising the sequence of any one of SEQ ID NOs: 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, or 132, or a sequence with at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least
15 about 97%, at least about 98%, at least about 99%, or about 100% identity thereto.

The polynucleotides of the invention can comprise combinations of features. For example, the ORF can be flanked by a 5'UTR that comprises a strong Kozak translational initiation signal and/or a 3'UTR comprising an oligo(dT) sequence for templated addition of a poly-A tail. A 5'UTR can comprise a first polynucleotide
20 fragment and a second polynucleotide fragment from the same and/or different UTRs (see, e.g., US2010/0293625, herein incorporated by reference in its entirety).

Other non-UTR sequences can be used as regions or subregions within the polynucleotides of the invention. For example, introns or portions of intron sequences can be incorporated into the polynucleotides of the invention. Incorporation of intronic
25 sequences can increase protein production as well as polynucleotide expression levels. In some embodiments, the polynucleotide of the invention comprises an internal ribosome entry site (IRES) instead of or in addition to a UTR (see, e.g., Yakubov et al., Biochem. Biophys. Res. Commun. 2010 394(1):189-193, the contents of which are incorporated herein by reference in their entirety). In some embodiments, the
30 polynucleotide comprises an IRES instead of a 5' UTR sequence. In some embodiments, the polynucleotide comprises an ORF and a viral capsid sequence. In

some embodiments, the polynucleotide comprises a synthetic 5' UTR in combination with a non-synthetic 3' UTR.

In some embodiments, the UTR can also include at least one translation enhancer polynucleotide, translation enhancer element, or translational enhancer elements (collectively, "TEE," which refers to nucleic acid sequences that increase the amount of polypeptide or protein produced from a polynucleotide. As a non-limiting example, the TEE can be located between the transcription promoter and the start codon. In some embodiments, the 5' UTR comprises a TEE.

In one aspect, a TEE is a conserved element in a UTR that can promote translational activity of a nucleic acid such as, but not limited to, cap-dependent or cap-independent translation.

Regions having a 5' cap

The disclosure also includes a polynucleotide that comprises both a 5' Cap and a polynucleotide of the present invention (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule).

The 5' cap structure of a natural mRNA is involved in nuclear export, increasing mRNA stability and binds the mRNA Cap Binding Protein (CBP), which is responsible for mRNA stability in the cell and translation competency through the association of CBP with poly(A) binding protein to form the mature cyclic mRNA species. The cap further assists the removal of 5' proximal introns during mRNA splicing.

Endogenous mRNA molecules can be 5'-end capped generating a 5'-ppp-5'-triphosphate linkage between a terminal guanosine cap residue and the 5'-terminal transcribed sense nucleotide of the mRNA molecule. This 5'-guanylate cap can then be methylated to generate an N7-methyl-guanylate residue. The ribose sugars of the terminal and/or ante-terminal transcribed nucleotides of the 5' end of the mRNA can optionally also be 2'-O-methylated. 5'-decapping through hydrolysis and cleavage of the guanylate cap structure can target a nucleic acid molecule, such as an mRNA molecule, for degradation.

In some embodiments, the polynucleotides of the present invention (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule) incorporate a cap moiety.

5 In some embodiments, polynucleotides of the present invention (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule) comprise a non-hydrolyzable cap structure preventing decapping and thus increasing mRNA half-life. Because cap structure hydrolysis requires cleavage of 5'-ppp-5' phosphodiester
10 linkages, modified nucleotides can be used during the capping reaction. For example, a Vaccinia Capping Enzyme from New England Biolabs (Ipswich, MA) can be used with α -thio-guanosine nucleotides according to the manufacturer's instructions to create a phosphorothioate linkage in the 5'-ppp-5' cap. Additional modified guanosine nucleotides can be used such as α -methyl-phosphonate and seleno-phosphate
15 nucleotides.

Additional modifications include, but are not limited to, 2'-O-methylation of the ribose sugars of 5'-terminal and/or 5'-antiterminal nucleotides of the polynucleotide (as mentioned above) on the 2'-hydroxyl group of the sugar ring. Multiple distinct 5'-cap structures can be used to generate the 5'-cap of a nucleic acid molecule, such as a
20 polynucleotide that functions as an mRNA molecule. Cap analogs, which herein are also referred to as synthetic cap analogs, chemical caps, chemical cap analogs, or structural or functional cap analogs, differ from natural (i.e., endogenous, wild-type or physiological) 5'-caps in their chemical structure, while retaining cap function. Cap analogs can be chemically (i.e., non-enzymatically) or enzymatically synthesized and/or
25 linked to the polynucleotides of the invention.

For example, the Anti-Reverse Cap Analog (ARCA) cap contains two guanines linked by a 5'-5'-triphosphate group, wherein one guanine contains an N7 methyl group as well as a 3'-O-methyl group (i.e., N7,3'-O-dimethyl-guanosine-5'-triphosphate-5'-
30 m7G(5')ppp(5')G). The 3'-O atom of the other, unmodified, guanine becomes linked to

the 5'-terminal nucleotide of the capped polynucleotide. The N7- and 3'-O-methylated guanine provides the terminal moiety of the capped polynucleotide.

Another exemplary cap is mCAP, which is similar to ARCA but has a 2'-O-methyl group on guanosine (i.e., N7,2'-O-dimethyl-guanosine-5'-triphosphate-5'-
5 guanosine, m7Gm-ppp-G).

In some embodiments, the cap is a dinucleotide cap analog. As a non-limiting example, the dinucleotide cap analog can be modified at different phosphate positions with a boranophosphate group or a phosphoroselenoate group such as the dinucleotide cap analogs described in U.S. Patent No. US 8,519,110, the contents of which are herein
10 incorporated by reference in its entirety.

In another embodiment, the cap is a cap analog is a N7-(4-chlorophenoxyethyl) substituted dinucleotide form of a cap analog known in the art and/or described herein. Non-limiting examples of a N7-(4-chlorophenoxyethyl) substituted dinucleotide form of a cap analog include a N7-(4-chlorophenoxyethyl)-G(5')ppp(5')G and a N7-(4-
15 chlorophenoxyethyl)-m3'-OG(5')ppp(5')G cap analog (See, e.g., the various cap analogs and the methods of synthesizing cap analogs described in Kore et al. Bioorganic & Medicinal Chemistry 2013 21:4570-4574; the contents of which are herein incorporated by reference in its entirety). In another embodiment, a cap analog of the present invention is a 4-chloro/bromophenoxyethyl analog.

20 While cap analogs allow for the concomitant capping of a polynucleotide or a region thereof, in an in vitro transcription reaction, up to 20% of transcripts can remain uncapped. This, as well as the structural differences of a cap analog from an endogenous 5'-cap structures of nucleic acids produced by the endogenous, cellular transcription machinery, can lead to reduced translational competency and reduced
25 cellular stability.

Polynucleotides of the invention (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule) can also be capped post-manufacture (whether IVT or chemical synthesis), using enzymes, to generate more authentic 5'-cap structures. As
30 used herein, the phrase "more authentic" refers to a feature that closely mirrors or mimics, either structurally or functionally, an endogenous or wild type feature. That is,

a "more authentic" feature is better representative of an endogenous, wild-type, natural or physiological cellular function and/or structure as compared to synthetic features or analogs, etc., of the prior art, or which outperforms the corresponding endogenous, wild-type, natural or physiological feature in one or more respects. Non-limiting

5 examples of more authentic 5'cap structures of the present invention are those that, among other things, have enhanced binding of cap binding proteins, increased half-life, reduced susceptibility to 5' endonucleases and/or reduced 5'decapping, as compared to synthetic 5'cap structures known in the art (or to a wild-type, natural or physiological 5'cap structure). For example, recombinant Vaccinia Virus Capping Enzyme and

10 recombinant 2'-O-methyltransferase enzyme can create a canonical 5'-5'-triphosphate linkage between the 5'-terminal nucleotide of a polynucleotide and a guanine cap nucleotide wherein the cap guanine contains an N7 methylation and the 5'-terminal nucleotide of the mRNA contains a 2'-O-methyl. Such a structure is termed the Cap1 structure. This cap results in a higher translational-competency and cellular stability and

15 a reduced activation of cellular pro-inflammatory cytokines, as compared, e.g., to other 5'cap analog structures known in the art. Cap structures include, but are not limited to, 7mG(5')ppp(5')N,pN2p (cap 0), 7mG(5')ppp(5')NlmpNp (cap 1), and 7mG(5')-ppp(5')NlmpN2mp (cap 2).

As a non-limiting example, capping chimeric polynucleotides post-manufacture

20 can be more efficient as nearly 100% of the chimeric polynucleotides can be capped. This is in contrast to ~80% efficiency when a cap analog is linked to a chimeric polynucleotide during an in vitro transcription reaction.

According to the present invention, 5' terminal caps can include endogenous caps or cap analogs. According to the present invention, a 5' terminal cap can comprise

25 a guanine analog. Useful guanine analogs include, but are not limited to, inosine, N1-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

Poly A Tails

30 In some embodiments, the polynucleotides of the present disclosure (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or

prophylactic payload, an effector molecule and/or a tether molecule) further comprise a poly-A tail. In further embodiments, terminal groups on the poly-A tail can be incorporated for stabilization. In other embodiments, a poly-A tail comprises des-3' hydroxyl tails.

5 During RNA processing, a long chain of adenine nucleotides (poly-A tail) can be added to a polynucleotide such as an mRNA molecule to increase stability. Immediately after transcription, the 3' end of the transcript can be cleaved to free a 3' hydroxyl. Then poly-A polymerase adds a chain of adenine nucleotides to the RNA. The process, called polyadenylation, adds a poly-A tail that can be between, for
10 example, approximately 80 to approximately 250 residues long, including approximately 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240 or 250 residues long. In one embodiment, the poly-A tail is 100 nucleotides in length (SEQ ID NO:84).

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa

15 aaaaaaaaa aaaaaaaaaa aaaaaaaaaa (SEQ ID NO: 84)

PolyA tails can also be added after the construct is exported from the nucleus.

According to the present invention, terminal groups on the poly A tail can be incorporated for stabilization. Polynucleotides of the present invention can include des-
20 3' hydroxyl tails. They can also include structural moieties or 2'-O-methyl modifications as taught by Junjie Li, et al. (Current Biology, Vol. 15, 1501–1507, August 23, 2005, the contents of which are incorporated herein by reference in its entirety).

The polynucleotides of the present invention can be designed to encode transcripts with alternative polyA tail structures including histone mRNA. According to
25 Norbury, "Terminal uridylation has also been detected on human replication-dependent histone mRNAs. The turnover of these mRNAs is thought to be important for the prevention of potentially toxic histone accumulation following the completion or inhibition of chromosomal DNA replication. These mRNAs are distinguished by their lack of a 3' poly(A) tail, the function of which is instead assumed by a stable stem-loop
30 structure and its cognate stem-loop binding protein (SLBP); the latter carries out the same functions as those of PABP on polyadenylated mRNAs" (Norbury, "Cytoplasmic

RNA: a case of the tail wagging the dog," Nature Reviews Molecular Cell Biology; AOP, published online 29 August 2013; doi:10.1038/nrm3645) the contents of which are incorporated herein by reference in its entirety.

Unique poly-A tail lengths provide certain advantages to the polynucleotides of the present invention. Generally, the length of a poly-A tail, when present, is greater than 30 nucleotides in length. In another embodiment, the poly-A tail is greater than 35 nucleotides in length (e.g., at least or greater than about 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, and 3,000 nucleotides).

In some embodiments, the polynucleotide or region thereof includes from about 30 to about 3,000 nucleotides (e.g., from 30 to 50, from 30 to 100, from 30 to 250, from 30 to 500, from 30 to 750, from 30 to 1,000, from 30 to 1,500, from 30 to 2,000, from 30 to 2,500, from 50 to 100, from 50 to 250, from 50 to 500, from 50 to 750, from 50 to 1,000, from 50 to 1,500, from 50 to 2,000, from 50 to 2,500, from 50 to 3,000, from 100 to 500, from 100 to 750, from 100 to 1,000, from 100 to 1,500, from 100 to 2,000, from 100 to 2,500, from 100 to 3,000, from 500 to 750, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to 2,500, from 500 to 3,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 2,500, from 1,000 to 3,000, from 1,500 to 2,000, from 1,500 to 2,500, from 1,500 to 3,000, from 2,000 to 3,000, from 2,000 to 2,500, and from 2,500 to 3,000).

In some embodiments, the poly-A tail is designed relative to the length of the overall polynucleotide or the length of a particular region of the polynucleotide. This design can be based on the length of a coding region, the length of a particular feature or region or based on the length of the ultimate product expressed from the polynucleotides.

In this context, the poly-A tail can be 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100% greater in length than the polynucleotide or feature thereof. The poly-A tail can also be designed as a fraction of the polynucleotides to which it belongs. In this context, the poly-A tail can be 10, 20, 30, 40, 50, 60, 70, 80, or 90% or more of the total length of the construct, a construct region or the total length of the construct minus the poly-A

tail. Further, engineered binding sites and conjugation of polynucleotides for Poly-A binding protein can enhance expression.

Additionally, multiple distinct polynucleotides can be linked together via the PABP (Poly-A binding protein) through the 3'-end using modified nucleotides at the 3'-
 5 terminus of the poly-A tail. Transfection experiments can be conducted in relevant cell lines at and protein production can be assayed by ELISA at 12hr, 24hr, 48hr, 72hr and day 7 post-transfection.

In some embodiments, the polynucleotides of the present invention are designed to include a polyA-G Quartet region. The G-quartet is a cyclic hydrogen bonded array
 10 of four guanine nucleotides that can be formed by G-rich sequences in both DNA and RNA. In this embodiment, the G-quartet is incorporated at the end of the poly-A tail. The resultant polynucleotide is assayed for stability, protein production and other parameters including half-life at various time points. It has been discovered that the polyA-G quartet results in protein production from an mRNA equivalent to at least 75%
 15 of that seen using a poly-A tail of 120 nucleotides alone (SEQ ID NO:153).

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa (SEQ ID NO: 153)

Start codon region

20 The invention also includes a polynucleotide that comprises both a start codon region and the polynucleotide described herein (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule). In some embodiments, the polynucleotides of the present invention can have regions that are analogous to or function like a start codon
 25 region.

In some embodiments, the translation of a polynucleotide can initiate on a codon that is not the start codon AUG. Translation of the polynucleotide can initiate on an alternative start codon such as, but not limited to, ACG, AGG, AAG, CTG/CUG, GTG/GUG, ATA/AUA, ATT/AUU, TTG/UUG (see Touriol et al. *Biology of the Cell*
 30 95 (2003) 169-178 and Matsuda and Mauro *PLoS ONE*, 2010 5:11; the contents of each of which are herein incorporated by reference in its entirety).

As a non-limiting example, the translation of a polynucleotide begins on the alternative start codon ACG. As another non-limiting example, polynucleotide translation begins on the alternative start codon CTG or CUG. As another non-limiting example, the translation of a polynucleotide begins on the alternative start codon GTG or GUG.

Nucleotides flanking a codon that initiates translation such as, but not limited to, a start codon or an alternative start codon, are known to affect the translation efficiency, the length and/or the structure of the polynucleotide. (See, e.g., Matsuda and Mauro PLoS ONE, 2010 5:11; the contents of which are herein incorporated by reference in its entirety). Masking any of the nucleotides flanking a codon that initiates translation can be used to alter the position of translation initiation, translation efficiency, length and/or structure of a polynucleotide.

In some embodiments, a masking agent can be used near the start codon or alternative start codon to mask or hide the codon to reduce the probability of translation initiation at the masked start codon or alternative start codon. Non-limiting examples of masking agents include antisense locked nucleic acids (LNA) polynucleotides and exon-junction complexes (EJCs) (See, e.g., Matsuda and Mauro describing masking agents LNA polynucleotides and EJCs (PLoS ONE, 2010 5:11); the contents of which are herein incorporated by reference in its entirety).

In another embodiment, a masking agent can be used to mask a start codon of a polynucleotide to increase the likelihood that translation will initiate on an alternative start codon. In some embodiments, a masking agent can be used to mask a first start codon or alternative start codon to increase the chance that translation will initiate on a start codon or alternative start codon downstream to the masked start codon or alternative start codon.

In some embodiments, a start codon or alternative start codon can be located within a perfect complement for a miRNA binding site. The perfect complement of a miRNA binding site can help control the translation, length and/or structure of the polynucleotide similar to a masking agent. As a non-limiting example, the start codon or alternative start codon can be located in the middle of a perfect complement for a miRNA binding site. The start codon or alternative start codon can be located after the

first nucleotide, second nucleotide, third nucleotide, fourth nucleotide, fifth nucleotide, sixth nucleotide, seventh nucleotide, eighth nucleotide, ninth nucleotide, tenth nucleotide, eleventh nucleotide, twelfth nucleotide, thirteenth nucleotide, fourteenth nucleotide, fifteenth nucleotide, sixteenth nucleotide, seventeenth nucleotide, eighteenth nucleotide, nineteenth nucleotide, twentieth nucleotide or twenty-first nucleotide.

In another embodiment, the start codon of a polynucleotide can be removed from the polynucleotide sequence to have the translation of the polynucleotide begin on a codon that is not the start codon. Translation of the polynucleotide can begin on the codon following the removed start codon or on a downstream start codon or an alternative start codon. In a non-limiting example, the start codon ATG or AUG is removed as the first 3 nucleotides of the polynucleotide sequence to have translation initiate on a downstream start codon or alternative start codon. The polynucleotide sequence where the start codon was removed can further comprise at least one masking agent for the downstream start codon and/or alternative start codons to control or attempt to control the initiation of translation, the length of the polynucleotide and/or the structure of the polynucleotide.

Stop codon region

The invention also includes a polynucleotide that comprises both a stop codon region and the polynucleotide described herein (e.g., a polynucleotide comprising a nucleotide sequence encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule). In some embodiments, the polynucleotides of the present invention can include at least two stop codons before the 3' untranslated region (UTR). The stop codon can be selected from TGA, TAA and TAG in the case of DNA, or from UGA, UAA and UAG in the case of RNA. In some embodiments, the polynucleotides of the present invention include the stop codon TGA in the case of DNA, or the stop codon UGA in the case of RNA, and one additional stop codon. In a further embodiment the additional stop codon can be TAA or UAA. In another embodiment, the polynucleotides of the present invention include three consecutive stop codons, four stop codons, or more.

Protease cleavage site or IRES

The invention also includes a polynucleotide that comprises a protease cleavage site or an internal ribosomal entry site. In some embodiments, the LNP compositions or systems of the present disclosure comprise: (a) a first polynucleotide (e.g., mRNA) comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide (e.g., mRNA) comprising a sequence encoding: (1) an effector molecule, and (2) a polypeptide that recognizes the binding element (a tether molecule).

10 In some embodiments, the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

In some embodiments, the first polynucleotide and the second polynucleotide are separated by a protease cleavage site, e.g., a T2A site or P2A site or E2A site, or TPE (P2A-T2A-E2A, or other known protease cleavage sites.

15 In some embodiments, the first polynucleotide and the second polynucleotide are separated by an IRES.

Methods of making polynucleotides

The present disclosure also provides methods for making a polynucleotide disclosed herein or a complement thereof. In some aspects, a polynucleotide (e.g., an mRNA) disclosed herein encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule can be constructed using in vitro transcription.

In other aspects, a polynucleotide (e.g., an mRNA) disclosed herein encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule can be constructed by chemical synthesis using an oligonucleotide synthesizer. In other aspects, a polynucleotide (e.g., an mRNA) disclosed herein encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule is made by using a host cell. In certain aspects, a polynucleotide (e.g., an mRNA) disclosed herein encoding a therapeutic payload or prophylactic payload, an

effector molecule and/or a tether molecule is made by one or more combination of the IVT, chemical synthesis, host cell expression, or any other methods known in the art.

Naturally occurring nucleosides, non-naturally occurring nucleosides, or combinations thereof, can totally or partially naturally replace occurring nucleosides present in the candidate nucleotide sequence and can be incorporated into a sequence-
5 optimized nucleotide sequence (e.g., an mRNA) encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule. The resultant mRNAs can then be examined for their ability to produce protein and/or produce a therapeutic outcome.

10 Exemplary methods of making a polynucleotide disclosed herein include: in vitro transcription enzymatic synthesis and chemical synthesis which are disclosed in International PCT application WO 2017/201325, filed on 18 May 2017, the entire contents of which are hereby incorporated by reference.

15 Purification

In other aspects, a polynucleotide (e.g., an mRNA) disclosed herein encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule can be purified. Purification of the polynucleotides (e.g., mRNA) encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether
20 molecule described herein can include, but is not limited to, polynucleotide clean-up, quality assurance and quality control. Clean-up can be performed by methods known in the arts such as, but not limited to, AGENCOURT® beads (Beckman Coulter Genomics, Danvers, MA), poly-T beads, LNATM oligo-T capture probes (EXIQON® Inc, Vedbaek, Denmark) or HPLC based purification methods such as, but not limited
25 to, strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC). The term "purified" when used in relation to a polynucleotide such as a "purified polynucleotide" refers to one that is separated from at least one contaminant. As used herein, a "contaminant" is any substance which makes another unfit, impure or inferior. Thus, a purified
30 polynucleotide (e.g., DNA and RNA) is present in a form or setting different from that

in which it is found in nature, or a form or setting different from that which existed prior to subjecting it to a treatment or purification method.

In some embodiments, purification of a polynucleotide (e.g., mRNA) encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule of the disclosure removes impurities that can reduce or remove an unwanted immune response, e.g., reducing cytokine activity.

In some embodiments, the polynucleotide (e.g., mRNA) encoding a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule of the disclosure is purified prior to administration using column chromatography (e.g., strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), or (LCMS)). In some embodiments, a column chromatography (e.g., strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), or (LCMS)) purified polynucleotide, which encodes a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule disclosed herein increases expression of the therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule compared to polynucleotides encoding the therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule purified by a different purification method.

In some embodiments, a column chromatography (e.g., strong anion exchange HPLC, weak anion exchange HPLC, reverse phase HPLC (RP-HPLC), and hydrophobic interaction HPLC (HIC-HPLC), or (LCMS)) purified polynucleotide encodes a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule. In some embodiments, the purified polynucleotide encodes a therapeutic payload or prophylactic payload, an effector molecule and/or a tether molecule.

In some embodiments, the purified polynucleotide is at least about 80% pure, at least about 85% pure, at least about 90% pure, at least about 95% pure, at least about 96% pure, at least about 97% pure, at least about 98% pure, at least about 99% pure, or about 100% pure.

A quality assurance and/or quality control check can be conducted using methods such as, but not limited to, gel electrophoresis, UV absorbance, or analytical HPLC.

In another embodiment, the polynucleotides can be sequenced by methods including, but not limited to reverse-transcriptase-PCR.

Chemical modifications of polynucleotides

The present disclosure provides for modified nucleosides and nucleotides of a nucleic acid (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids). A “nucleoside” refers to a compound containing a sugar molecule (*e.g.*, a pentose or ribose) or a derivative thereof in combination with an organic base (*e.g.*, a purine or pyrimidine) or a derivative thereof (also referred to herein as “nucleobase”). A “nucleotide” refers to a nucleoside, including a phosphate group. Modified nucleotides may be synthesized by any useful method, such as, for example, chemically, enzymatically, or recombinantly, to include one or more modified or non-natural nucleosides. Nucleic acids can comprise a region or regions of linked nucleosides. Such regions may have variable backbone linkages. The linkages can be standard phosphodiester linkages, in which case the nucleic acids would comprise regions of nucleotides.

Modified nucleotide base pairing encompasses not only the standard adenosine-thymine, adenosine-uracil, or guanosine-cytosine base pairs, but also base pairs formed between nucleotides and/or modified nucleotides comprising non-standard or modified bases, wherein the arrangement of hydrogen bond donors and hydrogen bond acceptors permits hydrogen bonding between a non-standard base and a standard base or between two complementary non-standard base structures, such as, for example, in those nucleic acids having at least one chemical modification. One example of such non-standard base pairing is the base pairing between the modified nucleotide inosine and adenine, cytosine or uracil. Any combination of base/sugar or linker may be incorporated into nucleic acids of the present disclosure.

In some embodiments, modified nucleobases in nucleic acids (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids) comprise N1-methyl-pseudouridine (m1ψ), 1-ethyl-pseudouridine (e1ψ), 5-methoxy-uridine (mo5U), 5-methyl-cytidine (m5C), and/or

pseudouridine (ψ). In some embodiments, modified nucleobases in nucleic acids (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids) comprise 5-methoxymethyl uridine, 5-methylthio uridine, 1-methoxymethyl pseudouridine, 5-methyl cytidine, and/or 5-methoxy cytidine. In some embodiments, the polyribonucleotide includes a
5 combination of at least two (*e.g.*, 2, 3, 4 or more) of any of the aforementioned modified nucleobases, including but not limited to chemical modifications.

In some embodiments, a RNA nucleic acid of the disclosure comprises N1-methyl-pseudouridine (m1 ψ) substitutions at one or more or all uridine positions of the nucleic acid.

10 In some embodiments, a RNA nucleic acid of the disclosure comprises N1-methyl-pseudouridine (m1 ψ) substitutions at one or more or all uridine positions of the nucleic acid and 5-methyl cytidine substitutions at one or more or all cytidine positions of the nucleic acid.

In some embodiments, a RNA nucleic acid of the disclosure comprises
15 pseudouridine (ψ) substitutions at one or more or all uridine positions of the nucleic acid.

In some embodiments, a RNA nucleic acid of the disclosure comprises pseudouridine (ψ) substitutions at one or more or all uridine positions of the nucleic acid and 5-methyl cytidine substitutions at one or more or all cytidine positions of the
20 nucleic acid.

In some embodiments, a RNA nucleic acid of the disclosure comprises uridine at one or more or all uridine positions of the nucleic acid.

In some embodiments, nucleic acids (*e.g.*, RNA nucleic acids, such as mRNA nucleic acids) are uniformly modified (*e.g.*, fully modified, modified throughout the
25 entire sequence) for a particular modification. For example, a nucleic acid can be uniformly modified with N1-methyl-pseudouridine, meaning that all uridine residues in the mRNA sequence are replaced with N1-methyl-pseudouridine. Similarly, a nucleic acid can be uniformly modified for any type of nucleoside residue present in the sequence by replacement with a modified residue such as those set forth above.

30 The nucleic acids of the present disclosure may be partially or fully modified along the entire length of the molecule. For example, one or more or all or a given type

of nucleotide (*e.g.*, purine or pyrimidine, or any one or more or all of A, G, U, C) may be uniformly modified in a nucleic acid of the disclosure, or in a predetermined sequence region thereof (*e.g.*, in the mRNA including or excluding the polyA tail). In some embodiments, all nucleotides X in a nucleic acid of the present disclosure (or in a sequence region thereof) are modified nucleotides, wherein X may be any one of nucleotides A, G, U, C, or any one of the combinations A+G, A+U, A+C, G+U, G+C, U+C, A+G+U, A+G+C, G+U+C or A+G+C.

The nucleic acid may contain from about 1% to about 100% modified nucleotides (either in relation to overall nucleotide content, or in relation to one or more types of nucleotide, *i.e.*, any one or more of A, G, U or C) or any intervening percentage (*e.g.*, from 1% to 20%, from 1% to 25%, from 1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 20% to 70%, from 20% to 80%, from 20% to 90%, from 20% to 95%, from 20% to 100%, from 50% to 60%, from 50% to 70%, from 50% to 80%, from 50% to 90%, from 50% to 95%, from 50% to 100%, from 70% to 80%, from 70% to 90%, from 70% to 95%, from 70% to 100%, from 80% to 90%, from 80% to 95%, from 80% to 100%, from 90% to 95%, from 90% to 100%, and from 95% to 100%). It will be understood that any remaining percentage is accounted for by the presence of unmodified A, G, U, or C.

The nucleic acids may contain at a minimum 1% and at maximum 100% modified nucleotides, or any intervening percentage, such as at least 5% modified nucleotides, at least 10% modified nucleotides, at least 25% modified nucleotides, at least 50% modified nucleotides, at least 80% modified nucleotides, or at least 90% modified nucleotides. For example, the nucleic acids may contain a modified pyrimidine such as a modified uracil or cytosine. In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 80%, at least 90% or 100% of the uracil in the nucleic acid is replaced with a modified uracil (*e.g.*, a 5-substituted uracil). The modified uracil can be replaced by a compound having a single unique structure, or can be replaced by a plurality of compounds having different structures (*e.g.*, 2, 3, 4 or more

unique structures). In some embodiments, at least 5%, at least 10%, at least 25%, at least 50%, at least 80%, at least 90% or 100% of the cytosine in the nucleic acid is replaced with a modified cytosine (*e.g.*, a 5-substituted cytosine). The modified cytosine can be replaced by a compound having a single unique structure, or can be replaced by a plurality of compounds having different structures (*e.g.*, 2, 3, 4 or more unique structures).

Pharmaceutical compositions

The present disclosure provides pharmaceutical formulations comprising any of the systems, or LNP compositions disclosed herein, *e.g.*, a system or an LNP composition comprising: (a) a first polynucleotide (*e.g.*, mRNA) comprising: (1) a sequence encoding a therapeutic payload or prophylactic payload, and (2) a binding element; and (b) a second polynucleotide (*e.g.*, mRNA) comprising a sequence encoding: (1) an effector molecule, and (2) a polypeptide that recognizes the binding element (a tether molecule).

In some embodiments of the disclosure, the polynucleotide are formulated in compositions and complexes in combination with one or more pharmaceutically acceptable excipients. Pharmaceutical compositions can optionally comprise one or more additional active substances, *e.g.* therapeutically and/or prophylactically active substances. Pharmaceutical compositions of the present disclosure can be sterile and/or pyrogen-free. General considerations in the formulation and/or manufacture of pharmaceutical agents can be found, for example, in Remington: The Science and Practice of Pharmacy 21st ed., Lippincott Williams & Wilkins, 2005.

In some embodiments, compositions are administered to humans, human patients or subjects. For the purposes of the present disclosure, the phrase "active ingredient" generally refers to polynucleotides to be delivered as described herein.

Although the descriptions of pharmaceutical compositions provided herein are principally directed to pharmaceutical compositions which are suitable for administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to any other animal, *e.g.*, to non-

human animals, e.g. non-human mammals. Modification of pharmaceutical compositions suitable for administration to humans in order to render the compositions suitable for administration to various animals is well understood, and the ordinarily skilled veterinary pharmacologist can design and/or perform such modification with
5 merely ordinary, if any, experimentation. Subjects to which administration of the pharmaceutical compositions is contemplated include, but are not limited to, humans and/or other primates; mammals.

In some embodiments, the polynucleotide of the present disclosure is formulated for subcutaneous, intravenous, intraperitoneal, intramuscular, intra-articular, intra-
10 synovial, intrasternal, intrathecal, intrahepatic, intralesional, intracranial, intraventricular, oral, inhalation spray, pulmonary, topical, rectal, nasal, buccal, vaginal, or implanted reservoir intramuscular, subcutaneous, or intradermal delivery. In other embodiments, the polynucleotide is formulated for subcutaneous or intravenous delivery.

15 Formulations of the pharmaceutical compositions described herein can be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of bringing the active ingredient into association with an excipient and/or one or more other accessory ingredients, and then, if necessary and/or desirable, dividing, shaping and/or packaging the product into a
20 desired single- or multi-dose unit.

Relative amounts of the active ingredient, the pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the disclosure will vary, depending upon the identity, size, and/or condition of the subject treated and further depending upon the route by which the
25 composition is to be administered. By way of example, the composition can comprise between 0.1% and 100%, e.g., between 0.5% and 50%, between 1% and 30%, between 5% and 80%, or at least 80% (w/w) active ingredient.

Formulations and delivery

30 The polynucleotide comprising an mRNA of the disclosure can be formulated using one or more excipients.

The function of the one or more excipients is, e.g., to: (1) increase stability; (2) increase cell transfection; (3) permit the sustained or delayed release (e.g., from a depot formulation of the polynucleotide); (4) alter the biodistribution (e.g., target the polynucleotide to specific tissues or cell types); (5) increase the translation of encoded protein in vivo; and/or (6) alter the release profile of encoded protein in vivo. In addition to traditional excipients such as any and all solvents, dispersion media, diluents, or other liquid vehicles, dispersion or suspension aids, surface active agents, isotonic agents, thickening or emulsifying agents, preservatives, excipients of the present disclosure can include, without limitation, lipidoids, liposomes, lipid nanoparticles, polymers, lipoplexes, core-shell nanoparticles, peptides, proteins, cells transfected with polynucleotides (e.g., for transplantation into a subject), hyaluronidase, nanoparticle mimics and combinations thereof. Accordingly, the formulations of the disclosure can include one or more excipients, each in an amount that together increases the stability of the polynucleotide, increases cell transfection by the polynucleotide, increases the expression of polynucleotides encoded protein, and/or alters the release profile of polynucleotide encoded proteins. Further, the polynucleotides of the present disclosure can be formulated using self-assembled nucleic acid nanoparticles.

Formulations of the pharmaceutical compositions described herein can be prepared by any method known or hereafter developed in the art of pharmacology. In general, such preparatory methods include the step of associating the active ingredient with an excipient and/or one or more other accessory ingredients.

A pharmaceutical composition in accordance with the present disclosure can be prepared, packaged, and/or sold in bulk, as a single unit dose, and/or as a plurality of single unit doses. As used herein, a "unit dose" refers to a discrete amount of the pharmaceutical composition comprising a predetermined amount of the active ingredient. The amount of the active ingredient is generally equal to the dosage of the active ingredient which would be administered to a subject and/or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage.

Relative amounts of the active ingredient, the pharmaceutically acceptable excipient, and/or any additional ingredients in a pharmaceutical composition in accordance with the present disclosure can vary, depending upon the identity, size,

and/or condition of the subject being treated and further depending upon the route by which the composition is to be administered. For example, the composition can comprise between 0.1% and 99% (w/w) of the active ingredient. By way of example, the composition can comprise between 0.1% and 100%, e.g., between .5 and 50%,
5 between 1-30%, between 5-80%, at least 80% (w/w) active ingredient.

In some embodiments, the formulations described herein contain at least one polynucleotide. As a non-limiting example, the formulations contain 1, 2, 3, 4 or 5 polynucleotides.

Pharmaceutical formulations can additionally comprise a pharmaceutically
10 acceptable excipient, which, as used herein, includes, but is not limited to, any and all solvents, dispersion media, diluents, or other liquid vehicles, dispersion or suspension aids, surface active agents, isotonic agents, thickening or emulsifying agents, preservatives, and the like, as suited to the particular dosage form desired. Various excipients for formulating pharmaceutical compositions and techniques for preparing
15 the composition are known in the art (see Remington: The Science and Practice of Pharmacy, 21st Edition, A. R. Gennaro, Lippincott, Williams & Wilkins, Baltimore, MD, 2006). The use of a conventional excipient medium can be contemplated within the scope of the present disclosure, except insofar as any conventional excipient medium can be incompatible with a substance or its derivatives, such as by producing
20 any undesirable biological effect or otherwise interacting in a deleterious manner with any other component(s) of the pharmaceutical composition.

In some embodiments, the particle size of the lipid nanoparticle is increased and/or decreased. The change in particle size can be able to help counter biological reaction such as, but not limited to, inflammation or can increase the biological effect of
25 the modified mRNA delivered to mammals.

Pharmaceutically acceptable excipients used in the manufacture of pharmaceutical compositions include, but are not limited to, inert diluents, surface active agents and/or emulsifiers, preservatives, buffering agents, lubricating agents, and/or oils. Such excipients can optionally be included in the pharmaceutical
30 formulations of the disclosure.

In some embodiments, the polynucleotides is administered in or with, formulated in or delivered with nanostructures that can sequester molecules such as cholesterol. Non-limiting examples of these nanostructures and methods of making these nanostructures are described in US Patent Publication No. US20130195759. Exemplary structures of these nanostructures are shown in US Patent Publication No. US20130195759, and can include a core and a shell surrounding the core.

A polynucleotide comprising an mRNA of the disclosure can be delivered to a cell using any method known in the art. For example, the polynucleotide comprising an mRNA of the disclosure can be delivered to a cell by a lipid-based delivery, e.g., transfection, or by electroporation.

Table 6: Miscellaneous Sequences

SEQ ID NO	Sequence information	Sequence
80	EPO (aa)	MGVHECPAWLWLLLSLLSLPLGLPVLGAPRLICDSRVLERYLLEA KEAENITTGCAEHCSLNENITVPDTKVNIFYAWKRMEVGGQAVEVWQG LALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRALGA QKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKCLKLYTGEACRTG DR
81	EPO (nt)	ATGGGAGTGCACGAGTGTCCCGCGTGGTTGTGGTTGCTGCTGTC GCTCTTGAGCCTCCCACTGGGACTGCCTGTGCTGGGGGCACCACCA GATTGATCTGCGACTCACGGGTACTTGAGAGGTACCTTCTTGAAGCC AAAGAAGCCGAAAACATCACAAACCGGATGCGCCGAGCACTGCTCCC TCAATGAGAACATTACTGTACCGGATACAAAGGTCAATTTCTATGCA TGGAAGAGAATGGAAGTAGGACAGCAGGCCGTCGAAGTGTGGCAG GGGCTCGCGCTTTTGTGCGGAGGCGGTGTTGCGGGGTCAGGCCCTCT CGTCAACTCATCACAGCCGTGGGAGCCCCCTCCAACCTTCATGTCGATA AAGCGGTGTGCGGGGCTCCGCAGCTTGACGACGTTGCTTCGGGCTCTG GGCGCACAAAAGGAGGCTATTTTCGCCGCCTGACGCGGCCTCCGCGG CACCCCTCCGAACGATCACCGCGGACACGTTTAGGAAGCTTTTAGA GTGTACAGCAATTTCTCCGCGGAAAGCTGAAATTGTATACTGGTGA AGCGGTGTAGGACAGGGGATCGC
82	deg-GFP target (aa)	MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYGKLT KFICTTGKLPVPWPTLVTTITYGVQCFSRYPDHMKQHDFFKSAMPEGY VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKL EYNYNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIG DGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELY KRSRDISHGFPPAVAAQDDGTLPMSCAQESGMDRHPAACASARINV
83	deg-GFP target (nt)	ATGGTGTCCAAGGGTGAGGAATTGTTTACCGGGTGTTGCTGCCTAT TCTCGTCGAACCTTGACGGGGATGTGAATGGACACAAGTTTTTCGGTAT CCGGAGAAGGAGAGGGTGACGCCACATACGGAAGCTTACACTCAA ATTCATCTGTACGACGGGGAACTGCCCGTACCTGGCCTACGCTCG TAACCACGCTGACTTATGGAGTGCAGTGCTTTAGCAGATACCCCGAC CATATGAAGCAGCACGACTTCTTCAAGTCGGCGATGCCCCGAGGGT

		ACGTGCAAGAGAGGACCATTTTCTTCAAAGACGATGGCAATTACAA AACACGCGCAGAAGTCAAGTTTGAGGGCGATACTCTGGTCAATCGG ATCGAATTGAAGGGAATCGATTTCAAAGAAGATGGAAACATCCTTG GCCATAAGCTCGAGTACAACATACTCGCATAATGTCTATATCATG GCTGACAAGCAGAAAAACGGTATCAAAGTCAACTTTAAGATCCGAC ACAATATTGAGGACGGTTCGGTGCAGCTTGCAGGACCACTATCAACA GAATACGCCGATTGGGGATGGTCCGGTCCTTTTGCCGGATAACCATT ATCTCTCAACCCAGTCAGCCCTGAGCAAAGATCCAAACGAGAAGAG GGACCACATGGTCTTGCTCGAATTCTGTGACAGCGGCAGGGATCACTC TGGGAATGGACGAGTTGTACAAGAGATCTCGAGATATCAGCCATGG CTTCCCGCCGGCGGTGGCGGCGCAGGATGATGGCACGCTGCCCATGT CTTGTGCCCAGGAGAGCGGGATGGACCGTCACCCTGCAGCCTGTGCT TCTGCTAGGATCAATGTG
84	polyA tail_100 (nt)	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA
85	Luc target (aa)	MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDA HIEVDITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSNSLQFFMPVLG ALFIGVAVAPANDIYNERELLSMISQPTVVVFVSKKGLQKILNVQKKL PIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALI MNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNIIPDTAILSVVPFHH GFGMFTTLGYLICGFRVVLMYRFEELFLRSLQDYKIQSALLVPTLFSFF AKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTE TTSAILITPEGDDKPGAVGKVVPPFEAKVVDLDTGKTLGVNQRGELCVR GPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLI KYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHG KTMTEKEIVDYVASQVTTAKKLRRGVVVFVDEVPKGLTGKLDARKIREI LIKAKKGGKIAV
86	Luc target (nt)	ATGGAAGATGCGAAGAACATCAAGAAGGGACCTGCCCCGTTTTA CCCTTTGGAGGACGGTACAGCAGGAGAACAGCTCCACAAGGCGATG AAACGCTACGCCCTGGTCCCCGGAACGATTGCGTTTACCGATGCACA TATTGAGGTAGACATCACATACGCAGAATACTTCGAAATGTCGGTGA GGCTGGCGGAAGCGATGAAGAGATATGGTCTTAACACTAATCACCG CATCGTGGTGTGTTTCGGAGAACTCATTGCAGTTTTTCATGCCGGTCCT TGGAGCACTTTTCATCGGGGTCGCAGTCGCGCCAGCGAACGACATCT ACAATGAGCGGGAACCTCTTGAATAGCATGGGAATCTCCCAGCCGAC GGTCGTGTTTGTCTCCAAAAAGGGGCTGCAGAAAATCCTCAACGTGC AGAAGAAGCTCCCCATTATTCAAAAGATCATCATTATGGATAGCAA GACAGATTACCAAGGGTTCAGTCGATGTATACCTTTGTGACATCGC ATTTGCCGCCAGGGTTTAACGAGTATGACTTCGTCCCCGAGTCATTT GACAGAGATAAAACCATCGCGCTGATTATGAATTCCTCGGGTAGCA CCGTTTGGCAAAGGGGGTGGCGTTGCCCCACCGCACTGCTTGTGTG CGTTTCTCGCACGCTAGGGATCCTATCTTTGGTAATCAGATCATTC CGACACAGCAATCCTGTCCGTGGTACCTTTTCATCACGGTTTTTGGCA TGTTACGACTCTCGGCTATTTGATTTGCGGTTTCAGGGTCGTACTTA TGTATCGGTTTCGAGGAAGAACTGTTTTTGAGATCCTTGCAAGATTAC AAGATCCAGTCGGCCCTCCTTGTGCCAACGCTTTTCTCATTCTTTGCG AAATCGACACTTATTGATAAGTATGACCTTTCCAATCTGCATGAGAT TGCCTCAGGGGGAGCGCCGCTTAGCAAGGAAGTCGGGGAGGCAGTG GCCAAGCGCTTCCACCTTCCCGGAATTCGGCAGGGATACGGGCTCAC GGAGACAACATCCGCGATCCTTATCACGCCGAGGGTGACGATAAG CCGGGAGCCGTCGGAAAAGTGGTCCCCTTCTTTGAAGCCAAGGTCGT AGACCTCGACACGGGAAAAACCTCGGAGTGAACCAGAGGGGCGA GCTCTGCGTGAGAGGGCCGATGATCATGTCAGGTTACGTGAATAACC CTGAAGCGACGAATGCGCTGATCGACAAGGATGGGTGGTTGCATTC

		GGGAGACATTGCCTATTGGGATGAGGATGAGCACTTCTTTATCGTAG ATCGACTTAAGAGCTTGATCAAATACAAAGGCTATCAGGTAGCGCCT GCCGAGCTCGAGTCAATCCTGCTCCAGCACCCCAACATTTTCGACGC CGGAGTGGCCGGGTTGCCCGATGACGACGCGGGTGAGCTGCCAGCG GCCGTGGTAGTCCTCGAACATGGGAAAACAATGACCGAAAAGGAGA TCGTGGACTACGTAGCATCACAAAGTGACGACTGCGAAGAACTGAG GGGAGGGGTAGTCTTTGTGGACGAGGTCCCGAAAGGCTTGACTGGG AAGCTTGACGCTCGCAAAATCCGGGAAATCCTGATTAAGGCAAAGA AAGGCGGGAAAATCGCTGTC
87	TPE element (aa)	GSGEGRGSLTTCGDVEENPGPGSGATNFSLLKQAGDVEENPGPGSG QCTNYALLKLAGDVESNPGP
88	TPE element (nt)	GGCTCTGGCGAAGGCCGGGGTAGCCTGCTGACCTGCGGCGACGT GGAGGAGAACCCTGGACCTGGATCCGGCGCCACCAACTTCAGCCTG CTCAAGCAGGCCGGAGACGTGGAAGAGAATCCAGGCCAGGCAGCG GTCAGTGCACCAACTACGCCCTGCTGAAGCTGGCCGGCGACGTAGA GAGCAACCCCGGCCCC
89	T2A self- cleaving peptide (aa)*	(GSG)EGRGSLTTCGDVEENPGP
90	P2A self- cleaving peptide (aa)*	(GSG)ATNFSLLKQAGDVEENPGP
91	E2A self- cleaving peptide (aa)*	(GSG)QCTNYALLKLAGDVESNPGP
92	F2A self- cleaving peptide (aa)*	(GSG)VKQTLNFDLLKLAGDVESNPGP
93	LUC_TPE_M BPmid2 (aa)	MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDA HIEVDITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLG ALFIGVAVAPANDIYNERELLSMISQPTVVFVSKKGLQKILNVQKKL PIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALI MNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILSVVPFHH GFGMFTTLGYLICGFRVVLMYRFEELFLRSLQDYKIQSALLVPTLFSFF AKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTE TTSAILITPEGDDKPGAVGKVPFFFEAKVVDLDTGKTLGVNQRGELCVR GPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLI KYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDDAGELPAAVVVLEHG KTMTEKEIVDYVASQVTTAKKLRRGGVVFVDEVPKGLTGKLDARKIREI LIKAKKGGKIAVSGEGRGSLTTCGDVEENPGPGSGATNFSLLKQAGD VEENPGPGSGQCTNYALLKLAGDVESNPGPASNFTQFVLVDNGGTGDV TVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKG AWRSYLNMEITIFATNSDCELVKAMQGLLDGNPIPSAIAANSIYIG GGGSDPTRLQGINCSPDFTPSFANLGRITLSTRGPPRGGPGGELPRGPAG LGPRRSQQGPRKEPRKIIATVLMTEIDIKLNAEKAWKPSSKRTAADKDR GEEDADGSKTQDLFRRVRSIANKATPQMAQQLMKQVTLAIDTEERLK GVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTVTNVNFRKLLL NRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGRLEKEELEEARDIAR RRSLGNIKFIGELFKLKMLTEAIMHDCVVKLLKNHDEESLECLCRLLTTI GKDLDFEKAKPRMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDRGSNW VPRRGDQGPKTIDQIHKEAEMEEHREHIKVQQLMAKGSDKRRGGPPGP

		PISRGLPLVDDGGWNTVPISKGSRPIDTSRLTKITKPGSIDSNNQLFAPGG RLSWGKGSSGGSGAKPSDAASEAARPATSTLNRFSALQQAV
94	LUC_TPE_M BPmid2 (nt)	ATGGAGGACGCCAAGAACATCAAGAAGGGGGCCGCTCCCTTCTA CCCGCTGGAGGACGGCACCCTGCGGAGCAGCTGCACAAGGCCATG AAGAGGTACGCACTGGTGCCCGGCACAATCGCCTTACCCGACGCCC ACATCGAGGTGGACATCACCTACGCCGAGTACTTCGAGATGAGCGT GCGGCTGGCCGAAGCCATGAAGCGGTACGGCCTGAACACCAACCAC CGCATCGTGGTGTGCTCCGAGAACTCCCTGCAGTTCTTCATGCCCCT TCTGGGCGCCCTGTTTCATTGGGGTGGCCGTGGCACCAGCCAACGACA TCTACAACGAGAGGGGAGCTGCTGAACAGCATGGGCATCAGCCAGCC CACCGTGGTGTTCGTGTCCAAGAAGGGCCTGCAGAAGATCCTGAAC GTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACA GCAAGACCGACTACCAGGGCTTCCAGTCCATGTACACCTTCGTGACC AGCCACCTGCCTCCCGGCTTCAACGAGTACGACTTCGTGCCCGAGTC CTTTCGACAGGGACAAGACCATCGCCCTGATCATGAACTCCAGCGGA AGCACCGGCCTGCCCCAAGGGTGTGGCTTACCCACCGGACCGCCTG TGTACGGTTCAGCCACGCCCCTGACCCATCTTCGGGAACCAAGATCA TCCCCGACACCGCCATCCTGAGCGTGGTGCCCTTCCACCAACGCGTTC GGCATGTTACACACCCTGGGCTACCTGATCTCGGGCTTCCGGGTGGT GCTGATGTACCGGTTTCAGAGGAGCTGTTCTCGGGAGCCTGCAG GACTACAAGATCCAGTCCGCCCTGCTGGTGGCCACCCTGTTCTCCTT CTTCGCCAAGTCCACCCTGATCGACAAGTACGACCTGAGCAACCTGC ACGAGATTGCCTCTGGCGGCGCACCTCTGTCCAAGGAGGTGCGGGA GGCCGTAGCCAAGCGGTTCCACCTGCCCGGCATCCGGCAAGGCTAC GGCCTGACCGAGACCACCAGCGCCATCCTGATCACCCCTGAGGGCG ACGACAAGCCCGGCGCAGTGGGGAAGGTGGTGGCATTCTTCGAGGC CAAGGTGGTGGACCTGGACACCGGCAAGACCCTGGGCGTGAACCAG CGGGGCGAGCTGTGCGTGCGGGGACCCATGATCATGTCCGGCTACG TGAACAACCCCGAGGCCACCAACGCCCTGATTGACAAGGACGGGTG GCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTC TTCATCGTGGACCGCCTGAAGTCCCTGATCAAGTACAAGGGGTACCA GGTGGCTCCCGCCGAGCTGGAGTCCATCCTGCTGCAGACCCCAACA TCTTTGACGCCGGCGTTGCCGACTGCCTGACGACGACGCTGGGGAA CTCCCCGAGCTGTGGTGGTGTGGAGCACGGCAAGACCATGACCG AGAAGGAGATCGTGGACTACGTGCGCTCCCAGGTGACCACCGCCAA GAAGTTACGCGGCGGCGTCGTGTTTCGTGGACGAGGTGCCCAAGGGC CTGACCGGCAAGCTGGACGCCCGCAAGATCCGGGAGATCCTGATCA AGGCCAAGAAGGGCGGGAAGATCGCCGTGGGCTCTGGCGAAGGCCG GGGTAGCCTGCTGACCTGCGGCGACGTGGAGGAGAACCCTGGACCT GGATCCGGCGCCACCAACTTCAGCCTGCTCAAGCAGGCCGGAGACG TGAAGAGAATCCAGGCCCAGGCAGCGGTCAAGTGCACCAACTACGC CCTGCTGAAGCTGGCCGGCGACGTAGAGAGCAACCCCGGCCCGCC AGCAACTTCACCCAGTTCTGTGGTGGTGGACAACCGCGGTACCGGAG ACGTGACCGTGGCCCTTCTAACTTCGCCAACGGCATCGCCGAGTGG ATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGACCTGCAGCG TGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATCAAGGTGGA GGTGCCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGGAGCTGACC ATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATCGTGAAGGC CATGCAGGGCCTGCTGAAGGACGGCAACCCCTATCCCTAGCGCCATC GCCGCCAATTACGGCATCTACGGTGGCGGAGGCAGCGACCCTACCA GACTGCAGGGCATCAACTGCGGCCCTGACTTCACTCCTTCTTTTCGA AACCTGGGCAGAACCAACCCTGAGCACCAGAGGCCCTCCTAGAGGTG GTCCCGGCGGAGAACTCCCCAGGGGTCTGCCGGCCTGGGCCCTAG ACGCTCCCAGCAAGGTCTAGAAAGGAGCCAAGGAAGATCATCGCC ACCGTGCTGATGACCGAGGACATCAAGCTGAACAAAGCTGAGAAGG

		CCTGGAAGCCTAGCAGCAAGAGAACCGCCGCCGACAAGGACAGAG GCGAGGAGGACGCCGACGGATCCAAGACCCAGGACCTGTTTCAGAAAG AGTGAGAAGCATCGCCAACAAGGCGACCCCTCAGATGGCGCAGCAG CTGATGAAGCAGGTGACGCAGCTCGCCATCGACACCGAGGAGAGAC TGAAGGGCGTGATCGACCTGATCTTTGAGAAGGCTATCTCAGAGCCT AACTTCAGCGTGGCCTACGCCAACATGTGCCGTTGCCTGATGGCATT GAAGGTGCCAACCACCGAGAAGCCTACTGTGACCGTCAATTTCCGTA AACTGCTGCTGAACCGGTGCCAGAAAGAGTTCGAGAAGGATAAGGA CGACGACGAGGTCTTCGAGAAGAAACAGAAAGAAATGGACGAGGC CGCCACCGCAGAGGAAAGGGGGCCGATTAAAGGAGGAGCTGGAGGA GGCCAGAGACATCGCCAGACGGCGTTCTCTGGGCAACATCAAGTTC ATAGGTGAGCTGTTCAAGCTAAAGATGCTCACCAGAGGCCATAATGC ACGACTGCGTGGTGAAGCTACTGAAGAACCACGACGAAGAAAGCCT GGAGTGCCTGTGCAGACTGCTGACCACCATCGGCAAGGACCTGGAC TTCGAGAAGGCAAAGCCTCGAATGGACCAGTACTTCAACCAGATGG AGAAGATTATCAAGGAGAAGAAGACCAGCAGCAGAATCAGATTTCAT GCTGCAGGACGTACTGGACCTGCGCGGAAGCAACTGGGTGCCAAGG AGAGGGGACCAAGGACCAAGACCATCGACGACCATCCACAAGGA GCGGAGATGGAGGAGCACAGAGAGCACATAAAGGTGCAGCAGCTT ATGGCCAAGGGCAGCGACAAGCGAAGAGGCGGCCCGCCGGACCTC CTATCAGCAGAGGCCTTCTCTGGTAGACGACGGCGGCTGGAACAC CGTGCCTATCTCTAAGGGCTCCAGACCTATCGACACTTCCCGTCTTA CCAAGATCACCAGCCAGGATCTATTGACAGCAACAACCAGCTGTT CGCCCCAGGAGGAAGACTTAGCTGGGGCAAGGGAAGTTCCGGCGGA TCCGGCGCCAAGCCTTCCGACGCCCGCCAGCGAGGCTGCCAGACCTG CCACCAGCACCTTGAACCGCTTTTCCGCTCTGCAGCAAGCTGTG
95	LUC-TPE-MBPLACZ (aa)	MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDA HIEVDITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLG ALFIGVAVAPANDIYNERELLSMISQPTVVFVSKKGLQKILNVQKKL PIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALI MNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILSVVPFHH GFGMFTTLGYLICGFRVVLMYRFEELFLRSLQDYKIQSALLVPTLFSFF AKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGRQGYGLTE TTSAILITPEGDDKPGAVGKVVPPFEAKVVDLDTGKTLGVNQRGELCVR GPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLI KYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHG KTMTEKEIVDYVASQVTTAKKLRRGVVVFVDEVPKGLTGKLDARKIREI LIKAKKGGKIAVSGEGRGSLLTCGDVEENPGPGSGATNFSLLKQAGD VEENPGPGSGQCTNYALLKLADVESNPGPASNFTQFVLVDNGGTGDV TVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKG AWRSYLNMELTIFATNSDCELVKAMQGLLKDGNIPIPSAIAANSIYG GGGSSFTLTNKNVIFVAGLGGIGLDTSKELLKRDVPVLRQRRDWENPGVT QLNRLAAHPPFASWRNSEEARTDRPSQQLRSLNGEWRFAWFPAPAEAVP ESWLECDLPEADTVVVP SNWQMHGYDAPIYTNVTYPITVNPFFVPTENP TGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRL PSEFDLSAFLRAGENRLAVMVLRWSDGSYLEDDQDMWRMSGIFRDVSL HKPTTQISDFHVATR FNDDFSRAVLEAEVQMC GELRDYLRVTVSLWQG ETQVASGTAPFGGEIIDERGGYADRVTLRLNVENPKLWSAEIPNLYRAV VELHTADGTLIEAEACDVGFREVRIENGLLLLNGKPLLRGVNRHEHP LHGQVMDEQTMVQDILLMKQNNFN AVRC SHYPNHLWYTLCDRYGL YVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHP SVII WSLGNESGHGANHDALYRWIKSVDP SRPVQYEGGGADTTATDIICPMY ARVDEDQPFPAVPKWSIKKWLSPGETRPLILCEYAHAMGNSLGGFAK YWQAFRQYPRLQGGFVWDWVDQSLIKYDENG NPWSAYGGDFGDT PN DRQFCMNGLVFADRTPHPALTEAKHQQQFFQFRLSGQTIEVTSEYLF RH

		SDNELLHWMVALDYGKPLASGEVPLDVAPQGKQLIELPELPQESAGQL WLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPHLTT SEMDFCIELGNKRWQFNRSQGLSQMWIGDKKQLLTPLRDQFTRAPLD NDIGVSEATRDPNAWVERWKAAGHYQAEAALLQCTADTLADAVLITT AHAWQHQGKTLFISRKTYRIDGSGQMAITVDVEVASDTPHPARIGLNC QLAQVAERVNWLGLGPQENYPDRLTAACFDRWDLPLSDMYTPYVFPS ENGLRCGTRELNYGPHQWRGDFQFNISRSQQQLMETSHRHLLHAEEG TWLNIDGFHMGIGGDDSWSPSVSAELQLSAGRYHYQLVWCQK
96	LUC-TPE- MBPLACZ (nt)	ATGGAGGACGCCAAGAACATCAAGAAGGGGCCCGCTCCCTTCTA CCCGCTGGAGGACGGCACCCTGGCGAGCAGCTGCACAAGGCCATG AAGAGGTACGCACTGGTGCCCCGGCACAATCGCCTTCACCGACGCCC ACATCGAGGTGGACATCACCTACGCCGAGTACTTCGAGATGAGCGT GCGGCTGGCCGAAGCCATGAAGCGGTACGGCCTGAACACCAACCAC CGCATCGTGGTGTGCTCCGAGAACTCCCTGCAGTTCTTCATGCCCCG TCTGGGCGCCCTGTTCAATTGGGGTGGCCGTGGCACCAGCCAACGACA TCTACAACGAGAGGGAGCTGCTGAACAGCATGGGCATCAGCCAGCC CACCGTGGTGTTCGTGTCCAAGAAGGGCCTGCAGAAGATCCTGAAC GTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACA GCAAGACCGACTACCAAGGGCTTCCAGTCCATGTACACCTTCGTGACC AGCCACCTGCCTCCCGGCTTCAACGAGTACGACTTCGTGCCCGAGTC CTTCGACAGGGACAAGACCATCGCCCTGATCATGAACCTCAGCGGA AGCACCGGCCTGCCAAGGGTGTGGCTCTACCCACCGGACCGCCTG TGTACGGTTCAGCCACGCCCCTGACCCCATCTTCGGGAACCAGATCA TCCCCGACACCGCCATCCTGAGCGTGGTGCCCTTCCACCACGGCTTC GGCATGTTACCAACCCTGGGCTACCTGATCTGCGGCTTCCGGGTGGT GCTGATGTACCGGTTTCAGAGGAGGAGCTGTTCTGCGGAGCCTGCAG GACTACAAGATCCAGTCCGCCCTGCTGGTGCCACCCCTGTTCTCCTT CTTCGCCAAGTCCACCCTGATCGACAAGTACGACCTGAGCAACCTGC ACGAGATTGCCTCTGGCGGCGCACCTCTGTCCAAGGAGGTGCGGGA GGCCGTAGCCAAGCGGTTCCACCTGCCCGGCATCCGGCAAGGCTAC GGCCTGACCGAGACCACCAGCGCCATCCTGATCACCCCTGAGGGCG ACGACAAGCCCGGCGCAGTGGGGAAGGTGGTGCCATTCTTCGAGGC CAAGGTGGTGGACCTGGACACCGGCAAGACCCTGGGCGTGAACCA CGGGGCGAGCTGTGCGTGCGGGGACCCATGATCATGTCCGGCTACG TGAACAACCCCGAGGCCACCAACGCCCTGATTGACAAGGACGGGTG GCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTC TTCATCGTGGACCGCCTGAAGTCCCTGATCAAGTACAAGGGGTACCA GGTGGCTCCCGCCGAGCTGGAGTCCATCCTGCTGCAGACCCCAACA TCTTTGACGCGGCGTTCGCCGACTGCCTGACGACGAGCTGGGGAA CTCCCCGAGCTGTGGTGGTGTGGAGCACGGCAAGACCATGACCG AGAAGGAGATCGTGGACTACGTGCCTCCCAGGTGACCACCGCCAA GAAGTTACGCGGCGGCGTCTGTTCTGAGACGAGGTGCCAAGGGC CTGACCGGCAAGCTGGACGCCCCGCAAGATCCGGGAGATCCTGATCA AGGCCAAGAAGGGCGGGAAGATCGCCGTGGGCTCTGGCGAAGGCCG GGGTAGCCTGCTGACCTGCGGCGACGTGGAGGAGAACCCTGGACCT GGATCCGGCGCCACCAACTTCAGCCTGCTCAAGCAGGCCGAGACG TGGAAGAGAATCCAGGCCAGGCAGCGGTCAGTGCACCAACTACGC CCTGCTGAAGCTGGCCGGCGACGTAGAGAGCAACCCCGGCCCGCC AGCAACTTCACCCAGTTCTGTGCTGGTGGACAACGGCGGTACCGGAG ACGTGACCGTGGCCCCCTTCTAACTTCGCCAACGGCATCGCCGAGTGG ATCAGCAGCAACAGCAGAAGCCAGGCCTACAAGGTGACCTGCAGCG TGAGACAGAGCAGCGCCAGAACAGAAAGTACACCATCAAGGTGGA GGTGCCTAAGGGCGCCTGGAGAAGCTACCTGAACATGGAGCTGACC ATCCCTATCTTCGCCACCAACAGCGACTGCGAGCTGATCGTGAAGGC CATGCAGGGCCTGCTGAAGGACGGCAACCCTATCCCTAGCGCCATC

		GCCGCCAATTCAGGCATCTACGGAGGCGGTGGAAGCTCGTTACGCT AACGAACAAGAACGTCATCTTCGTAGCGGGACTTGGCGGTATCGGC CTAGACACGTCGAAGGAACCTAAAGCGTGACCCGGTAGTCCTCC AACGTCGCGATTGGGAGAACCCGGGCGTAACGCAACTAAACCGTCT TGCGGCGCACCCGCCGTTTGCCTCGTGGCGTAACCTCGGAGGAGGCG CGAACGGATCGTCCGTCGCAACAACCTACGTTTCGCTCAACGGGGAGT GGCGCTTCGCGTGGTTCGCCGGCGCCGAGGCGGTACCGGAGTCGTG GCTCGAGTGCGATCTACCGGAGGCGGACACGGTCGTCTGACCGTCG AACTGGCAAATGCACGGTTACGACGCGCCGATATACACGAACGTCA CGTACCCGATAACGGTAAACCCGCCGTTTCGTCCCGACGGAGAACCC GACGGGGTGCTACTCGCTAACGTTCAACGTTGACGAGTCGTGGTTGC AAGAGGGTCAAACGCGTATCATATTCGACGGGTGTAACTCGGCGTTT CACCTGTGGTGCAACGGGCGCTGGGTAGGGTACGGCCAAGACTCGC GTCTACCGTCGGAGTTCGACCTATCGGCGTTTCTACGAGCGGGTGAG AACCGGCTAGCGGTCTGTTGCTACGTTGGTTCGACGGTTCGTACCT CGAGGACCAAGACATGTGGCGAATGTGCGGTATCTTCCGCGACGTA TCGCTCCTACACAAGCCGACGACGCAAATCTCGGACTTCCACGTTCG GACGCGTTTCAACGACGATTTCTCGCGGCGAGTCTAGAGGCGGAG GTCAAATGTGCGGGGAGCTACGTGACTACCTCCGTGTACGGGTATC GCTCTGGCAAGGTGAGACGCAAGTAGCGTCGGGTACGGCGCCGTTT GGCGGTGAGATCATCGACGAGCGTGGTGGGTACGCGGACCGTGTA CGCTACGTCTAAACGTCGAGAACCCGAAGCTCTGGTTCGGCGGAGAT CCCGAACCTATACCGTGCGGTCTGTCGAGCTACATACGGCGGACGGG ACGCTAATAGAGGCGGAAGCGTGCGACGTCGGGTTTCGAGAGGTTT GTATAGAGAACGGGCTGCTACTTCTAAACGGGAAGCCGTTGCTCATA CGTGGTGTCAACCGTCACGAGCACCACCCGCTACACGGTCAAGTAAT GGACGAGCAAACGATGGTACAAGACATCCTACTAATGAAGCAGAAC AACTTCAACGCGGTACGCTGTTTCGCATTACCCGAACCATCCGTTGTG GTACACGCTTTGCGACCGATACGGTCTATACGTCGTAGACGAGGCGA ACATAGAGACGACGCGGATGGTACCGATGAATCGCCTAACGGACGA CCCGCGTTGGCTACCGGCGATGTGCGGAGCGAGTCACGCGTATGGTCC AACGGGACCGTAACCAACCGTCGGTAATAATCTGGTTCGCTAGGCAA CGAATCGGGGCACGGGGCGAACCACGACGCGCTATACCGTTGGATC AAGTCGGTAGACCCGTCGCGTCCGGTACAATACGAAGGTGGCGGTG CGGACACGACGGCGACGGACATCATCTGCCCGATGTACGCGCGCGT CGACGAAGACCAACCGTTCCCGGGCGGTACCGAAGTGGTCGATCAAG AAGTGGCTCTCGTTGCCGGGTGAAACGCGTCCGTTGATACTTTGCGA GTACGCGCACGCGATGGGCAACTCGTTGGGTGGGTTCGCGAAGTAC TGGCAGGCGTTCCGTCAATACCCGCGTCTACAGGGTGGGTTCGTCTG GGACTGGGTAGACCAATCGCTAATCAAGTACGACGAGAACGGCAAC CCGTGGTTCGGCGTACGGTGGGGACTTCGGGGACACGCCGAACGACC GCCAATTCTGTATGAACGGCCTAGTCTTCGCGGACCGAACGCCGCAC CCGGCGTTGACGAGGCGAAGCATCAACAACAATTCTTCCAATTCCG TCTATCGGGGCAAACGATCGAGGTAACGTCGGAGTACTTGTTCGGC ACTCGGACAACGAGCTACTACACTGGATGGTAGCACTAGACGGCAA GCCGCTAGCGTCGGGAGAAGTCCCTTTGGACGTCGCGCCGCAAGGT AAGCAACTAATCGAGCTACCGGAGCTACCGCAACCGGAGTCGGCGG GTCAACTGTGGTTGACGGTCCGTGTCGTTCAACCGAACGCGACGGCG TGGTCGGAGGCGGGTACATCTCGGCGTGGCAGCAGTGGCGTCTAG CGGAGAACCTCTCGGTACGCTACCGGCGGCGTCGCACGCGATACC GCATCTAACGACGTCGGAGATGGACTTCTGCATCGAGTTGGGGAAC AAGAGGTGGCAGTTCAACCGTCAATCGGGATTCTATCGCAAATGTG GATAGGTGACAAGAAGCAACTACTAACGCCGCTACGTGATCAGTTC ACGCGTGCTCCGCTAGACAACGACATAGGTGTTTCGGAGGCGACGC GTATAGACCCGAACGCGTGGGTGGAGCGGTGGAAGGCGCGGGGCA

		CTACCAAGCGGAGGCGGCGCTACTACAGTGCACGGCGGACACGCTA GCGGACGCGGTATTGATCACGACGGCGCACGCGTGGCAACACCAGG GGAAGACGCTATTTCATCTCGCGTAAGACGTACCGTATCGACGGTTTC GGCCAAATGGCGATCACGGTCGACGTAGAGGTAGCGTCGGACACGC CGCATCCGGCGCGCATCGGTCTAAACTGCCAACTAGCGCAAGTAGC GGAGCGTGTAAACTGGCTAGGGCTAGGGCCGCAAGAGAACTATCCG GACCGCCTAACGGCGGCGTGTTCGACCGTTGGGACCTACCGCTTTC GGACATGTATACCCCGTACGTCTTCCCGTCGGAGAACGGGTTGAGGT GCGGGACGCGCGAGCTAAACTACGGGCCGACACAGTGGCGAGGGGA CTTCCAATTCAACATATCGCGTTACTCGCAACAACAATAATGGAGA CGTCGCACCGTCACCTACTACACGCGGAGGAGGGGACGTGGCTAAA CATCGACGGGTTCCACATGGGCATAGGTGGGGACGACTCGTGGTCG CCGTCGGTCTCGGCGGAGCTCCAACCTCTCGGCGGGTCTGTACCATTA CCAAC TAGTTTGGTGCCAGAAG
97	LacZ (aa)	MSFTLTNKNVIFVAGLGGIGLDTSKELLKRDPVVLQRRDWENPGVT QLNRLAAHPPFASWRNSEEARTDRPSQQLRSLNGEWRFAWFAPEAVP ESWLECDLPEADTVVVP SNWQM HGYDAPIYTNVTYPITVNPFPVPTENP TGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRL PSEFDLSAFLRAGENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDLVSL HKPTTQISDFHVATRFNDDFSRAVLEAEVQMCGLRDYLRVTVSLWQG ETQVASGTAPFGGEIIDERGGYADRVTLRLNVENPKLWSAEIPNLYRAV VELHTADGTLIEAEACDVGFREVRIENGLLLLNGKPLLRGVNRHEHP LHGQVMDEQTMVQDILLMKQNNFNNAVRC SHYPNHLWYTLCDRYGL YVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHP SVII WSLGNESGHGANHDALYRWIKSVDPSPRVQYEGGGADTTATDIICPMY ARVDEDQFPFAVPKWSIKKWL SLPGETRPLILCEYAHAMGNSLGGFAK YWQAFRQYPRLQGGFVWDWVDQSLIKYDENGPNWSAYGGDFGDT PN DRQFCMNGLVFADRTPHPALTEAKHQQQFFQFRLSGQTIEVTSEYLF RH SDNELLHWMVALDGKPLASGEVPLDVAPQKGQLIELPELPQESAGQL WLTVRVVPNATAWSEAGHISAWQQWRLAENLSVTLPAAASHAIPHLTT SEMDFCIELGNKRWQFNRSQGLSQMWIGDKKQLLTPLRDQFTRAPLD NDIGVSEATRDPNAWVERWKAAGHYQAEAALLQCTADTLADAVLITT AHAWQHQQKTLFISRKTYRIDGSGQMAITVDVEVASDTPH PARIGLNC QLAQVAERVNWLGLGPQENYPDRLTAACFDRWDLPLSDMYTPYVFPS ENGLRCGTRELNYGPHQWRGDFQFNISRYSQQLMETSHRHLLHAEEG TWLNIDGFHMGIGGDDSWSPSVSAELQLSAGRYHYQLVWCQK
98	LacZ (nt)	ATGAGCTTCACCTGACCAACAAGAACGTGATCTTCGTGGCCGG CCTGGGCGGCATCGGCCTGGACACCAGCAAGGAGCTGCTGAAGCGG GACCCCGTGGTGTGCAGCGCCGGGACTGGGAGAACCCCGGCGTGA CCCAGCTGAACAGGCTGGCCGCCCATCCTCCGTTCCGCTCCTGGCGG AACTCCGAGGAGGCCCCGACCGACCGGCCCTCCAGCAGCTGCGGT CTCTGAACGGCGAGTGGAGGTTTCGCTGGTTCCCCGCCCCAGAGGCC GTGCCTGAGAGCTGGCTGGAGTGCACGCTACGACGACACCCG TGGTGGTGCCTAGCAACTGGCAGATGCACGGCTACGACGCACCCAT CTACACCAACGTGACCTACCCTATCACCGTGAATCCGCCTTTCGTGC CCACCGAGAACCCACCGGCTGCTACAGCCTGACCTTCAACGTGGAC GAGAGTTGGCTGCAGGAGGGCCAGACCCGCATCATCTTCGACGGCG TGAACAGCGCCTTCCACCTGTGGTGCAACGGCCGGTGGGTGGGCTAC GGCCAGGACTCCCGGCTGCCCTCCGAGTTTGACCTGTCCGCCTTCT GCGCGCCGGCGAGAACCGGCTGGCCGTGATGGTGTGAGGTGGAGC GACGGCAGCTACCTGGAGGACCAGGACATGTGGCGCATGTCCGGAA TCTTCCGGGACGTGAGCCTGCTGCACAAGCCCACCACCCAGATCAGC GACTTCCACGTGGCCACCCGGTTCAACGACGACTTCAGCCGGGCGGT GCTGGAGGCCGAGGTGCAGATGTGCGGCGAGCTGCGGGACTACCTG CGCGTGACCGTGAGCCTGTGGCAGGGCGAGACCCAGGTGGCTTCCG

		GCACCGCCCCATTTCGGCGGCGAGATCATCGACGAGCGGGGCGGCTA CGCCGACCGGGTGACCCTGCGGCTGAACGTGGAGAACCCCAAGCTG TGGAGCGCCGAGATCCCCAACCTGTACCGGGCCGTGGTGGAGCTGC ACACCGCCGACGGCACCCCTGATCGAGGCCGAGGCCCTGCGACGTGGG CTTCCGGGAGGTGAGGATCGAGAACGGGCTGCTGCTGCTGAACGGG AAGCCACTGCTGATCAGGGGCGTGAACCGCCACGAGCACCATCCCC TGCACGGCCAGGTGATGGACGAGCAGACCATGGTGCAGGACATCCT GCTGATGAAGCAGAACAACTTCAACGCCGTGAGGTGCTCCCACTAC CCCAACCATCCCCGTGGTACACCCTGTGCGACCCGGTACGGCCTGTA CGTGGTGGACGAGGCCAACATCGAGACCCACGGGATGGTCCCAATG AACCGGCTGACTGACGACCCTCGCTGGCTGCCCGCCATGAGCGAGC GGGTGACCAGGATGGTGCAGCGCGACCGGAACCAACCCAGCGTGAT CATCTGGAGCCTGGGGAACGAGTCCGGCCACGGCGCCAACCACGAC GCCCTGTACCGCTGGATCAAGAGCGTGGACCCCTCCCGCCCCGTGCA GTACGAGGGCGGCGGCGCCGACACCACCGCCACCGACATCATCTGC CCCATGTACGCCCCGGGTGGACGAGGACCAGCCCTTCCCTGCCGTGCC CAAGTGGAGCATCAAGAAGTGGCTGAGCTGCCCCGGCGAGACACAGA CCGCTGATCCTGTGCGAGTACGCCACGCCATGGGCAACTCCCTGGG CGGCTTCGCCAAGTACTGGCAGGCCTTCCGGCAGTACCCGCGGCTGC AGGGCGGCTTCGTGTGGGACTGGGTGGACCAAGTCCCTGATCAAGTA CGACGAGAACGGCAACCCCTGGAGCGCTACGGCGGGGACTTCGGG GACACACCCAACGACCGCCAGTTCTGCATGAACGGCCTGGTGTTCGC CGACCGCACCCCGCACCCCTGCCCTGACCGAGGCCAAGCACCAGCAG CAGTTCTTCCAGTTCCGGCTGAGCGGGCAGACCATCGAGGTGACCAG CGAGTACCTGTTCCGGCACAGCGACAACGAGCTGCTGCACTGGATG GTGGCCCTGGACGGCAAGCCCCTGGCCAGCGGCGAGGTGCCTCTGG ACGTGGCTCCCCAGGGCAAGCAGCTGATCGAGCTGCCCGAGCTGCC CCAGCCCAGAGCGCCGCGCCAGCTGTGGCTGACCGTGAGGGTGGTG CAGCCCAACGCCACCGCCTGGAGCGAGGCCGCGCCACATCAGCGCCT GGCAACAGTGGCGGCTGGCCGAGAACCTGAGCGTGACCCTGCCCGC CGCCAGCCACGCCATCCCTCACCTGACCACCTCCGAGATGGACTTCT GCATCGAGCTGGGCAACAAGCGGTGGCAGTTCAACGGGCAGTCCGG CTTCTGAGCCAGATGTGGATCGGCGACAAGAAGCAGCTGCTGACC CCACTGCGGGACCAGTTCACCAGGGCCCCACTGGACAACGACATCG GCGTGAGCGAGGCCACCAGAATCGACCCCAACGCCTGGGTGGAGCG GTGGAAGGCCGCGGCCACTACCAGGCCGAGGCCGCCCTGCTGCAG TGCACCGCCGACACCCTGGCCGACGCCGTGCTGATCACCACCGCCCA CGCCTGGCAGCACCAGGGGAAGACCCTGTTTCATCAGCCGGAAGACC TACCGGATCGACGGGTCCGGCCAGATGGCCATCACCCTGGACGTGG AGGTGGCCTCCGACACCCCTACCCCGCCCGGATCGGCCTGAACTGT CAGCTGGCCCAGGTGGCCGAGCGGGTGAACCTGGCTGGGGCTGGGCC CACAGGAGAACTACCCCGACCGGCTGACCGCCGCTGCTTCGACCCG GTGGGACCTGCCCTGAGCGACATGTACACCCCATACGTGTTCCCT CTGAGAACGGCCTGCGGTGCGGCACCCGGGAGCTGAACTACGGGCC CCACCAAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGGTACAGCC AGCAGCAGCTGATGGAGACCAGCCACCGGCACCTGCTGCACGCCGA GGAGGGCACCTGGCTGAACATCGACGGCTTCCACATGGGCATCGGC GGCGACGACAGCTGGTCCCCATCCGTGAGCGCCGAGCTGCAGCTGA GCGCCGGCCGGTACCCTACCAGCTGGTGTGGTGCCAGAAG
153	PolyA tail_120 (nt)	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

* (GSG) is an optional element

Equivalents and Scope

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments in accordance with the disclosure described herein. The scope of the present disclosure is not intended
5 to be limited to the Description below, but rather is as set forth in the appended claims.

In the claims, articles such as “a,” “an,” and “the” may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include “or” between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in,
10 or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The disclosure includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The disclosure includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a
15 given product or process.

It is also noted that the term “comprising” is intended to be open and permits but does not require the inclusion of additional elements or steps. When the term “comprising” is used herein, the term “consisting of” is thus also encompassed and disclosed.

Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different
20 embodiments of the disclosure, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.
25

All cited sources, for example, references, publications, databases, database entries, and art cited herein, are incorporated into this application by reference, even if not expressly stated in the citation. In case of conflicting statements of a cited source and the instant application, the statement in the instant application shall control.

30

EXAMPLES

The disclosure will be more fully understood by reference to the following examples. They should not, however, be construed as limiting the scope of the disclosure. It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

10 **Example 1: Production of LNP compositions**

A. Production of nanoparticle compositions

In order to investigate safe and efficacious nanoparticle compositions for use in the delivery of therapeutic and/or prophylactics to cells, a range of formulations are prepared and tested. Specifically, the particular elements and ratios thereof in the lipid component of nanoparticle compositions are optimized.

Nanoparticles can be made with mixing processes such as microfluidics and T-junction mixing of two fluid streams, one of which contains the therapeutic and/or prophylactic and the other has the lipid components.

Lipid compositions are prepared by combining a lipid according to Formulae (I), (IA), (IB), (II), (IIa), (IIb), (IIc), (IId), (IIf), (IIg), (III), (IIIa1), (IIIa2), (IIIa3), (IIIa4), (IIIa5), (IIIa6), (IIIa7), or (IIIa8) or a non-cationic helper lipid (such as DOPE, or DSPC obtainable from Avanti Polar Lipids, Alabaster, AL), a PEG lipid (such as 1,2 dimyristoyl sn glycerol methoxypolyethylene glycol, also known as PEG-DMG, obtainable from Avanti Polar Lipids, Alabaster, AL), and a phytosterol (optionally including a structural lipid such as cholesterol) at concentrations of about, e.g., 50 mM in a solvent, e.g., ethanol. Solutions should be refrigerated for storage at, for example, -20° C. Lipids are combined to yield desired molar ratios (see, for example, Table 6 below) and diluted with water and ethanol to a final lipid concentration of e.g., between about 5.5 mM and about 25 mM. Phytosterol* in Table 6 refers to phytosterol or

optionally a combination of phytosterol and structural lipid such as beta-phytosterol and cholesterol.

Table 6: Exemplary formulations of LNP compositions

Composition (mol %)	Components
40:20:38.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
45:15:38.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
50:10:38.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
55:5:38.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
60:5:33.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
45:20:33.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
50:20:28.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
55:20:23.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
60:20:18.5:1.5	Ionizable lipidd:Phospholipid:Phytosterol*:PEG-DMG
40:15:43.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
50:15:33.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
55:15:28.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG

Composition (mol %)	Components
60:15:23.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
40:10:48.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
45:10:43.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
55:10:33.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
60:10:28.5:1.5	Ionizable lipidd:Phospholipid:Phytosterol*:PEG-DMG
40:5:53.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
45:5:48.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
50:5:43.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
40:20:40:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
45:20:35:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
50:20:30:0	Ionizable lipidd:Phospholipid:Phytosterol*:PEG-DMG
55:20:25:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
60:20:20:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
40:15:45:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG

Composition (mol %)	Components
45:15:40:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
50:15:35:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
55:15:30:0	Ionizable lipidPhospholipid:Phytosterol*:PEG-DMG
60:15:25:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
40:10:50:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
45:10:45:0	Ionizable lipidPhospholipid:Phytosterol*:PEG-DMG
50:0:48.5:1.5	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
50:10:40:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
55:10:35:0	Ionizable lipid:Phospholipid:Phytosterol*:PEG-DMG
60:10:30:0	Ionizable lipidPhospholipid:Phytosterol*:PEG-DMG

Nanoparticle compositions including a therapeutic and/or prophylactic and a lipid component are prepared by combining the lipid solution with a solution including the therapeutic and/or prophylactic at lipid component to therapeutic and/or prophylactic wt:wt ratios between about 5:1 and about 50:1. The lipid solution is rapidly injected using a NanoAssemblr microfluidic based system at flow rates between about 10 ml/min and about 18 ml/min into the therapeutic and/or prophylactic solution to produce a suspension with a water to ethanol ratio between about 1:1 and about 4:1.

For nanoparticle compositions including an RNA, solutions of the RNA at concentrations of 0.1 mg/ml in deionized water are diluted in a buffer, e.g., 50 mM sodium citrate buffer at a pH between 3 and 4 to form a stock solution.

Nanoparticle compositions can be processed by dialysis to remove ethanol and achieve buffer exchange. Formulations are dialyzed twice against phosphate buffered saline (PBS), pH 7.4, at volumes 200 times that of the primary product using Slide-A-Lyzer cassettes (Thermo Fisher Scientific Inc., Rockford, IL) with a molecular weight cutoff of 10 kDa. The first dialysis is carried out at room temperature for 3 hours. The formulations are then dialyzed overnight at 4° C. The resulting nanoparticle suspension is filtered through 0.2 µm sterile filters (Sarstedt, Nümbrecht, Germany) into glass vials and sealed with crimp closures. Nanoparticle composition solutions of 0.01 mg/ml to 0.10 mg/ml are generally obtained.

The method described above induces nano-precipitation and particle formation. Alternative processes including, but not limited to, T-junction and direct injection, may be used to achieve the same nano-precipitation.

B. Characterization of nanoparticle compositions

A Zetasizer Nano ZS (Malvern Instruments Ltd, Malvern, Worcestershire, UK) can be used to determine the particle size, the polydispersity index (PDI) and the zeta potential of the nanoparticle compositions in 1×PBS in determining particle size and 15 mM PBS in determining zeta potential.

Ultraviolet-visible spectroscopy can be used to determine the concentration of a therapeutic and/or prophylactic (e.g., RNA) in nanoparticle compositions. 100 µL of the diluted formulation in 1×PBS is added to 900 µL of a 4:1 (v/v) mixture of methanol and chloroform. After mixing, the absorbance spectrum of the solution is recorded, for example, between 230 nm and 330 nm on a DU 800 spectrophotometer (Beckman Coulter, Beckman Coulter, Inc., Brea, CA). The concentration of therapeutic and/or prophylactic in the nanoparticle composition can be calculated based on the extinction coefficient of the therapeutic and/or prophylactic used in the composition and on the difference between the absorbance at a wavelength of, for example, 260 nm and the baseline value at a wavelength of, for example, 330 nm.

For nanoparticle compositions including an RNA, a QUANT-IT™ RIBOGREEN® RNA assay (Invitrogen Corporation Carlsbad, CA) can be used to evaluate the encapsulation of an RNA by the nanoparticle composition. The samples are diluted to a concentration of approximately 5 µg/mL in a TE buffer solution (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). 50 µL of the diluted samples are transferred to a polystyrene 96 well plate and either 50 µL of TE buffer or 50 µL of a 2% Triton X-100 solution is added to the wells. The plate is incubated at a temperature of 37° C for 15 minutes. The RIBOGREEN® reagent is diluted 1:100 in TE buffer, and 100 µL of this solution is added to each well. The fluorescence intensity can be measured using a fluorescence plate reader (Wallac Victor 1420 Multilabel Counter; Perkin Elmer, Waltham, MA) at an excitation wavelength of, for example, about 480 nm and an emission wavelength of, for example, about 520 nm. The fluorescence values of the reagent blank are subtracted from that of each of the samples and the percentage of free RNA is determined by dividing the fluorescence intensity of the intact sample (without addition of Triton X-100) by the fluorescence value of the disrupted sample (caused by the addition of Triton X-100).

C. In vivo formulation studies

In order to monitor how effectively various nanoparticle compositions deliver therapeutic and/or prophylactics to targeted cells, different nanoparticle compositions including a particular therapeutic and/or prophylactic (for example, a modified or naturally occurring RNA such as an mRNA) are prepared and administered to rodent populations. Mice are intravenously, intramuscularly, subcutaneously, intraarterially, or intratumorally administered a single dose including a nanoparticle composition with a lipid nanoparticle formulation. In some instances, mice may be made to inhale doses. Dose sizes may range from 0.001 mg/kg to 10 mg/kg, where 10 mg/kg describes a dose including 10 mg of a therapeutic and/or prophylactic in a nanoparticle composition for each 1 kg of body mass of the mouse. A control composition including PBS may also be employed.

Upon administration of nanoparticle compositions to mice, dose delivery profiles, dose responses, and toxicity of particular formulations and doses thereof can be

measured by enzyme-linked immunosorbent assays (ELISA), bioluminescent imaging, or other methods. For nanoparticle compositions including mRNA, time courses of protein expression can also be evaluated. Samples collected from the rodents for evaluation may include blood, sera, and tissue (for example, muscle tissue from the site of an intramuscular injection and internal tissue); sample collection may involve sacrifice of the animals.

Nanoparticle compositions including mRNA are useful in the evaluation of the efficacy and usefulness of various formulations for the delivery of therapeutic and/or prophylactics. Higher levels of protein expression induced by administration of a composition including an mRNA will be indicative of higher mRNA translation and/or nanoparticle composition mRNA delivery efficiencies. As the non-RNA components are not thought to affect translational machineries themselves, a higher level of protein expression is likely indicative of a higher efficiency of delivery of the therapeutic and/or prophylactic by a given nanoparticle composition relative to other nanoparticle compositions or the absence thereof.

Example 2: Tethered eIF4G increases potency of target mRNA

This Example describes increased potency of a target mRNA (e.g., increased protein expression and/or duration of protein expression) when co-delivered with an RNA encoding a tethered effector protein, e.g., tethered eIF4G.

The system used in this Example is depicted in **FIG. 1**. The eIF4G mid-to-C-terminal domain can support ribosome complex formation, support or enhance cap-dependent translation when recruited to mRNA 3' UTR via cap-independent translation enhancers (CITEs), and stimulate translation of a capless mRNA when recruited via IRES elements (Kraft JJ, et al. (2013) *NAR*. 41(5):3398-413; Paek KY, et al. (2015) *PNAS* 112(4):1041-6; Pestova TV, et al. (1996) *Mol Cell Biol*. 16(12):6870-8). Thus, this domain (herein named eIF4GΔN) was used as the effector. With no eIF4E and PABP binding abilities, this domain may not bind other mRNAs in the cell, thereby reducing probable off-target effects. As a tether the MS2-MBP system was used since it has been extensively validated over the last few decades. Notably the MS2-MBP

interaction is expected to be very strong ($K_d \sim 1\text{-}10\text{ nM}$) and stable in a cell (Tutucci E, et al. (2018) Nat Methods. 15(1):81-89). Six MS2 loops were inserted in the 3'UTR of the target mRNA. The MS2 loops either replace the 3'UTR completely or are added before or after v1.1 UTR sequence (**FIG. 1**).

5 Initial cell-based experiments relied on using deg-GFP as the reporter in HeLa cells. Target reporter RNA was co-delivered with an mRNA encoding control protein, EPO; or an mRNA encoding tethered control protein, MBP-LacZ; or an mRNA encoding tethered effector protein, MBP-eIF4GΔN. The experiments were done in 10x molar excess of target RNA. As expected, minimal differences were observed in target mRNA
10 expression with 3'v1.1 UTR under different conditions (this RNA has no binding site for any of the encoded proteins). In all 3 instances of target mRNAs that have binding sites for the tethered protein, tethered effector significantly increased protein output from the mRNA (about 3-6x increase in protein output). Maximum benefits were observed when the tether was immediately adjacent to the polyA sequence (**FIGs. 2A-2D**). This boost in
15 expression was reproduced across multiple cell-types (**FIGs. 3A-3C**) ranging from 3-fold (modest) to 80-fold (high).

Finally, at fixed target mRNA amount, the potency boost with tethered effector increased with increasing amounts of effector. This was also evident when the data was fitted to a half-life model that predicted half-life of target mRNA increased with
20 increasing effector concentrations (**FIGs. 4A-4B**). At a 9x molar excess of target, ~ 8-10 fold increase in half-life of the target mRNA was observed in cells delivered the target mRNA with the tethered effector (green bars), as compared to cells delivered the target mRNA with a tethered control (orange bars) (see **FIG. 4B**).

The boost in protein expression with tethered effector was observed for two more
25 reporters, Luc and NPI-Luc. The shapes of the total-intensity curves and model fits both suggest an increased half-life for the target mRNA in the presence of tethered effector protein.

In summary, the data provided in this experiment demonstrates that the expression
30 of a target mRNA and its encoded protein can be significantly increased when said target mRNA is co-delivered with a tethered effector protein.

Example 3: Tethered eIF4G increases mRNA half-life

This Example describes increased half-life of a target mRNA when co-delivered with a tethered effector, e.g., tethered eIF4G.

5 A similar system as used in Example 2 was used in this Example. Briefly, HEK293 cells were electroporated with target reporter RNA and an mRNA encoding control protein (EPO); or an mRNA encoding tethered control protein (MBP-LacZ); or an mRNA encoding tethered effector protein (MBP-eIF4GΔN). The experiments were done in 3x molar excess of target RNA.

10 As shown in **FIGs. 5C-5D**, half-life of the target mRNA was increased in the presence of the tethered effector as compared to the tethered control or non-tethered control. Maximum benefits were observed when the tether was immediately adjacent to the polyA sequence. **FIG. 5D** shows that about 80% of the target mRNA (co-delivered with the tethered effector) is present/can be detected at 24 hours post electroporation
15 compared to control conditions (target mRNA with tethered control or with non-tethered control).

This data shows that the half-life of a target mRNA can be increased when said target mRNA is co-delivered with a tethered effector, e.g., tethered eIF4G.

20 **Example 4: Evaluation of target mRNA translation with tethered effector**

This Example describes the effects of a tethered effector on the translation of target mRNA.

For this Example, the nascent peptide imaging (NPI)-Luc reporter system (used
25 to image mRNA, translating mRNA and protein) was utilized to evaluate target mRNA translation with a tethered effector. The NPI-Luc reporter encodes for a Luc ORF tagged with V5 peptide on the N-terminal end, and a nuclear localization signal. In using this reporter, by assessing spots/ signal positive for both V5 (newly forming protein) and smFISH (that detects LUC RNA), one is able to measure translating
30 mRNAs. This is done by assessing the spots that are positive for both NPI+ (or V5

positive) smFISH+ where NPI reflects the nascent peptide, and the FISH probes are specific for Luc.

All experiments were done with target or control RNA electroporation under different conditions. In HeLa, imaging results were obtained for target or control mRNA with non-tethered, tethered control, and tethered effector RNAs (only images for tethered control and effector RNA groups are shown). Tethered effector exhibited higher translation (colocalized NPI+ smFISH+ spots) for MS2 containing target RNAs, especially at later time points in HeLa. This experiment was repeated in HeLa and Hep3b with NPI-Luc mRNA with MS2 binding sites in the UTR. The quantified data is shown in **FIGs. 6A-6E, 7A-7D, 16 and 17A-17C**.

In Hep3b cells, co-transfection with effector RNA led to i) decreased mRNA loss with time; ii) decreased translating mRNA loss with time; iii) robust translation maintenance in cells over time and iv) translation maintenance per mRNA over time(**FIGs. 6A-6E**). **FIG. 16** shows that the tethered effector decreases loss of translating mRNAs over a time period of 48 hours post transfection. In HeLa cells, translation overall appeared more permissive. A minimal impact was observed on cytosolic mRNAs with time, and translation per mRNA with time at any of the time points under any condition. However, tethered effector reduced the loss of translating mRNAs with time and showed more robust translation in a higher fraction of cells with time (**FIGs. 7A-7D**). The difference in magnitude of impact observed could be cell specific, or could depend on the dynamic range of the assay in the two cell types.

In summary, the data provided in this Example shows that a tethered effector can prevent loss of mRNA with time; prevent a reduction in, or maintain translation on mRNAs with time; and promote translation in cells over time. Therefore, in some embodiments, co-delivery of a tethered effector and target mRNA can be used to increase the translation output of a target mRNA or increase the duration of translation output from an mRNA, thereby increasing the amount of target protein.

Example 5: Identifying the domain of eIF4G required for effector function

This Example describes an analysis of the domains of eIF4G which are required for effector function. The schematic of the different constructs used in this Example is provided in **FIG. 8A**.

Target deg-GFP RNA was co-delivered with an mRNA encoding control protein, EPO; or an mRNA encoding tethered control protein, MBP-LacZ; or an mRNA encoding tethered effector protein, MBP-eIF4GΔN, or mRNAs encoding MBP-fused to eIF4G truncations and mutations as illustrated in **FIG. 8A**. The experiment was done in 10x molar excess of target RNA. As expected, minimal differences were observed in target mRNA expression with 3'v1.1 UTR under different conditions.

As shown in **FIGs. 8B-8E**, the eIF4G dN, eIF4G dN2, eIF4G dN3, eIF4G mid1, eIF4G mid2, eIF4G mid3, eIF4G mid4, eIF4G C1 and eIF4G C2 constructs resulted in increased protein output from the target mRNA with MS2-binding sites in the UTR. In contrast **FIGs. 8B and 8D** show that the eIF4G dN 3A construct (which is unable to bind eIF3 and eIF4A through the mid-domain (Imataka H, Sonenberg N. (1997) *Mol Cell Biol.* 17(12):6940-7)) showed no appreciable benefit over tethered control. **FIG. 9** provides a summary of expected binding activity for the various constructs that were tested.

Taken together, this data shows that ribosome binding of eIF4G via EIF3/EIF4A is important for effector function.

Example 6: Tethered eIF4G rescues protein expression and increases mRNA stability

This Example describes increasing the protein expression (or stability) of an unstable / translationally inactive target mRNA when co-delivered with an RNA encoding a tethered effector.

A similar system as used in Example 2 was used in this Example. HeLa cells were electroporated with 2 RNAs. Each sample has target degGFP encoding RNA co-transfected with a control (EPO) (i.e., 3'UTR + EPO) or effector protein encoding RNA

(i.e., 3'UTR + MBP-eIF4G). The experiments were done in 3x molar excess of target RNA. The 3' UTR used is either 3'UTR with 24 MS2 hairpins (3'UTR_24 MS2) or 3' UTR_v1.1. As shown in **FIG. 10B**, an increased number of MS2 hairpin loops (24 MS2) in the 3' UTR of the target RNA reduced protein expression (and possibly the mRNA half life) compared to the 3' UTR v1.1. However, mRNA expression was restored even for the mRNA with 24 MS2 loop, when it was tethered to an activator protein (eIF4G1(623-1599)) using the MS2-binding protein. In addition, there was a notable extension of the functional half-life of the target RNA in the presence of the activator protein. **FIG. 10C** shows the area under the curve (AUC) for the same data in **FIG. 10B**.

This data shows that the protein expression (and possibly half-life of a target mRNA) can be increased when the target mRNA is co-delivered with an RNA encoding a tethered effector.

15 **Example 7: Tailless target RNAs can be rescued by effectors recruited using RBP-RNA (MBP-MS2) tethers**

FIG. 11 provides a schematic of a system to recruit potential effectors to target RNA with an MS2 binding protein (MBP)-MS2 tether to rescue tailless target RNA. The target mRNA has MS2 loops in the 3'UTR and lacks a polyA tail (A0). The second mRNA encodes MS2-binding protein (MBP) fused to an effector/eIF4G-mid. The MBP-MS2 binding interaction recruits the effector protein to the target RNA. This system can be coupled to a miRNA-dependent switch gate to permit tethering in specific cells, thereby turning ON expression of the target protein. This system can also be coupled to an effector encoding DNA molecule under the control of a tissue-specific promoter to turn on expression in specific cells. Finally this system can be encoded such that the effector protein expression and/or recruitment is under the control of a trigger (receptor-mediated activation, a change in pH, hypoxia, etc) in specific microenvironments and/or specific cell types (the presence of specific microRNA in that cell-type). Thus, in some embodiments, even if the effector is made, it can not be recruited to the target RNA without a specific trigger.

The eIF4G mid-to-C-terminal domain can support ribosome complex formation, support or enhance cap-dependent translation when recruited to mRNA 3' UTR via cap-independent translation enhancers (CITEs), and stimulate translation of a capless mRNA when recruited via IRES elements (Kraft JJ, et al. (2013) *NAR*. 41(5):3398-413; Paek KY, et al. (2015) *PNAS* 112(4):1041-6; Pestova TV, et al. (1996) *Mol Cell Biol*. 16(12):6870-8). Thus, this domain (herein named eIF4GAN) can be used as the effector. The domain eIF4-mid can also be used as the effector. With no eIF4E and PABP binding abilities, this domain may not bind other mRNAs in the cell, thereby reducing probable off-target effects. As a tether, the MS2-MBP system can be used since it has been extensively validated over the last few decades. Notably the MS2-MBP interaction is expected to be very strong ($K_d \sim 1-10$ nM) and stable in a cell (Tutucci E, et al. (2018) *Nat Methods*. 15(1):81-89). Six MS2 loops were inserted in the 3' UTR of the target mRNA.

15 **Example 8: Tethered eIF4G increases mRNA stability of tailless mRNA**

This Example describes increasing translation and stability of a tailless mRNA when co-delivered with a tethered effector.

Hep3b cells or HeLa cells were transfected with 2 RNAs. Each sample has a target degGFP co-transfected with an mRNA encoding a tethered protein (Effector, eIF4G-mid2) or a tethered control (MBP-LacZ). The target mRNA was tailless (A0) (i.e., 3'v1.1_MS2_A0) or has an A100 tail (v1.1_A100). The experiments were done in 10x molar excess of target RNA and protein expression was measured by Incucyte. As shown in **FIGs. 12A** and **12C**, while standard mRNAs with A100 tails show robust expression, no detectable expression is seen for tailless mRNAs with MS2 UTRs co-delivered with tethered control (Test (A0) target RNA + t-ctrl). The translation of tailless targets is restored in the presence of the eIF4G mid domain (Test (A0) target RNA + t-eff). The shape of the curve suggests an increase in RNA half-life. **FIGs. 12B** and **12D** show the area under the curve (AUC) for the same data in **FIGs. 12A** and **12B**.

30

Example 9: Tailless target RNAs with idT tails can be rescued by effectors recruited using RBP-RNA (MBP-MS2) tethers

To illustrate that the system of **FIG. 11** can be used to rescue modified tailless target RNA expression, HEP3B cells were transfected with 3'v1.1_MS2_A0 or
 5 3'v1.1_MS2_A0-idT + Effector/Control deg-GFP constructs. In this experiment, tailless target RNA was modified with an idT tail, *i.e.*, A20 oligo followed by an inverted idT. Control RNA encodes for MBP-LacZ and effector RNA encodes for MBP-eIF4G-mid. Target RNA was in 10x molar excess of effector RNA. As seen in **FIG. 13**, A0 or A0-idT constructs showed no detectable expression when co-transfected
 10 with control RNA. Tethering effector increased expression from A0 RNA. Tethering in the presence of idT rescued expression to a much higher level. These results demonstrate that translation from an A0 RNA can improve in the presence of idT, along with a tether.

Example 10: Capless target RNAs can be rescued by effectors recruited using RBP-RNA (MBP-MS2) tethers

To illustrate that the system of **FIG. 11** can be used to rescue capless target RNA, HeLa cells were transfected with 5' triphosphate (5'ppp) ended NpiLuc-
 _3'v1.1_MS2_A100 target RNAs + Effector/Control NpiLuc constructs. Target RNA
 20 was in 1.5x molar excess of effector RNA. As seen in the **FIG. 14**, while the capless RNA by itself showed no appreciable signal above background with a control tether RNA (MBP-mid2-Mut, encodes for a mutant mid domain), appreciable expression is restored on this RNA when tethered with effector (MBP-mid2)+ Effector/Control deg-GFP construct, and 3'v1.1_MS2_A0 + Control or Effector.

25

Example 11: Capless-Tailless target RNAs can be rescued by effectors recruited using RBP-RNA (MBP-MS2) tethers

To illustrate that the system of **FIG. 11** can be used to rescue capless-tailless target RNA, HeLa cells were transfected with 5'ppp ended NpiLuc_3'v1.1_MS2_A0
 30 target RNAs + Effector/Control NpiLuc constructs. Target RNA was in 1.5x molar

excess of effector RNA. As seen in the **FIG. 15**, while the capless-tailless RNA by itself showed no appreciable signal above background with a control tether RNA (MBP-mid2-Mut, encodes for a mutant mid domain), appreciable expression was restored on this RNA when delivered with the tethered-effector construct (MBP-mid2).

5

Example 12: Tethered effector maintains more translating mRNAs in more cells over time

This Example describes maintenance of more translating mRNAs in more cells over time when the mRNA is co-delivered with a tethered effector, e.g., tethered
 10 eIF4GΔN. Hep3b cells were electroporated with the target RNA (3'v1.1MS2_A100 or 3'v1.1_A100) in combination with non-tethered control (nt-ctrl; LacZ), tethered control (t-ctrl; MBP-LacZ) or t-effector (eIF4GΔN). Tethered effector exhibited a higher percentage of translating mRNA (NPI+ smFISH+ spots) for a longer period of time, as seen in **FIGs. 17A-17C**.

15

Example 13: Tethering increases secreted protein expression and translation

This Example demonstrates that secreted protein expression and translation is increased when the mRNA is co-delivered with a tethered effector, e.g., tethered
 20 eIF4GΔN. v1.1 target RNA constructs with optimized reading frames for the light chain and heavy chain pairs of two secreted antibodies, (Ab1 and Ab2) were co-transfected into Hek293 cells with t-ctrl (MBP-LacZ) or t-eff (MBP-eIF4GΔN). Experiments were done with 5x molar target excess (Ab1) or 1x molar target excess (Ab2). In each of the antibody experiments, the tethered effector is tethered to two separate light and heavy chain encoding RNAs. **FIG. 18** shows the concentration (μg/ml) of Ab1 or Ab2 over
 25 time and indicates that tethering increases secreted protein expression and translation.

Example 14: A single RNA tethering system enhances target expression *in vivo*

This Example demonstrates that a single RNA tethering system can increase the efficiency of target RNA translation *in vivo*. **FIG. 19** provides a schematic of the single
 30 RNA tethering system. The mRNA molecule from 5' to 3' includes a CAP, a 5'UTR,

target ORF, 3 protease cleavage sites in tandem: T2A-P2A-E2A/ TPE (red), another ORF encoding for the RNA binding protein fused to an effector (orange-RBP, -green-Effector), and MS2 loops (orange stripes) in the 3'UTR. The MS2 loops are added after v1.1 UTR sequence. The TPE protease cleavage site leads to ribosome skipping during translation in a cell. The end product of this translation is that the target ORF encoded protein gets a C-terminal tag, and the RNA binding binding-effector fusion starts with a residual Proline (C-terminal amino acid) from TPE. For specifically TPE cleavage site, this arrangement is expected to lead to 1 Target: 0.05 effector protein molecules (Liu et al, 2017 Scientific Reports 7: 2193). In the 1 RNA system, the effector protein binding to the RNA 3'UTR via the RNA-binding-protein interaction confers the benefit of effector protein binding to this one RNA encoding for both target and effector ORFs.

FIG. 20A shows the level of luminescence over the various timepoints indicated for mice that were injected with a t-eff construct (Effector) or a t-ctrl construct (Control). **FIG. 20B** shows the cumulative luminiscence plotted as total Area Under Curve; AUC) corresponding to **FIG. 20B**. **FIG. 20C** shows the luminescence over the time points indicated for mice that were injected with a t-eff construct (Effector) or a t-ctrl construct (Control). Data is shown as mean +/- SEM. The delivery system used here is an LNP.

Example 15: Identifying other useful effector proteins or domains thereof

This Example describes an analysis of the homologues or domains of eIF4G which are required for effector function.

Target deg-GFP RNA was co-delivered with an mRNA encoding control protein, LacZ; or an mRNA encoding tethered control protein, MBP-LacZ; or an mRNA encoding tethered effector proteins as depicted in **FIGs. 21A-21B** and **FIGs. 22A-22B**. The experiment was done in 10x molar excess of target RNA in Hep3b.

As shown in **FIGs. 21A-21B**, both eIF4G1-fl and eIF4G3-fl, and domains thereof that contained the mid domain (binds eIF4A-3) increase protein output from target RNA. As shown in **FIGs. 22A-22B**. PolyA binding protein and polynucleotidyl

transferases such as Gld2, TENT4A and TENT4B all increase protein output from target RNA.

5 Taken together, this data shows that ribosome binding of eIF4G homologues via eIF3/eIF4A is important for effector function. Additionally, PABP and polynucleotidyl transferases also serve as effectors that increase protein expression and duration of expression from target RNAs.

Other Embodiments

10 It is to be understood that while the present disclosure has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the present disclosure, which is defined by the scope of the appended claims. Other aspects, advantages, and alterations are within the scope of the following claims. All references described herein are incorporated by reference in their entireties.

What is claimed is:

1. A lipid nanoparticle (LNP) composition comprising:

(a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or a prophylactic payload, and (2) a binding element; and

(b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and/ or (2) a polypeptide that recognizes the binding element (a tether molecule),

optionally wherein, (a) and (b) each comprise an mRNA.

2. A lipid nanoparticle (LNP) composition comprising:

(a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or a prophylactic payload, and (2) a binding element; and

(b) a second polynucleotide comprising a sequence encoding an effector molecule,

optionally wherein, (a) and (b) each comprise an mRNA.

3. The LNP composition of claim 2, wherein the effector molecule further comprises a polypeptide that recognizes the binding element (a tether molecule).

4. The LNP composition of claim 2, wherein the effector molecule recognizes the binding element.

5. The LNP composition of any one of claims 1-4, wherein the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

6. The LNP composition of claim 5, wherein the first polynucleotide and the second polynucleotide are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A site) or an internal ribosomal entry site.

7. The LNP composition any one of claims 1-4, wherein the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.
8. The LNP composition of claim 7, wherein (a) and (b) are in the same LNP.
9. The LNP composition of claim 7, wherein (a) and (b) are in different LNPs.
10. The LNP composition of any one of claims 1-4, wherein the second polynucleotide is DNA.
11. The LNP composition of claim 10, wherein the sequence encoding the effector molecule is under the control of a tissue-specific promoter.
12. The LNP composition of any one of claims 1-4, wherein expression or recruitment of the effector molecule is under the control of a trigger in a specific microenvironment or specific cell-type.
13. The LNP composition of claim 12, wherein the trigger is microRNA, receptor-mediated activation, and/or a change in pH and/or hypoxia.
14. A system comprising:
 - (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or a prophylactic payload, and (2) a binding element; and/or
 - (b) a second polynucleotide comprising a sequence encoding: (1) an effector molecule, and (2) a polypeptide that recognizes the binding element (a tether molecule), optionally wherein, (a) and (b) each comprise an mRNA.
15. A system comprising:
 - (a) a first polynucleotide comprising: (1) a sequence encoding a therapeutic payload or a prophylactic payload, and (2) a binding element; and/or

(b) a second polynucleotide comprising a sequence encoding an effector molecule,

optionally wherein, (a) and (b) each comprise an mRNA.

16. The system of claim 15, wherein the effector molecule further comprises a polypeptide that recognizes the binding element (a tether molecule).

17. The system of claim 15, wherein the effector molecule recognizes the binding element.

18. The system of any one of claims 14-17, wherein the system comprises (a).

19. The system of any one of claims 14-18, wherein the system comprises (b).

20. The system of any one of claims 14-19, wherein the system comprises (a) and (b).

21. The system of any one of claims 14-20, wherein the first polynucleotide and the second polynucleotide are disposed in the same polynucleotide.

22. The system of claim 21, wherein the first polynucleotide and the second polynucleotide are separated by a protease cleavage site (e.g., a P2A, T2A, E2A, or TPE (P2A-T2A-E2A site) or an internal ribosomal entry site.

23. The system of any one of claims 14-20, wherein the first polynucleotide and the second polynucleotide are disposed in different polynucleotides.

24. The system of any one of claims 14-23, wherein at least one of (a) or (b) is formulated as a lipid nanoparticle (LNP).

25. The system of claim 24, wherein (a) is formulated as an LNP.

26. The system of claim 24, wherein (b) is formulated as an LNP.
27. The system of claim 24, wherein (a) and (b) both are formulated as LNPs, e.g., the same LNP or different LNPs.
28. The system of any one of claims 14-17, wherein the second polynucleotide is DNA.
29. The system of claim 15, wherein the effector molecule is under the control of a tissue-specific promoter.
30. The system of any one of claims 14-17, wherein expression of the effector molecule or recruitment of the effector molecule is under the control of a trigger in a specific microenvironment or specific cell-type.
31. The system of claim 30, wherein the trigger is microRNA, receptor-mediated activation, and/or a change in pH and/or hypoxia.
32. A pharmaceutical composition comprising the system, or LNP composition of any one of the preceding claims.
33. A cell comprising a system, or LNP composition of any one of the preceding claims.
34. The cell of claim 33, which has been contacted with the system, or LNP composition of any of claims 1-31.
35. The cell of claim 33 or 34, which is maintained under conditions sufficient to allow for expression of one or both polynucleotides of the system, or LNP composition.
36. A method of increasing expression of a therapeutic payload or a prophylactic payload in a cell, comprising administering to the cell a system, or LNP composition of any one of claims 1-27.

37. A method of increasing expression of a therapeutic payload or a prophylactic payload, in a subject, comprising administering to the subject an effective amount of a system or LNP composition of any one of claims 1-27.

38. A method of delivering a system, or LNP composition of any one of claims 1-27, to a cell.

39. The method of claim 38, comprising contacting the cell *in vitro*, *in vivo*, or *ex vivo* with the system or LNP composition.

40. A method of delivering a system or LNP composition of any one of claims 1-27, to a subject having a disease or disorder, e.g., as described herein.

41. A method of modulating an immune response in a subject, comprising administering to the subject in need thereof an effective amount of a system, or LNP composition of any one of claims 1-27.

42. A method of treating, preventing, or preventing a symptom of, a disease or disorder comprising administering to a subject in need thereof an effective amount of a system, or LNP composition of any one of claims 1-27.

43. The method of any one of claims 36-42, wherein the first polynucleotide or the second polynucleotide of the system is formulated as an LNP.

44. The method of claim 43, wherein both the first and the second polynucleotides of the system are formulated as LNPs, e.g., the same or different LNPs.

45. The method of any one of claims 36-44, wherein the LNP comprising (a) and the LNP comprising (b) are administered simultaneously, e.g., substantially simultaneously.

46. The method of any one of claims 36-44, wherein the LNP comprising (a) and the LNP comprising (b) are administered sequentially.

47. The system, LNP composition, cell or method of any one of the preceding claims, wherein the tether molecule of the second polynucleotide comprises an RNA binding protein or a fragment thereof, which binds to, e.g., recognizes, the binding element of the first polynucleotide.

48. The system, LNP composition, cell or method of any one of the preceding claims, wherein the binding element of the first polynucleotide is situated upstream (5') or downstream (3') or in the open reading frame of the sequence encoding the therapeutic payload or a prophylactic payload.

49. The system, LNP composition, cell or method of any one of the preceding claims, wherein the binding element of the first polynucleotide is situated upstream (5') or downstream (3') of a 5' UTR of the first polynucleotide.

50. The system, LNP composition, cell or method of any one of the preceding claims, wherein the binding element of the first polynucleotide is situated upstream (5') or downstream (3') of a 3' UTR of the first polynucleotide.

51. The system, LNP composition, cell or method of any one of the preceding claims, wherein the binding element of the first polynucleotide is situated downstream of a 3' UTR of the first polynucleotide.

52. The system, LNP composition, cell or method of any one of the preceding claims, wherein the binding element of the first polynucleotide is bound by the tether molecule of the second polynucleotide, e.g., a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, 15.5kd, LARP7 or a variant or fragment thereof.

53. The system, LNP composition, cell or method of any one of the preceding claims, wherein the binding element is chosen from a binding element provided in Table 1, e.g., MS2 (e.g., wildtype MS2, or a variant or fragment thereof).

54. The system, LNP composition, cell or method of any one of the preceding claims, wherein the binding element comprises a sequence comprising 19 nucleotides, e.g., a MS2 binding element nucleotide sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

55. The system, LNP composition, cell or method of any one of the preceding claims, wherein the effector molecule is chosen from: a translation factor, a splicing factor, an RNA stabilizing factor, an RNA editing factor, an RNA-binding factor, an RNA localizing factor, or a combination thereof.

56. The system, LNP composition, cell or method of any one of the preceding claims, wherein the effector molecule is a translation factor, e.g., a translation factor provided in Table 4, e.g., eIF4G; Poly A binding protein (PABP); eIF3d or a component thereof; Dazl, or a fragment, or variant or combination thereof.

57. The system, LNP composition, cell or method of any one of the preceding claims, wherein the effector molecule is a translation factor which modulates, e.g., facilitates, ribosome binding, e.g., recruitment, pre-initiation complex formation, or RNA unwinding.

58. The system, LNP composition, cell or method of any one of the preceding claims, wherein the effector molecule comprises eIF4G, e.g., wildtype eIF4G, a variant of eIF4G, or a fragment thereof.

59. The system, LNP composition, cell or method of claim 58, wherein the eIF4G variant retains ribosome binding, e.g., recruitment.

60. The system, LNP composition, cell or method of any one of the preceding claims, wherein the effector molecule comprises RNA modifying enzymes such as terminal transferases e.g., TENT4A, TENT4B, Gld2, a variant or fragment thereof .

61. The system, LNP composition, cell or method of any one of the preceding claims, wherein the effector molecule comprises an amino acid sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

62. The system, LNP composition, cell or method of any one of the preceding claims, wherein the effector molecule is encoded by a nucleotide sequence provided in Table 2, or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

63. The system, LNP composition, cell or method of any one of the preceding claims, wherein the tether molecule binds to a binding element in the first polynucleotide.

64. The system, LNP composition, cell or method of any one of the preceding claims, wherein the tether molecule comprises a tether molecule provided in Table 1, e.g., MBP, PCP, Lambda N, U1A or PUF, 15.5kd, LARP7 or a variant or fragment thereof.

65. The system, LNP composition, cell or method of any one of the preceding claims, wherein the tether molecule comprises MBP.

66. The system, LNP composition, cell or method of any one of the preceding claims, wherein the tether molecule comprises an amino acid sequence provided in Table 2 or a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

67. The system, LNP composition, cell or method of any one of the preceding claims, wherein the tether molecule is encoded by a nucleotide sequence provided in Table 2 or

a sequence with at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity thereof.

68. The system, LNP composition, cell or method of any one of the preceding claims, wherein the therapeutic payload or a prophylactic payload comprises an mRNA encoding: a secreted protein, a membrane-bound protein; or an intercellular protein.

69. The system, LNP composition, cell or method of claim 65, wherein the therapeutic payload or a prophylactic payload is chosen from a cytokine, an antibody, a vaccine (e.g., an antigen, an immunogenic epitope), a receptor, an enzyme, a hormone, a transcription factor, a ligand, a membrane transporter, a structural protein, a nuclease, a growth factor, an immune modulator, or a component, variant or fragment (e.g., a biologically active fragment) thereof.

70. The system, LNP composition, cell or method of any one of the preceding claims, which results in one, two, three, four, five, six or all, or any combination thereof, of the following in a cell (e.g., in a cell contacted with the system or LNP composition):

- (i) increased expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (ii) sustained expression and/or level of mRNA encoding the therapeutic payload or prophylactic payload;
- (iii) increased expression and/or level of therapeutic payload or prophylactic payload;
- (iv) sustained expression and/or level of therapeutic payload or prophylactic payload;
- (v) increased stability of mRNA encoding the therapeutic payload or prophylactic payload;
- (vi) increased resistance of translation of therapeutic payload or prophylactic payload to cellular environment, e.g., stress or nutrient deprivation or translation factor availability;
- (vii) reduced dosing of the therapeutic payload or prophylactic payload; or

(viii) reduced toxicity, e.g., reduced modulation of a protein translated from endogenous mRNA in a cell.

71. The system, LNP composition, cell or method of any one of the preceding claims, wherein:

(a) the first polynucleotide comprises:

(1) a sequence encoding a therapeutic payload or a prophylactic payload, and

(2) a binding element comprising an MS2 sequence, e.g., 6 MS2 sequences of 19 nucleotides separated by spacers of 20 nucleotides in length;

(b) the second polynucleotide comprises a sequence encoding:

(1) an effector molecule comprising eIF4G, e.g., wildtype eIF4G, a variant or a fragment thereof; and

(2) a tether molecule comprising MBP, e.g., wildtype MBP, a variant or fragment thereof.

72. The system, LNP composition, cell or method of any one of the preceding claims, wherein the first polynucleotide comprises an mRNA comprising at least one chemical modification.

73. The system, LNP composition, cell or method of any one of the preceding claims, wherein the second polynucleotide comprises an mRNA comprising at least one chemical modification.

74. The system, LNP composition, cell or method of any one of claims 1-71, wherein the first polynucleotide comprises an mRNA which does not have any chemical modification.

75. The system, LNP composition, cell or method of any one of claims 1-71 and 74, wherein the second polynucleotide comprises an mRNA which does not have any chemical modification.

76. The system, LNP composition, cell or method of any one of claims 72-75, wherein the chemical modification is selected from the group consisting of pseudouridine, N1-methylpseudouridine, 2-thiouridine, 4'-thiouridine, 5-methylcytosine, 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-methyluridine, 5-methoxyuridine, and 2'-O-methyl uridine.

77. The system, LNP composition, cell or method of any one of the preceding claims, wherein the LNP composition comprises: (i) an ionizable lipid, e.g., an amino lipid; (ii) a sterol or other structural lipid; (iii) a non-cationic helper lipid or phospholipid; and (iv) a PEG-lipid.

78. The system, LNP composition, cell or method of claim 77, wherein the ionizable lipid comprises an amino lipid.

79. The system, LNP composition, cell or method of claim 77 or 78, wherein the ionizable lipid comprises a compound of any of Formulae (I), (IA), (IB), (II), (IIa), (IIb), (IIc), (IId), (IIe), (IIf), (IIg), (III), (IIIa1), (IIIa2), (IIIa3), (IIIa4), (IIIa5), (IIIa6), (IIIa7), or (IIIa8).

80. The system, LNP composition, cell or method of any one of claims 77-79, wherein the non-cationic helper lipid or phospholipid comprises a compound chosen from DSPC (e.g., a variant of DSPC, e.g., a compound of Formula (IV)); DPPC; or DOPC.

81. The system, LNP composition, cell or method of any one of claims 77-80, wherein the structural lipid is alpha-tocopherol, β -sitosterol or cholesterol.

82. The system, LNP composition, cell or method of any one of claims 77-81, wherein the PEG lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof.

83. The system, LNP composition, cell or method of any one of claims 77-82, wherein the PEG lipid is selected from the group consisting of PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC and PEG-DSPE lipid.

84. The system, LNP composition, cell or method of any one of claims 77-81, wherein the PEG lipid is chosen from of a compound of Formula (V), Formula (VI-A), Formula (VI-B), Formula (VI-C) or Formula (VI-D).

85. The system, LNP composition, cell or method of claim 84, wherein the PEG-lipid is a compound of Formula (VI-A).

86. The system, LNP composition, cell or method of claim 84, wherein the PEG-lipid is a compound of Formula (VI-B).

87. The system, LNP composition, cell or method of claim 84, wherein the PEG-lipid is a compound of Formula (VI-C).

88. The system, LNP composition, cell or method of claim 84, wherein the PEG-lipid is a compound of Formula (VI-D).

89. The system, LNP composition, cell or method of any one of claims 77-88, wherein the LNP comprises a molar ratio of about 20-60% ionizable lipid: 5-25% phospholipid: 25-55% cholesterol; and 0.5-15% PEG lipid.

90. The system, LNP composition, cell or method of any one of the preceding claims, which is formulated for intravenous, subcutaneous, intramuscular, intranasal, intraocular, rectal, pulmonary or oral delivery.

91. The system, LNP composition, cell or method of any one of the preceding claims, wherein the subject is a mammal, e.g., a human.

Fig. 1

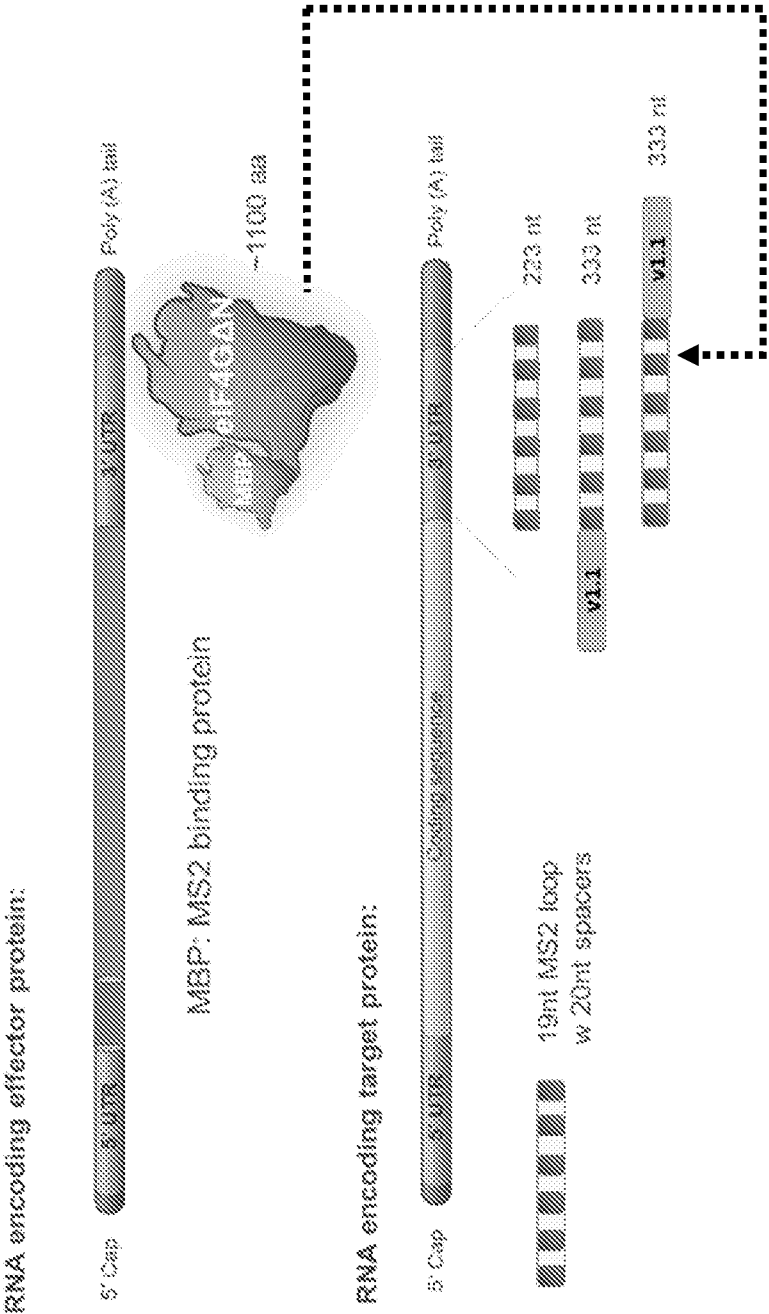


Fig. 2A

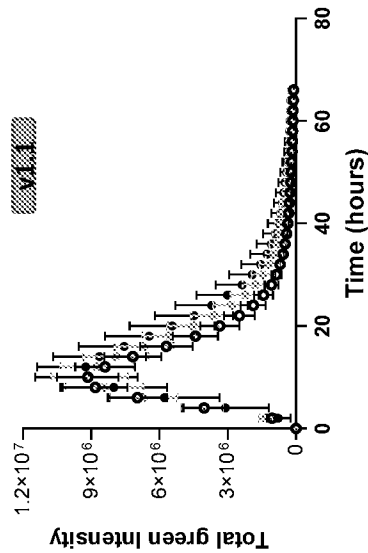


Fig. 2C

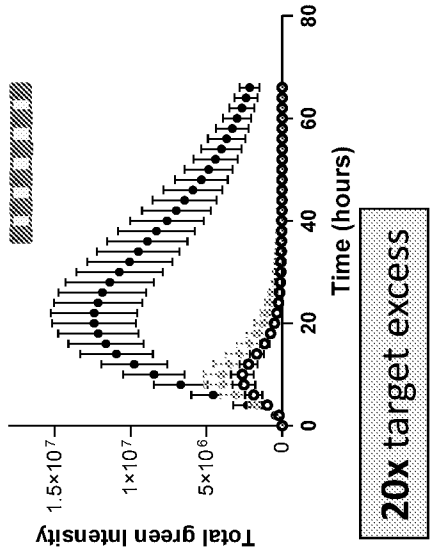


Fig. 2B

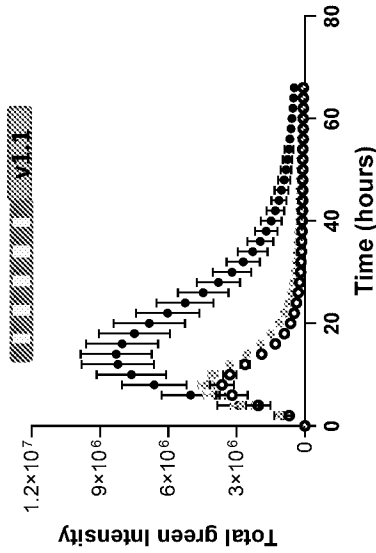
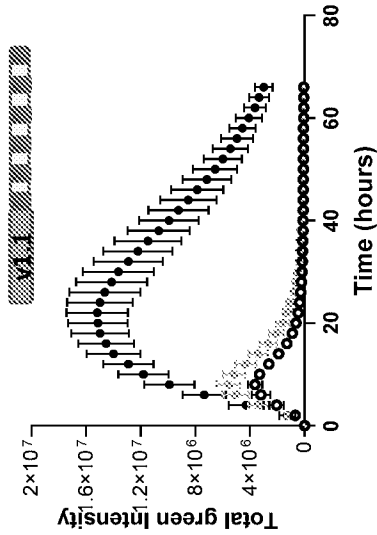


Fig. 2D



○ + non-tethered control
● + tethered control
● + tethered effector

Fig. 3A

HeLa

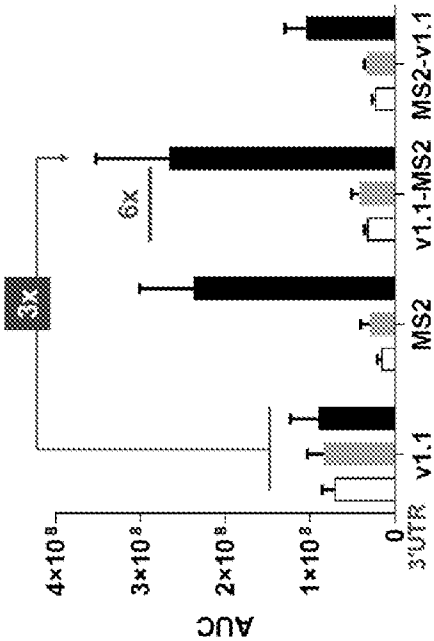


Fig. 3B

HEK293

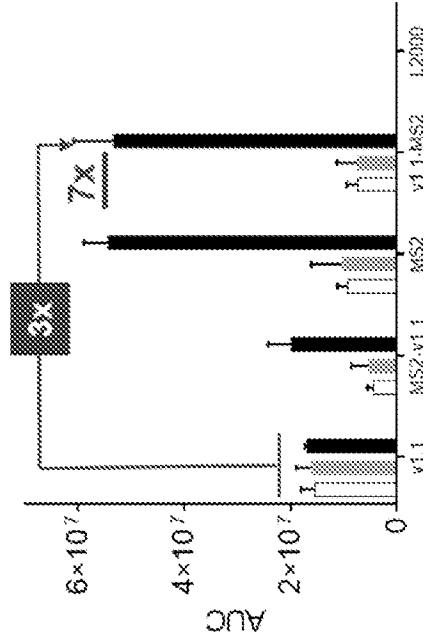
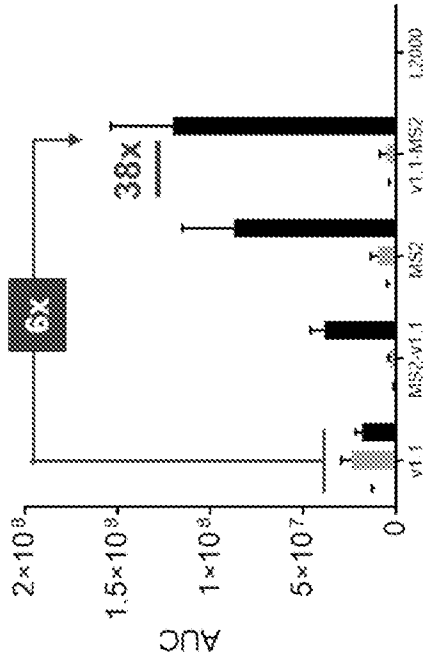


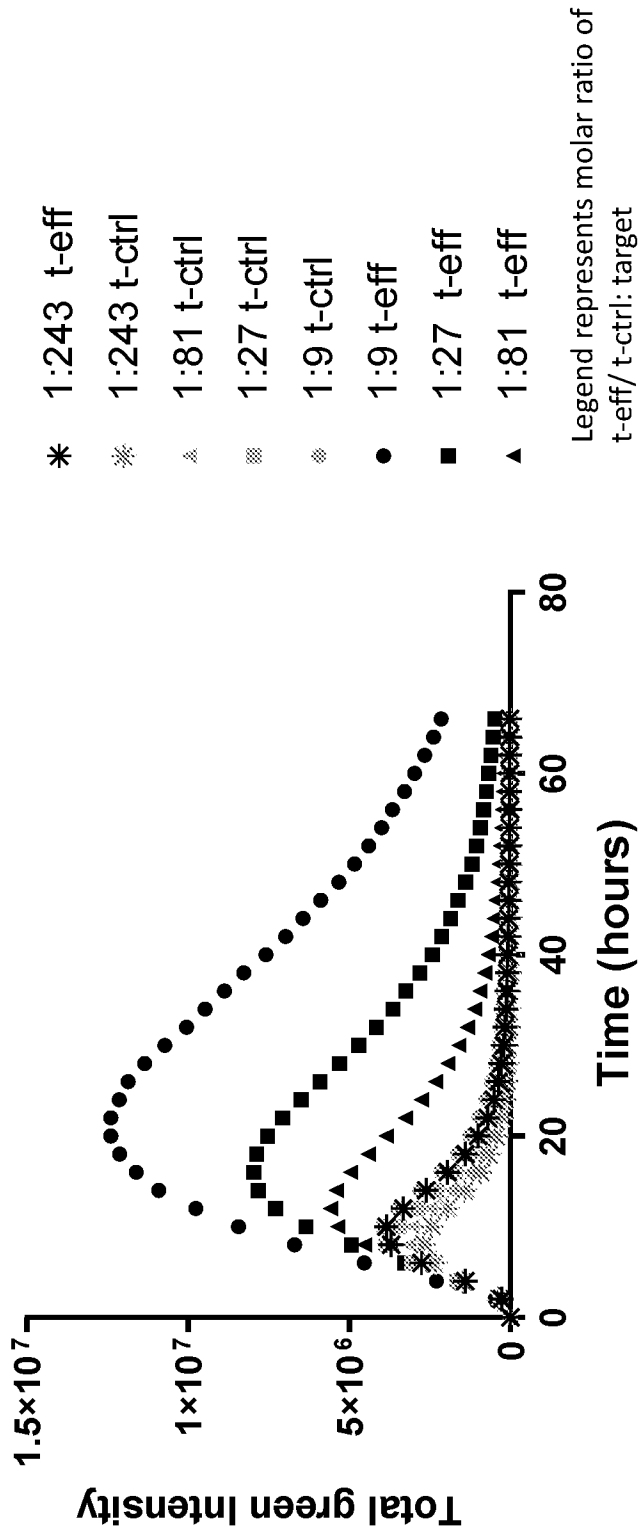
Fig. 3C

Hep3b



- + non-tethered control
- ◐ + tethered control
- + tethered effector

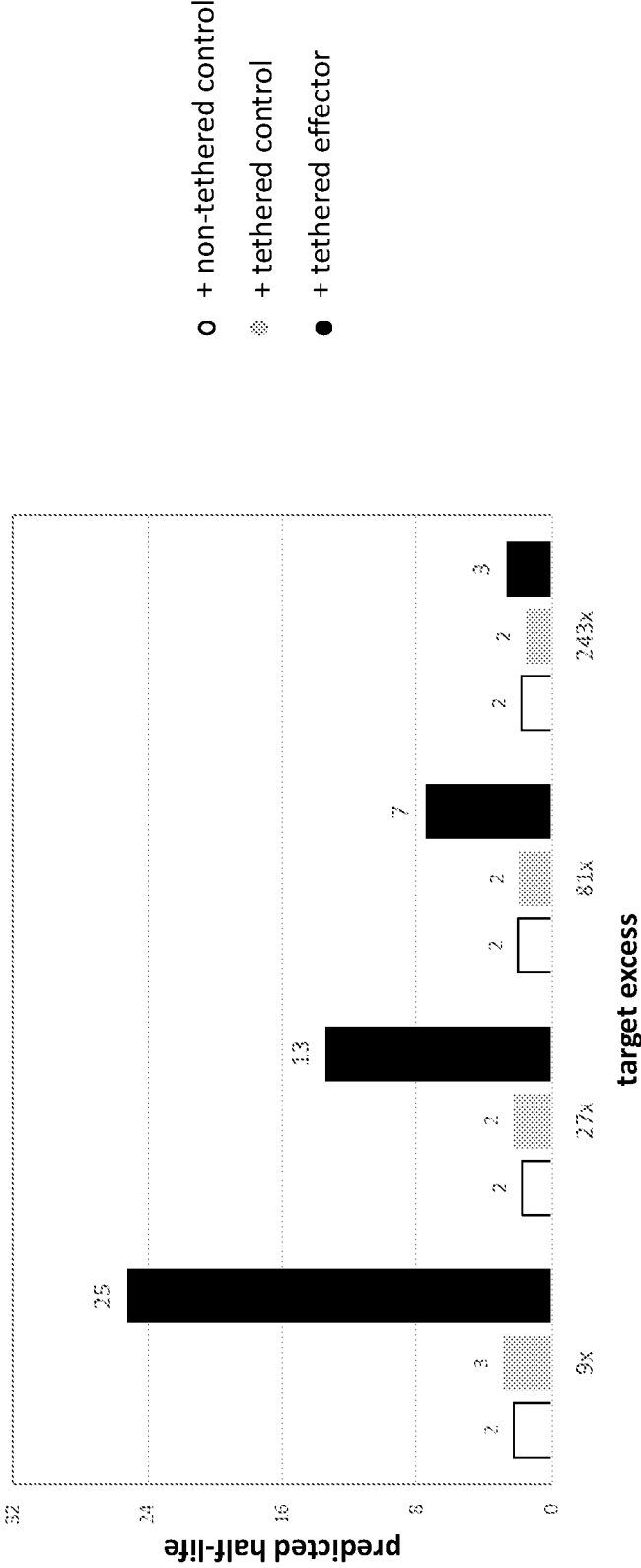
Fig. 4A



Fixed target RNA + increasing amounts of effector
(20 ng) + (~0.28 ng → 6.8 ng)

3'UTR:  pA

Fig. 4B



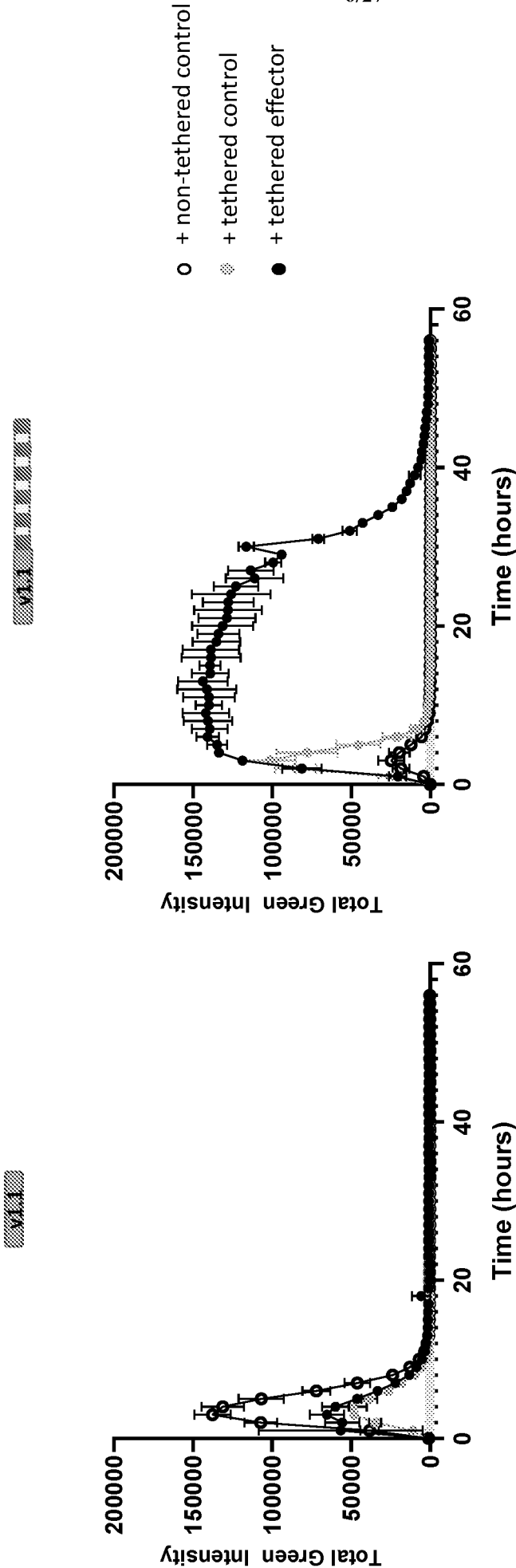


Fig. 5B

Fig. 5A

Fig. 5C

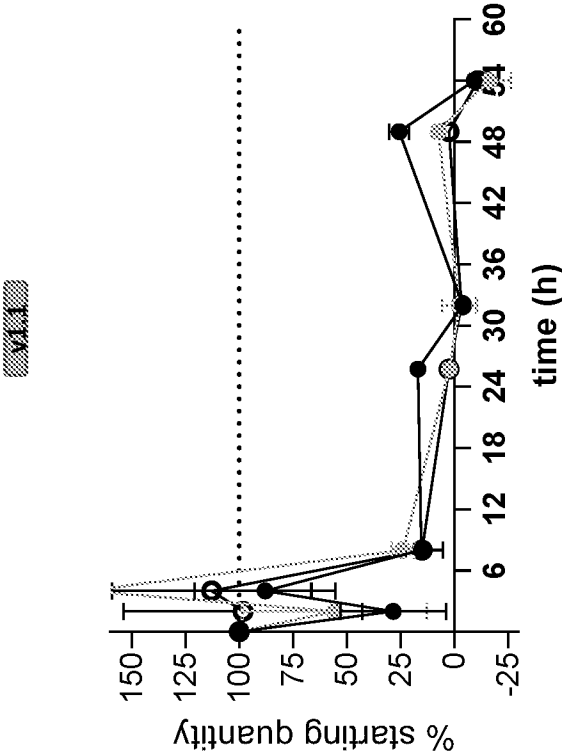
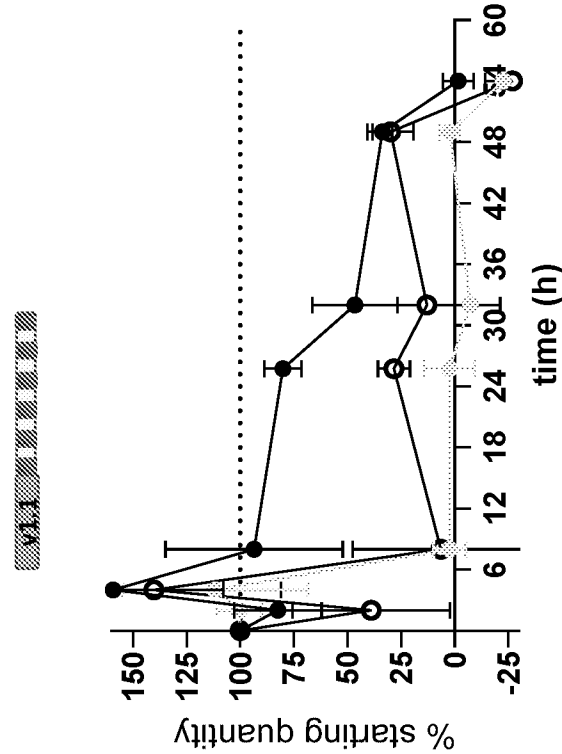


Fig. 5D



- + non-tethered control
- + tethered control
- + tethered effector

8/27

Fig. 6A

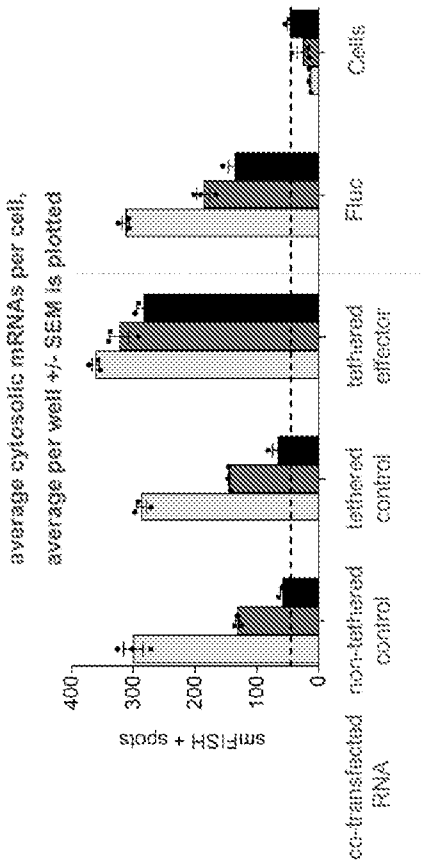
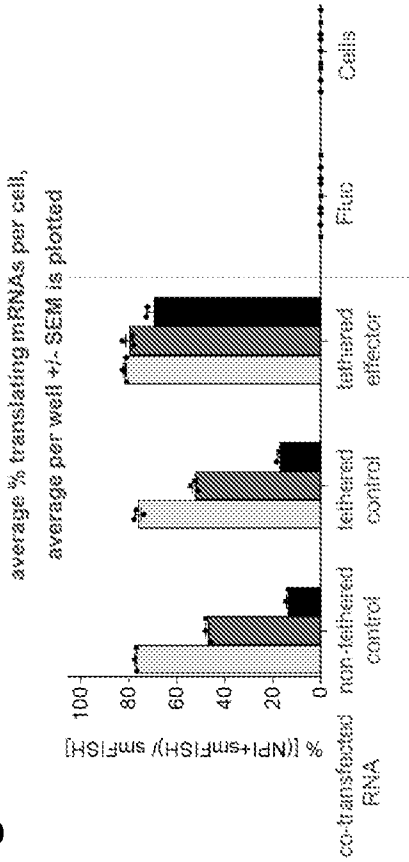


Fig. 6B



4h
6h
9h

Fig. 6C

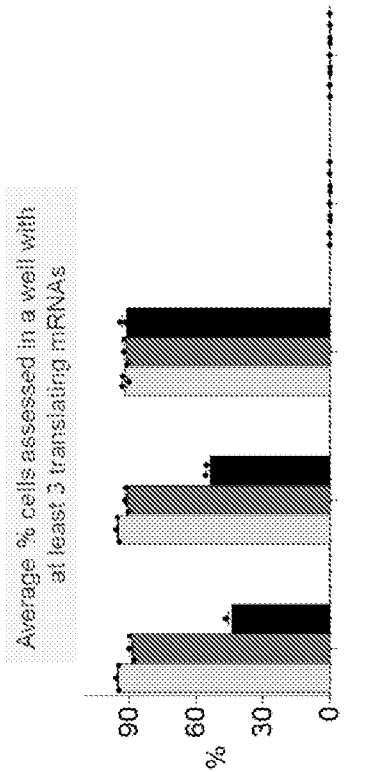


Fig. 6D

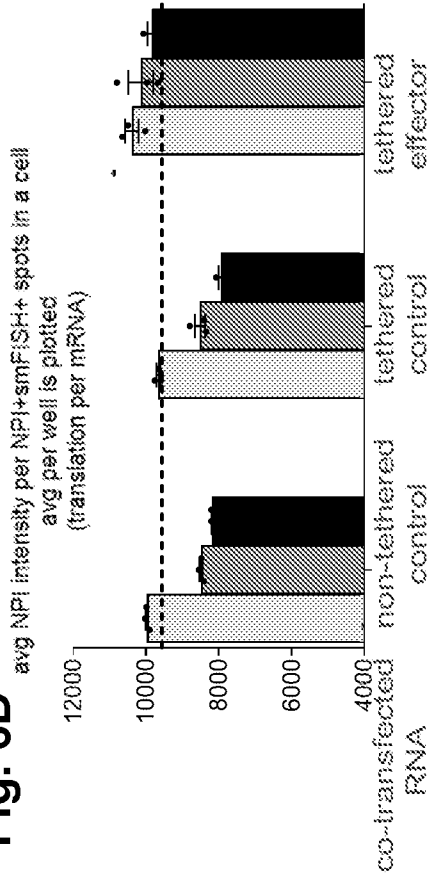


Fig. 6E

3'UTR: 

average Npi intensity of Npi+ snFISH+ spots per cell, plotted is individual cells

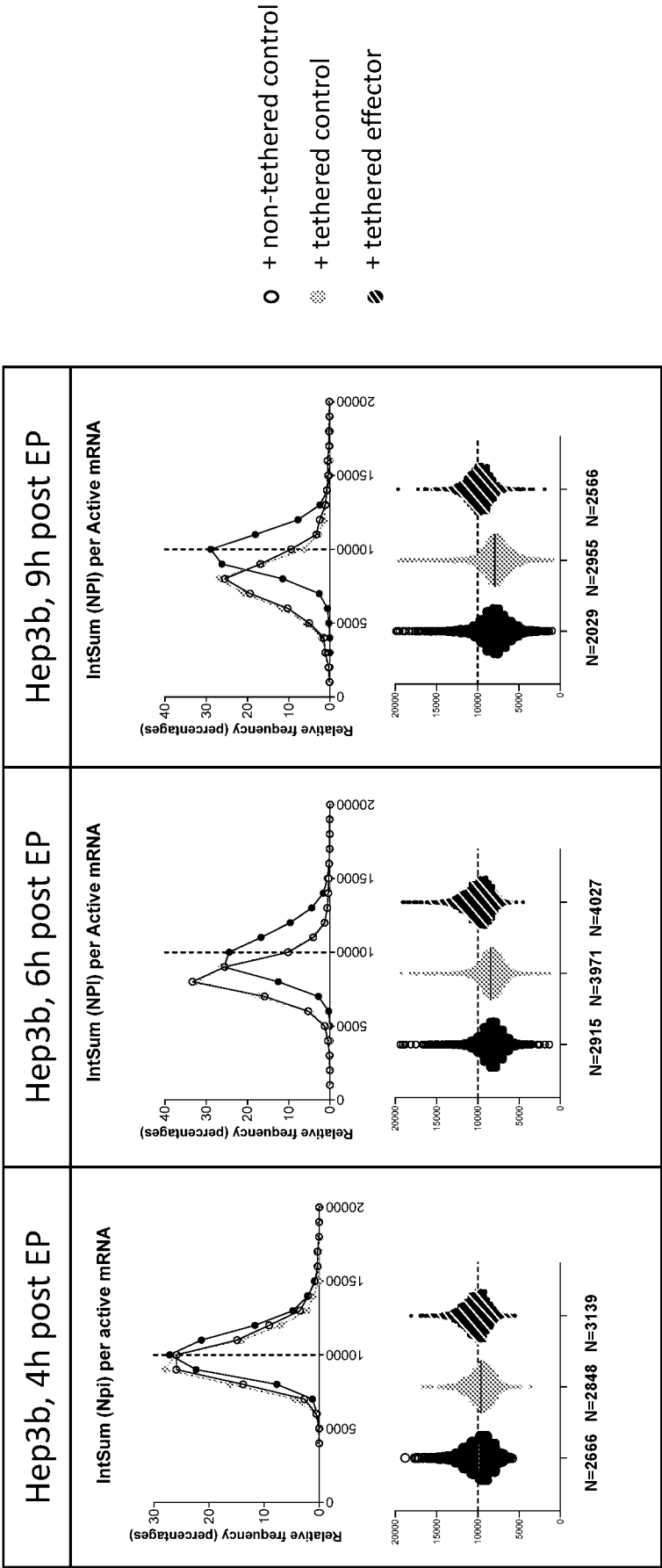


Fig. 7A

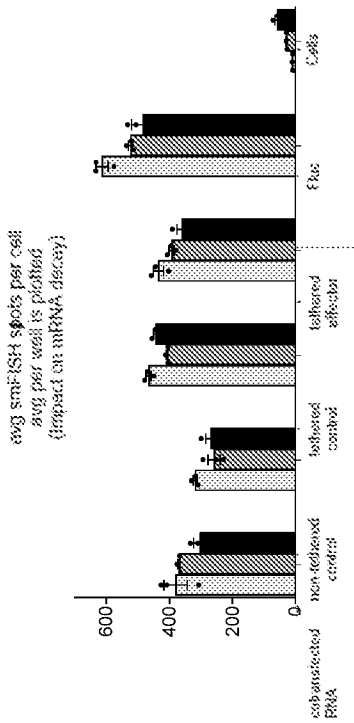


Fig. 7B

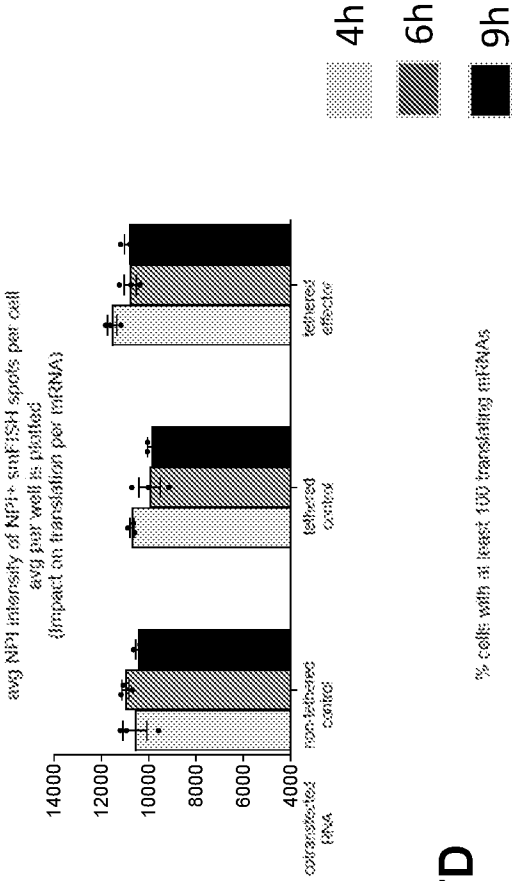


Fig. 7C

avg % NP1+smFISH+ smFISH+ spots per cell
avg per well is plotted
(impact on % of translating mRNAs)

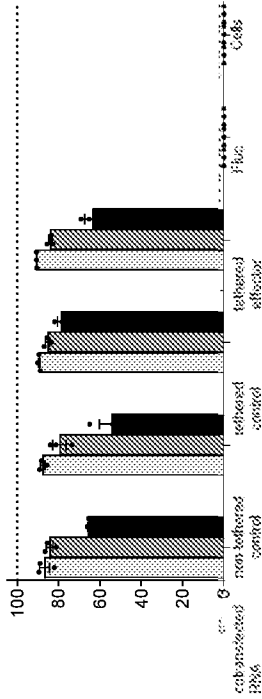


Fig. 7D

% cells with at least 100 translating mRNAs

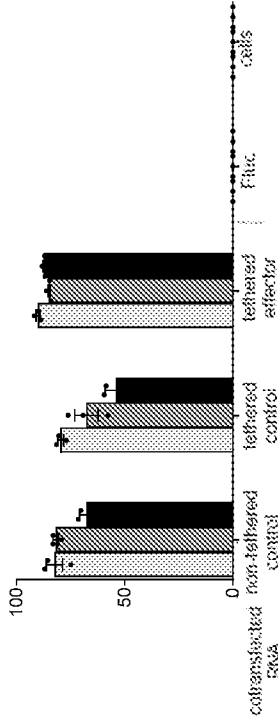


Fig. 8A

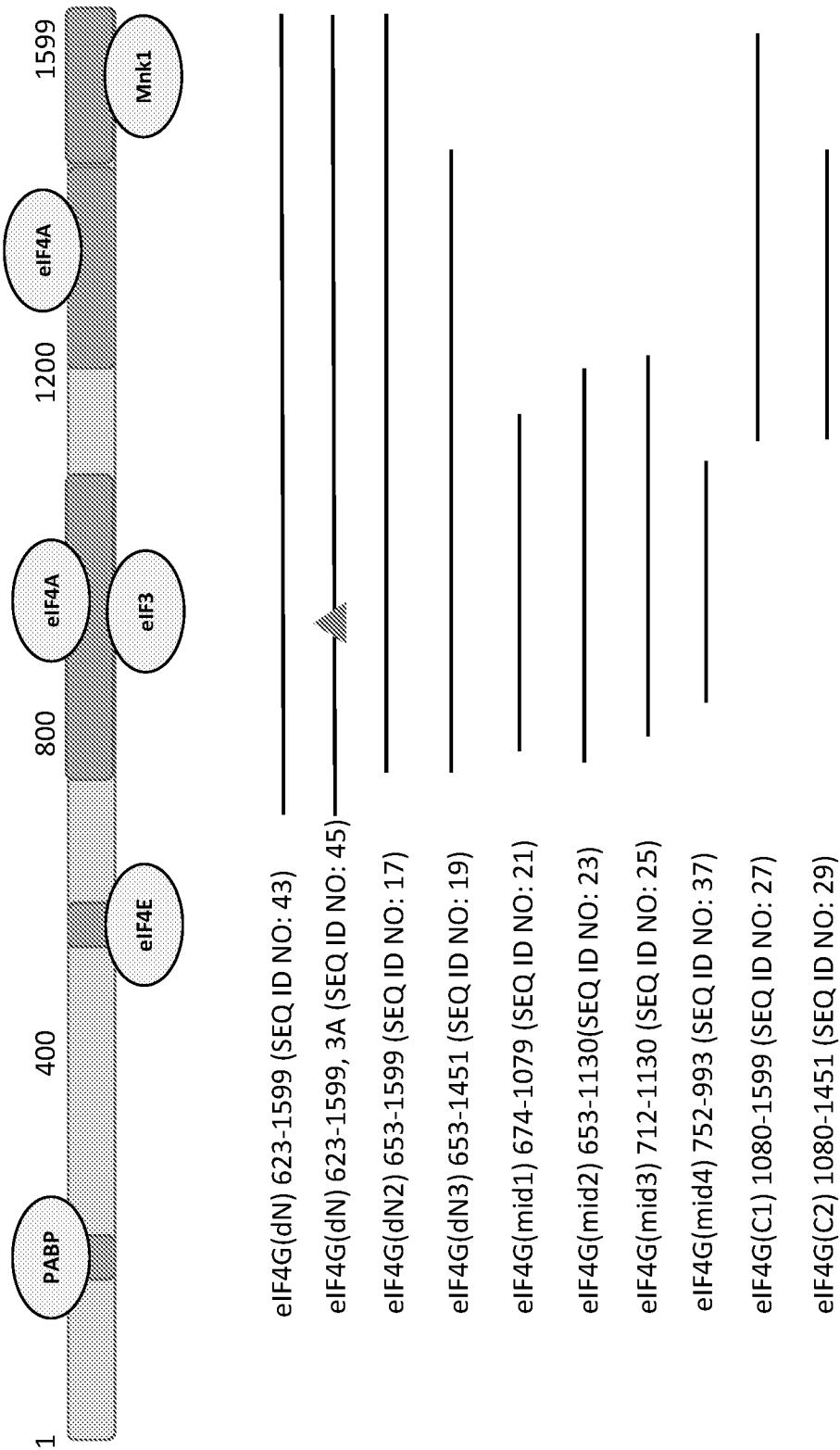


Fig. 8B

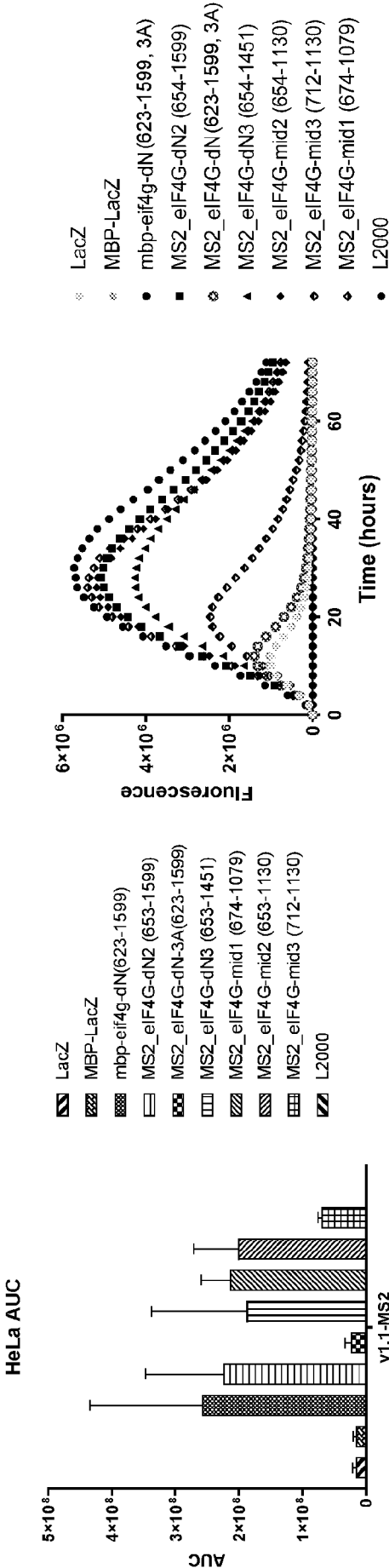


Fig. 8C

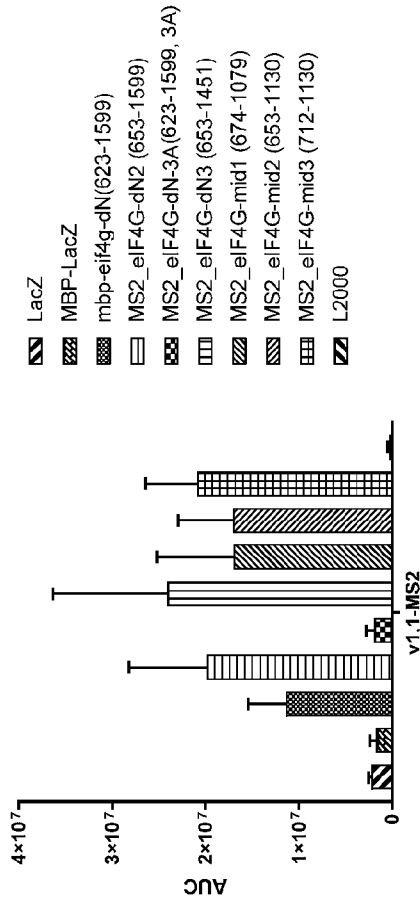


Fig. 8D

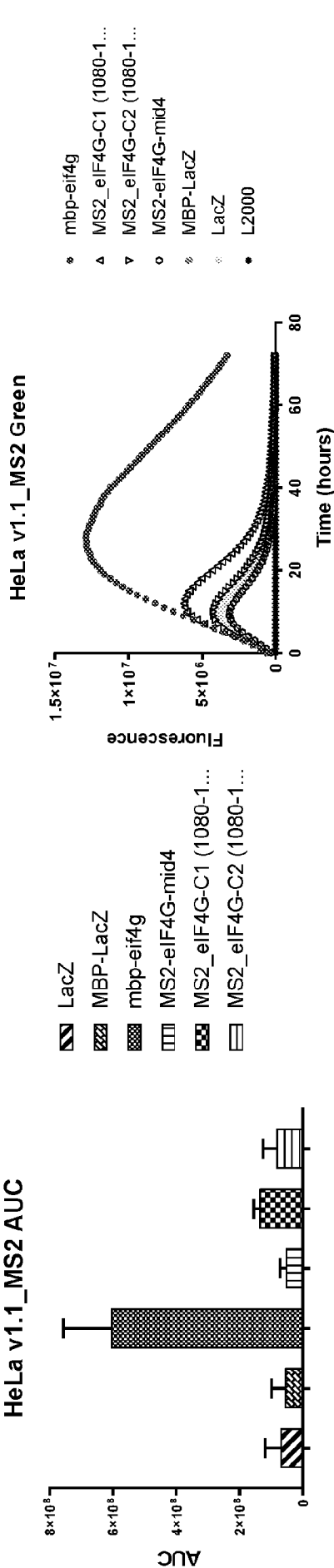


Fig. 8E

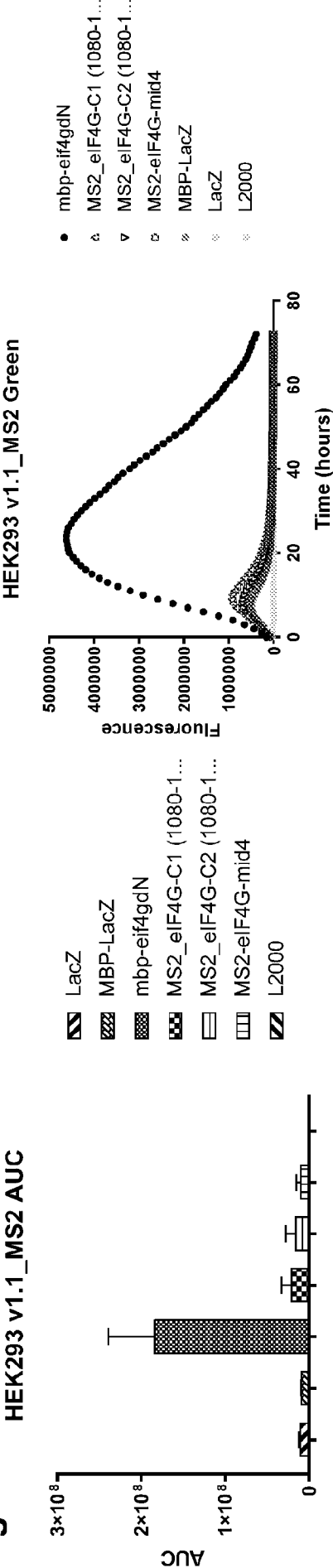


Fig. 9

binding sites for	PABP	elF4e	elF4A ₁ /elF3	elF4A ₁ , elF3	Mnk1	sequence details	length (aa)	effective
elF4G-fl						1-1599	1599	Not tested
elF4G-dN						623-1599	976	Yes
elF4G-dN-mid_mutant						623-1599,768-771-776->A	976	No
elF4G-dN2						653-1599	945	Yes
elF4G-dN3						653-1451	797	Yes
elF4G-mid1						674-1079	405	Yes
elF4G-mid2						653-1130	476	Yes
elF4G-mid3						712-1130	418	Yes
elF4G-mid4*						752-993	241	No
elF4G-C1						1080-1599	519	small boost?
elF4G-C2						1080-1451	371	No

*unlikely to bind eIF3
binding site absent
binding site present
mutant binding site

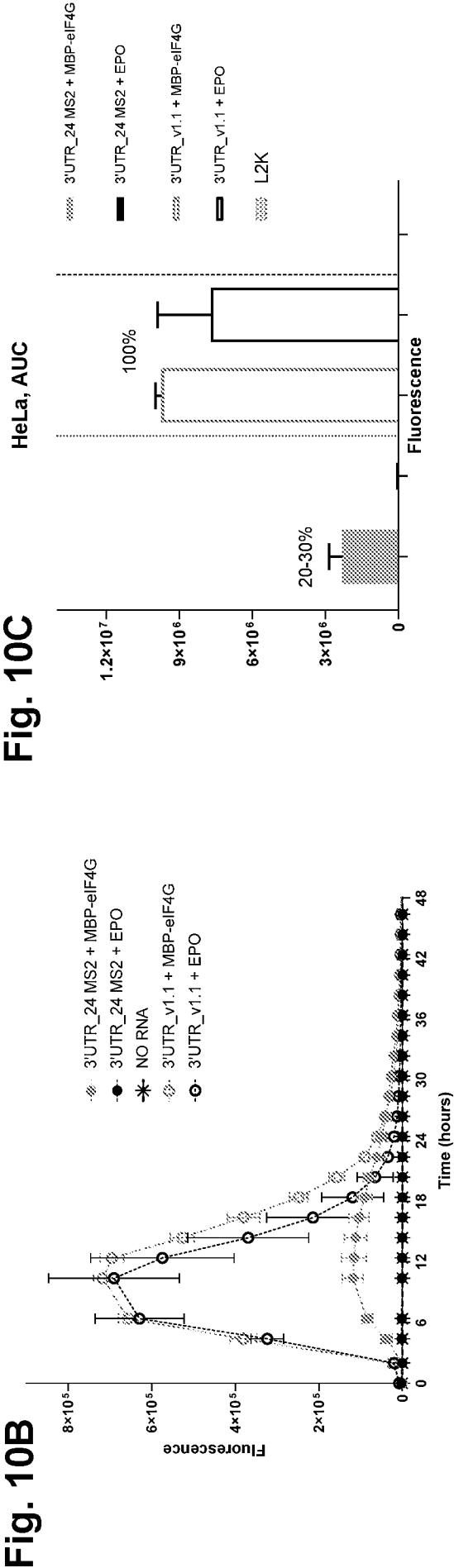
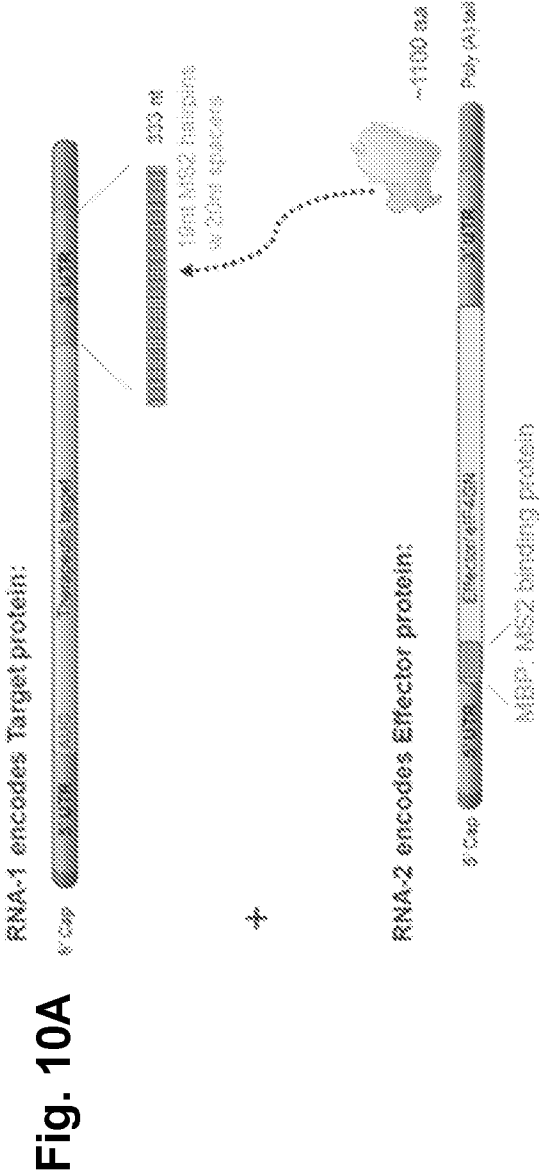


Fig. 11

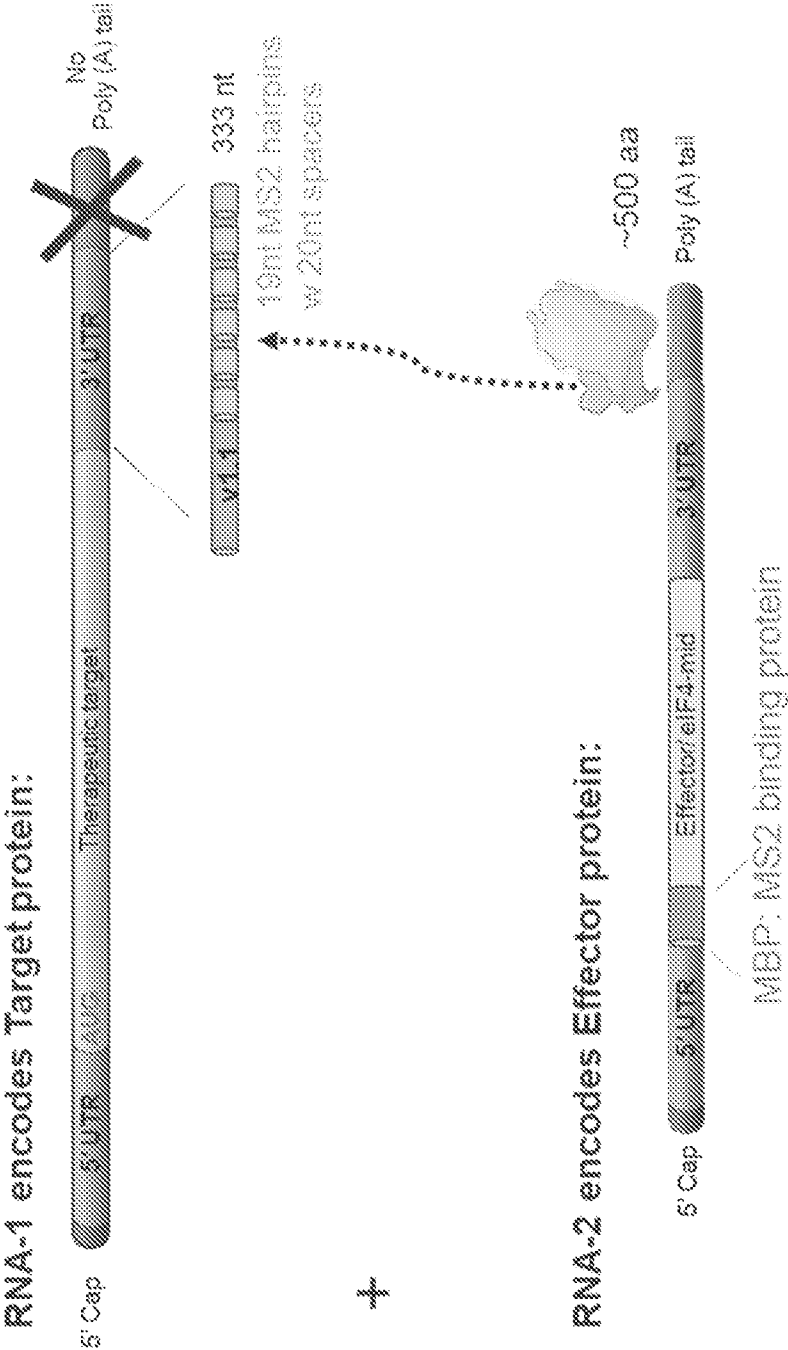


Fig. 12A

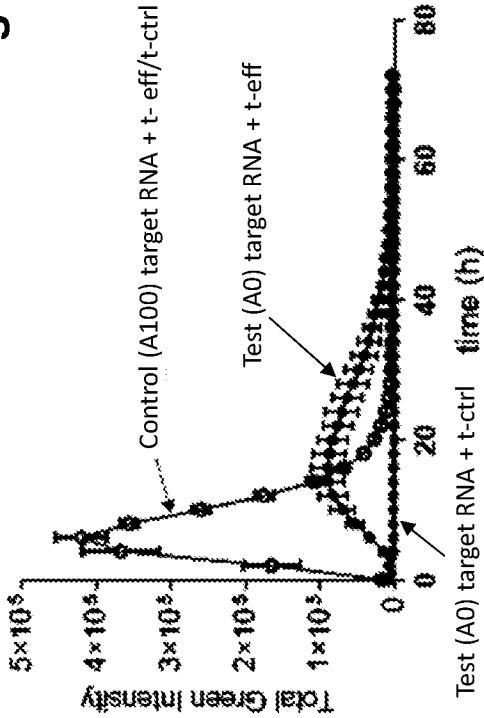


Fig. 12B

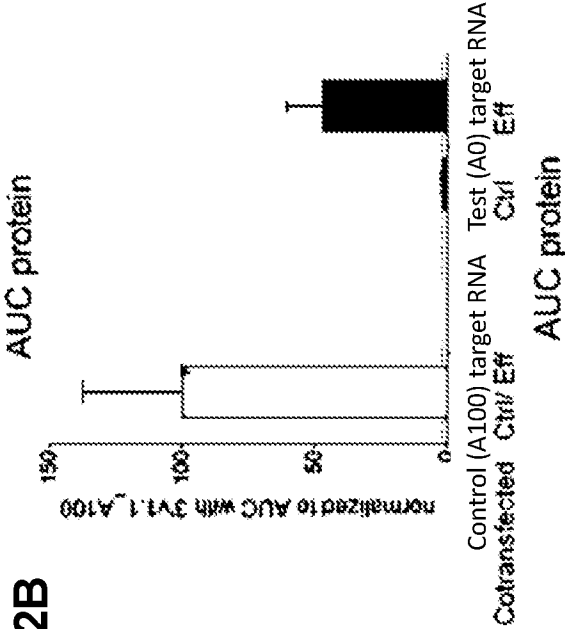


Fig. 12C

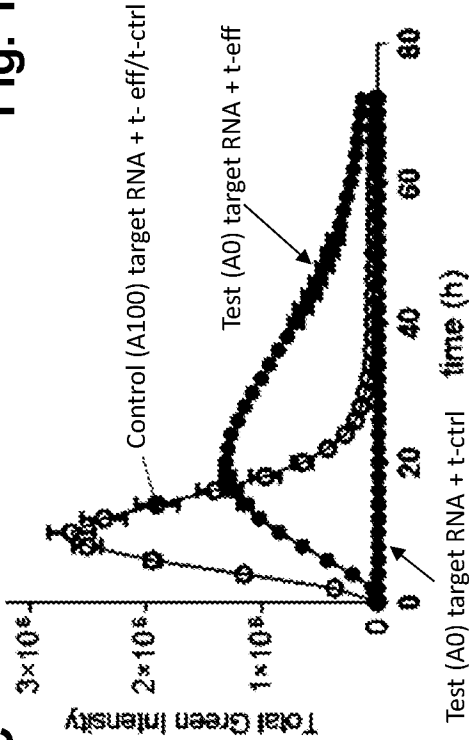


Fig. 12D

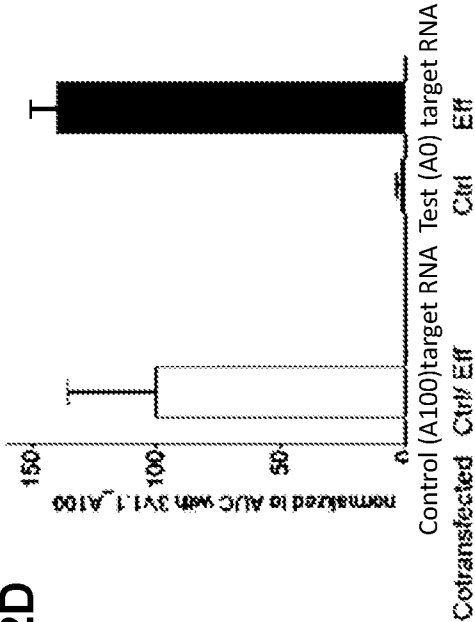


Fig. 13

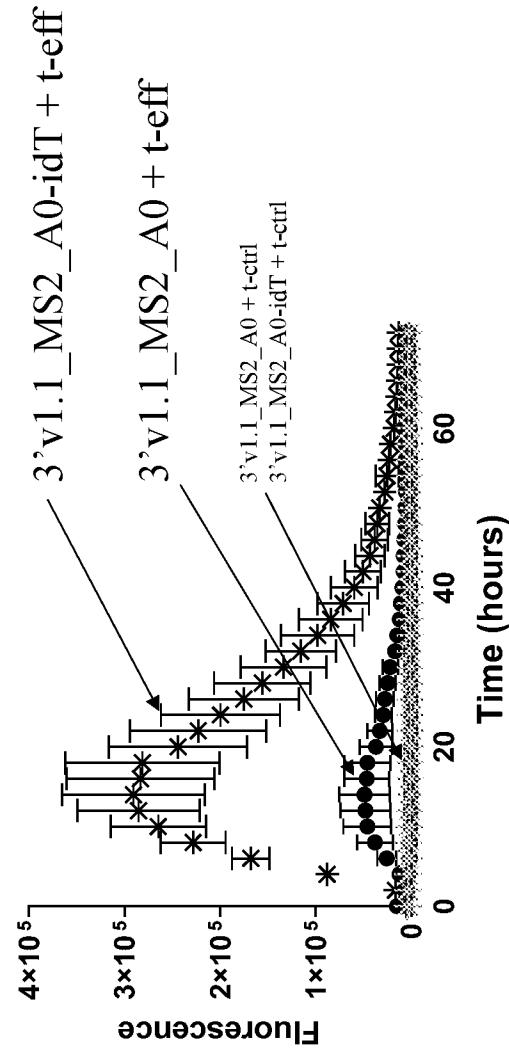


Fig. 14

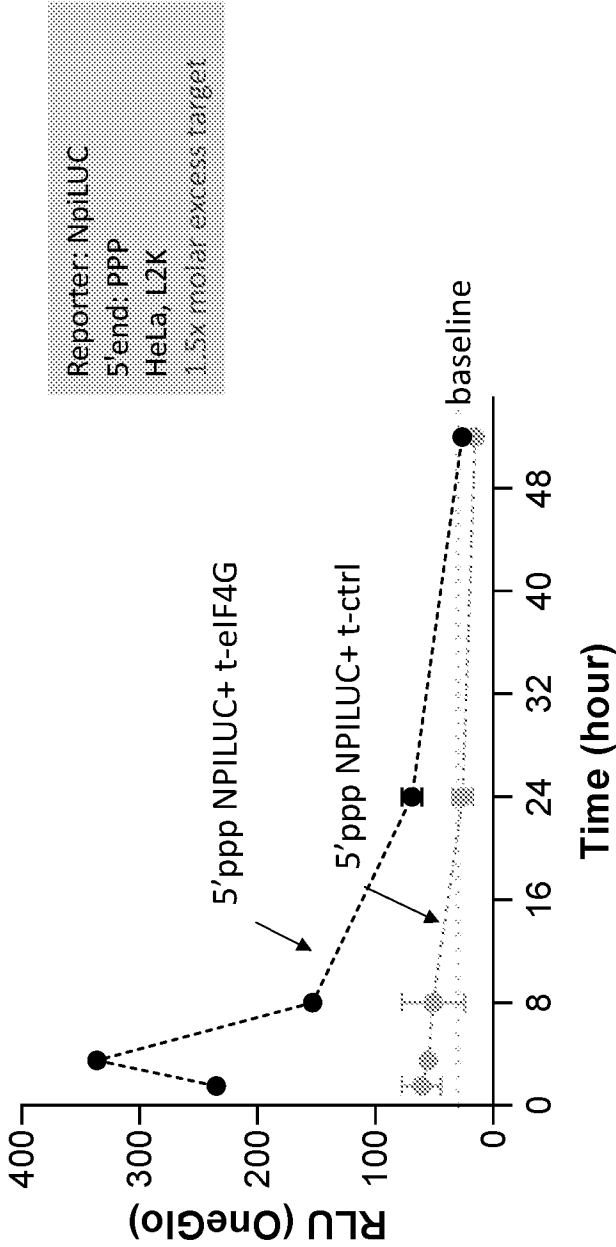


Fig. 15

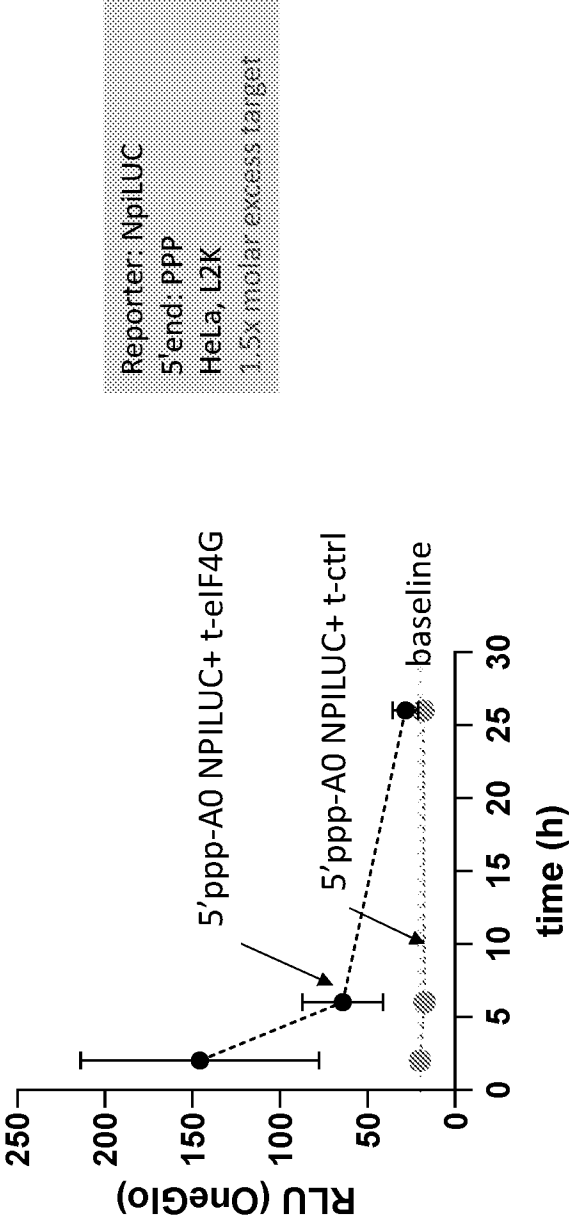


Fig. 16

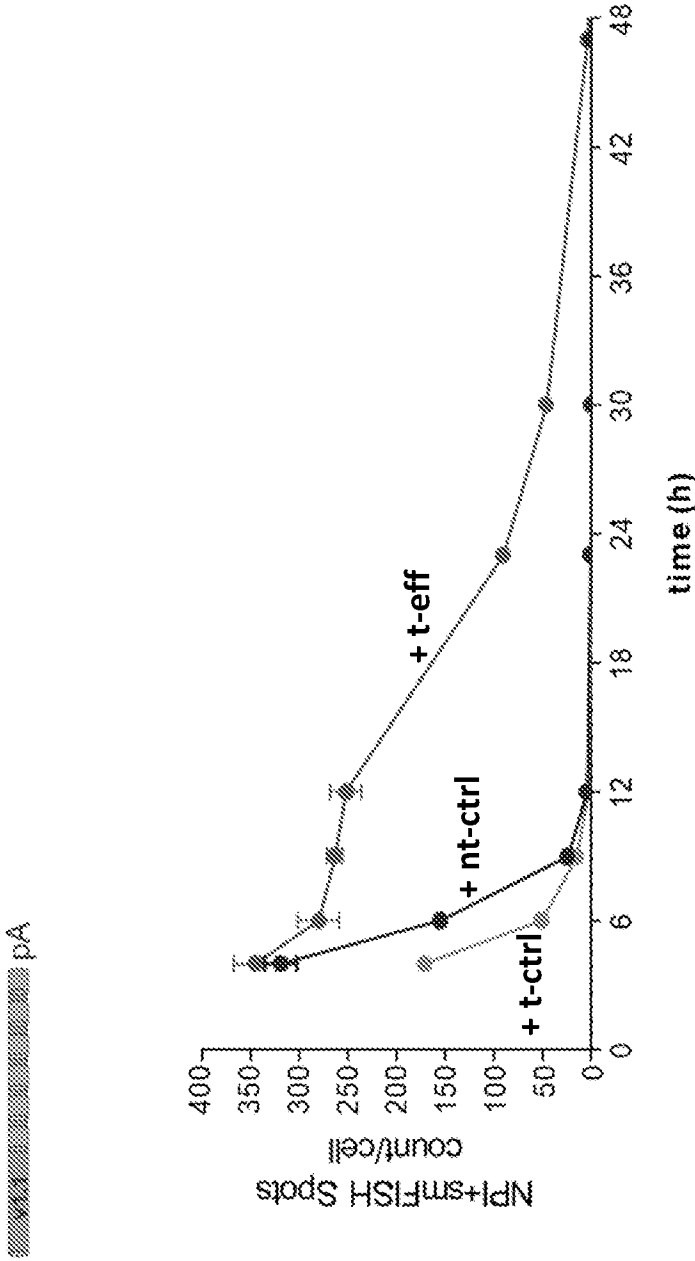


Fig. 17A

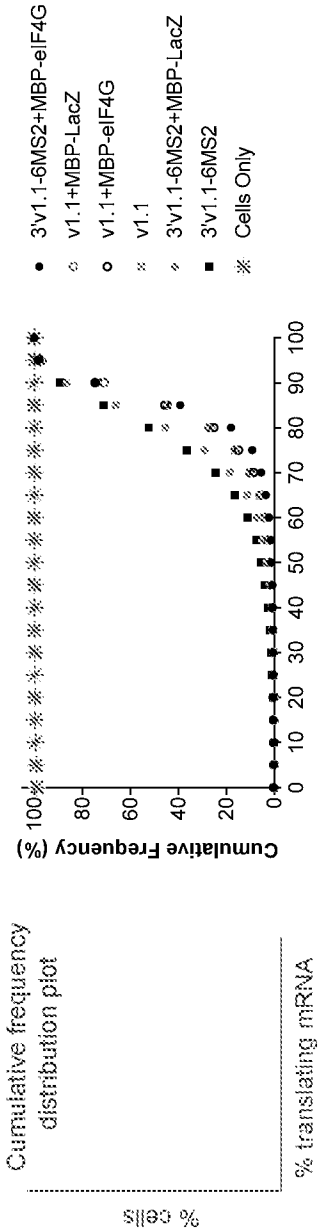


Fig. 17B

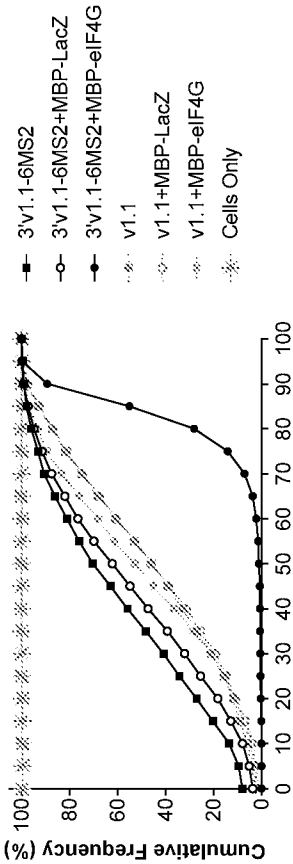


Fig. 17C

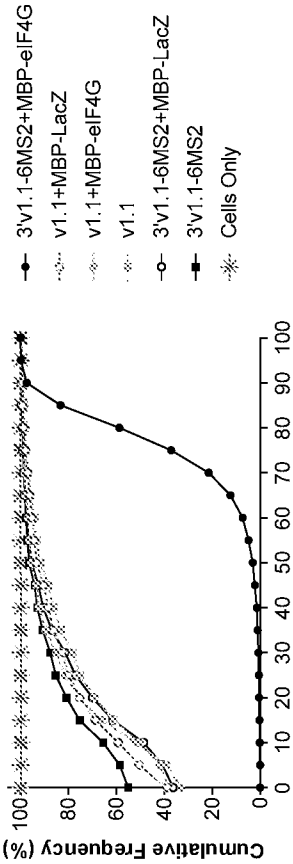
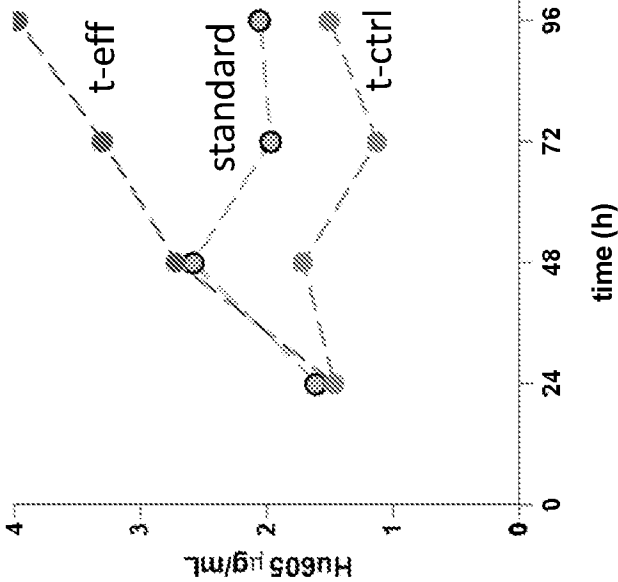


Fig. 18

Optimized orf, Secreted, 2 targetRNAs encoding for Hc and Lc of the target Antibody



Hek293, L2K
Eff= eIF4GΔN, ctrl=LACZ

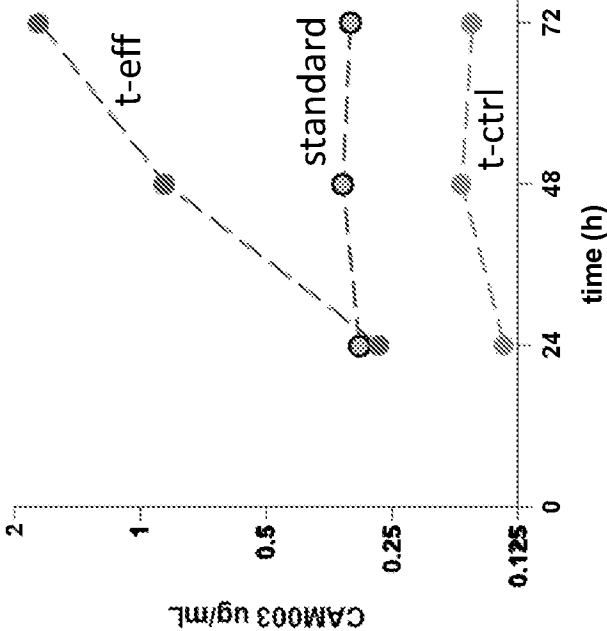


Fig. 19

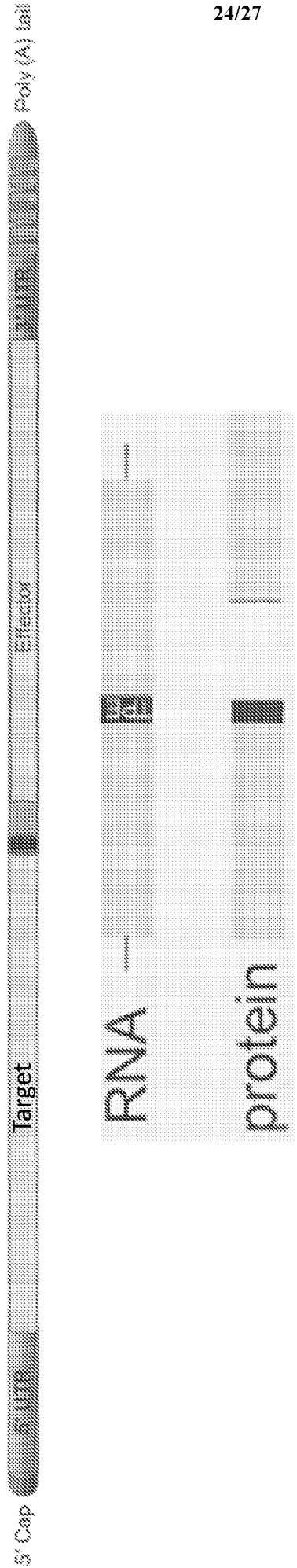


Fig. 20A

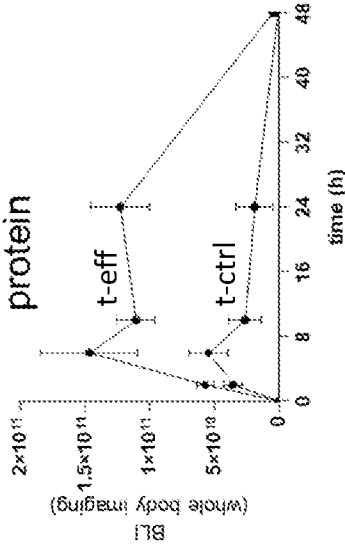


Fig. 20B

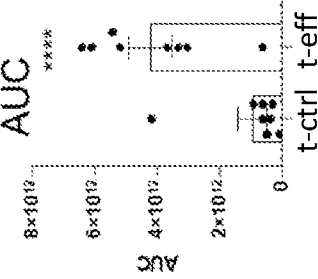


Fig. 20C

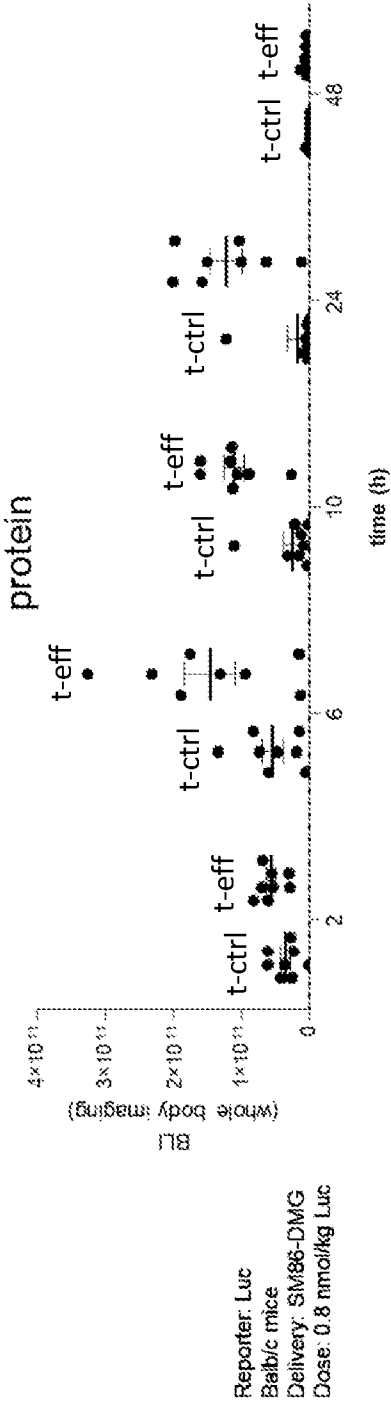


Fig. 21A

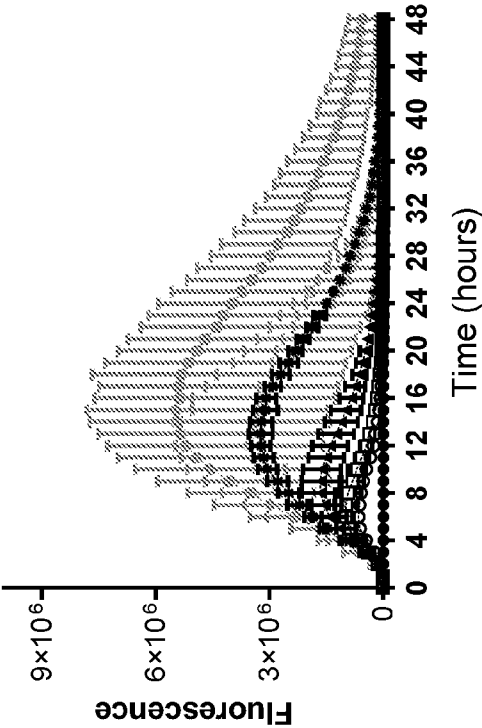


Fig. 21B

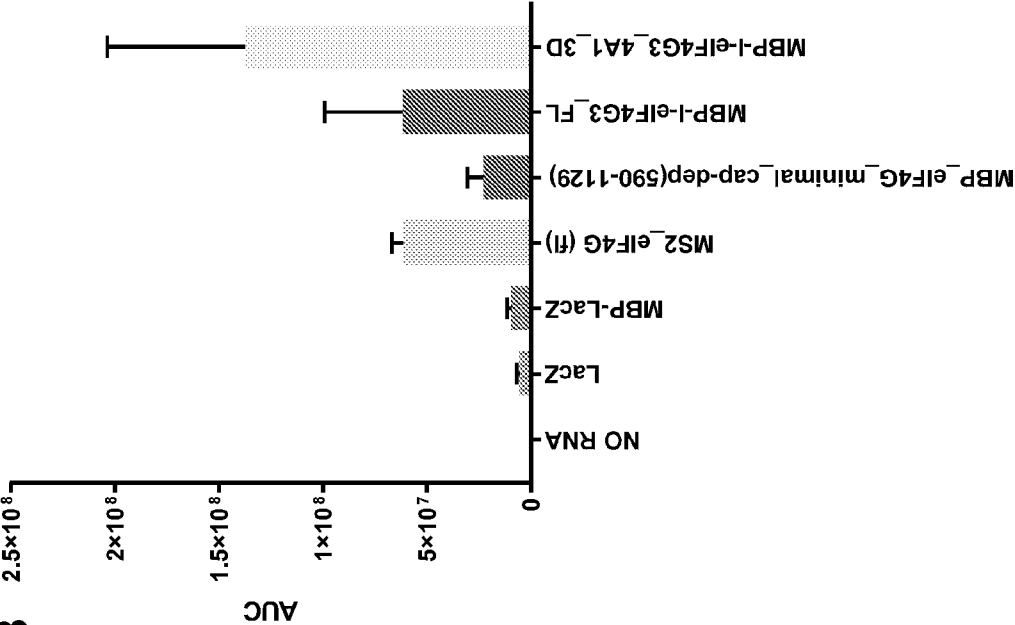


Fig. 22B

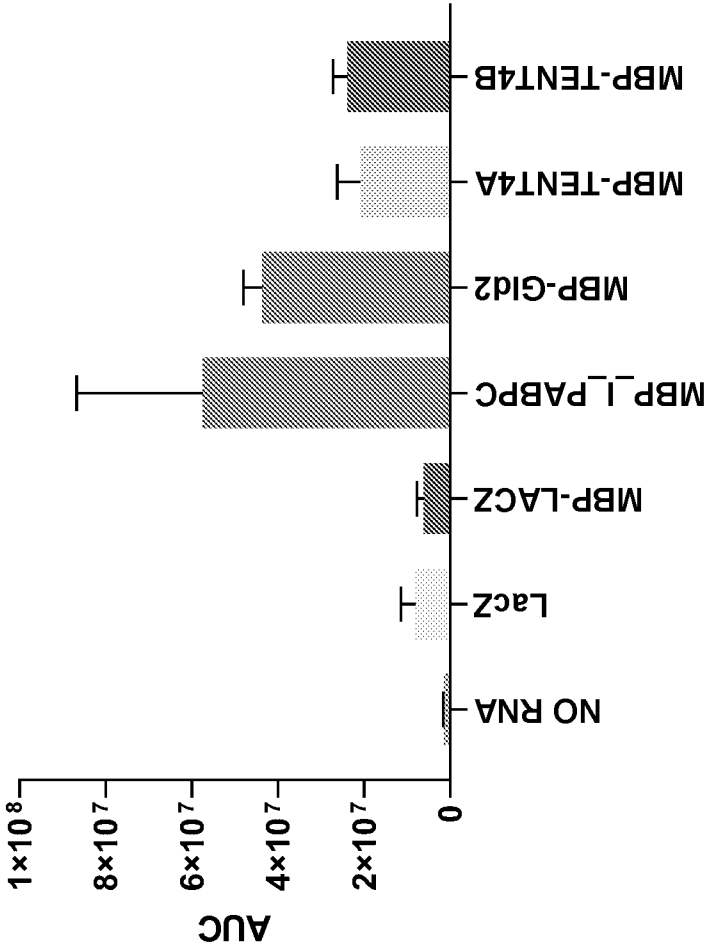
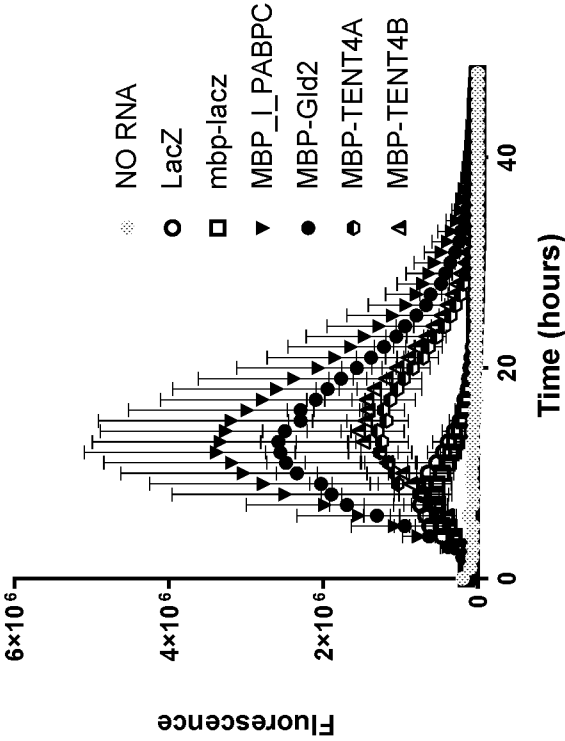


Fig. 22A



SEQUENCE LISTING

<110> MODERNATX, INC.

<120> LNP COMPOSITIONS COMPRISING AN MRNA THERAPEUTIC AND AN EFFECTOR MOLECULE

<130> 45817-0103W01

<140>

<141>

<150> 63/183,119

<151> 2021-05-03

<150> 63/024,862

<151> 2020-05-14

<160> 157

<170> PatentIn version 3.5

<210> 1

<211> 333

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 1

tgataatagg ctggagcctc ggtggcctag cttcttgccc cttgggcctc cccccagccc	60
ctcctcccct tcctgcaccc gtacccccgt ggtctttgaa taaagtctga gtgggcggcc	120
aggtatgatt acaacctgtg ctactttaga tgactattga aagaccatta ggctttctaa	180
cggacgctca cgagagaggg gcaccaatag ggccccacgg cacttcaaag gtttggcttg	240
gagtagtaac ccaagcagca acagttttga ctttcggacc accatcaggg gtcccacgtt	300
gggaacacgt aactctccta ctaacaagag gag	333

<210> 2

<211> 333

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 2

tgataatagc aggtatgatt acaacctgtg ctactttaga tgactattga aagaccatta	60
ggctttctaa cggacgctca cgagagaggg gcaccaatag ggccccacgg cacttcaaag	120
gtttggcttg gagtagtaac ccaagcagca acagttttga ctttcggacc accatcaggg	180
gtcccacgtt gggaacacgt aactctccta ctaacaagag gaggtctggag cctcggtggc	240
ctagcttctt gcccttggg cctcccccca gccctcctc cccttcctgc acccgtaacc	300
ccgtggtctt tgaataaagt ctgagtgggc ggc	333

<210> 3

<211> 223

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 3

tgataatagc aggtatgatt acaacctgtg ctactttaga tgactattga aagaccatta	60
ggctttctaa cggacgctca cgagagaggg gcaccaatag ggccccacgg cacttcaaag	120
gtttggcttg gagtagtaac ccaagcagca acagttttga ctttcggacc accatcaggg	180
gtcccacgtt gggaacacgt aactctccta ctaacaagag gag	223

<210> 4

<211> 119

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 4

tgataatagg ctggagcctc ggtggcctag cttcttgccc cttgggcctc cccccagccc	60
ctcctcccct tcctgcaccc gtacccccgt ggtctttgaa taaagtctga gtgggcggc	119

<210> 5

<211> 57

<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 5
gggaaataag agagaaaaga agagtaagaa gaaatataag accccggcgc cgccacc

57

<210> 6
<211> 117
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polypeptide"

<400> 6
Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr
115

<210> 7
<211> 351
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 7
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgagag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta c 351

<210> 8
<211> 977
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polypeptide"

<400> 8
Ile Phe Ala Ser Met Gln Lys Pro Glu Gly Leu Pro His Ile Ser Asp
1 5 10 15

Val Val Leu Asp Lys Ala Asn Lys Thr Pro Leu Arg Pro Leu Asp Pro
20 25 30

Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe
35 40 45

Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly
50 55 60

Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg
65 70 75 80

Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr
85 90 95

Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp
100 105 110

Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu
115 120 125

Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser
130 135 140

Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys Gln
145 150 155 160

Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile
165 170 175

Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala
180 185 190

Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr
195 200 205

Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg
210 215 220

Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe Glu
225 230 235 240

Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly
245 250 255

Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg
260 265 270

Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met
275 280 285

Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu Lys Asn
290 295 300

His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile
305 310 315 320

Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr
325 330 335

Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg
340 345 350

Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp
355 360 365

Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His
370 375 380

Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln
385 390 395 400

Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro
405 410 415

Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr
420 425 430

Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr
435 440 445

Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala
450 455 460

Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly
465 470 475 480

Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser
485 490 495

Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu Ser
500 505 510

Thr Asp Asn Arg Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu Arg
515 520 525

Gly Glu Lys Ala Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu Arg
530 535 540

Gly Gly Asp Arg Gly Asp Arg Leu Asp Arg Ala Arg Thr Pro Ala Thr
545 550 555 560

Lys Arg Ser Phe Ser Lys Glu Val Glu Glu Arg Ser Arg Glu Arg Pro
565 570 575

Ser Gln Pro Glu Gly Leu Arg Lys Ala Ala Ser Leu Thr Glu Asp Arg
580 585 590

Asp Arg Gly Arg Asp Ala Val Lys Arg Glu Ala Ala Leu Pro Pro Val
595 600 605

Ser Pro Leu Lys Ala Ala Leu Ser Glu Glu Glu Leu Glu Lys Lys Ser
610 615 620

Lys Ala Ile Ile Glu Glu Tyr Leu His Leu Asn Asp Met Lys Glu Ala
625 630 635 640

Val Gln Cys Val Gln Glu Leu Ala Ser Pro Ser Leu Leu Phe Ile Phe
645 650 655

Val Arg His Gly Val Glu Ser Thr Leu Glu Arg Ser Ala Ile Ala Arg
660 665 670

Glu His Met Gly Gln Leu Leu His Gln Leu Leu Cys Ala Gly His Leu
675 680 685

Ser Thr Ala Gln Tyr Tyr Gln Gly Leu Tyr Glu Ile Leu Glu Leu Ala
690 695 700

Glu Asp Met Glu Ile Asp Ile Pro His Val Trp Leu Tyr Leu Ala Glu
705 710 715 720

Leu Val Thr Pro Ile Leu Gln Glu Gly Gly Val Pro Met Gly Glu Leu
725 730 735

Phe Arg Glu Ile Thr Lys Pro Leu Arg Pro Leu Gly Lys Ala Ala Ser
740 745 750

Leu Leu Leu Glu Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro Lys
755 760 765

Lys Val Gly Thr Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu Phe
770 775 780

Leu Pro Glu Gly Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys Val
785 790 795 800

Glu Tyr Thr Leu Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala Leu
805 810 815

Pro Ser Glu Glu Leu Asn Arg Gln Leu Glu Lys Leu Leu Lys Glu Gly
820 825 830

Ser Ser Asn Gln Arg Val Phe Asp Trp Ile Glu Ala Asn Leu Ser Glu
835 840 845

Gln Gln Ile Val Ser Asn Thr Leu Val Arg Ala Leu Met Thr Ala Val
850 855 860

Cys Tyr Ser Ala Ile Ile Phe Glu Thr Pro Leu Arg Val Asp Val Ala
865 870 875 880

Val Leu Lys Ala Arg Ala Lys Leu Leu Gln Lys Tyr Leu Cys Asp Glu
885 890 895

Gln Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Leu Val Val Thr
900 905 910

Leu Glu Gln Pro Pro Asn Leu Leu Arg Met Phe Phe Asp Ala Leu Tyr
915 920 925

Asp Glu Asp Val Val Lys Glu Asp Ala Phe Tyr Ser Trp Glu Ser Ser
930 935 940

Lys Asp Pro Ala Glu Gln Gln Gly Lys Gly Val Ala Leu Lys Ser Val
 945 950 955 960

Thr Ala Phe Phe Lys Trp Leu Arg Glu Ala Glu Glu Glu Ser Asp His
 965 970 975

Asn

<210> 9
 <211> 2931
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 9
 atattcgcga gcatgcagaa gcctgaaggt ctgcctcaca tcagcgacgt ggtgctggac 60
 aaggccaaca agaccctct tagacctctg gaccctacca gactgcaggg catcaactgc 120
 ggccctgact tcactccttc tttcgcaaac ctgggcagaa ccaccctgag caccagaggc 180
 cctcctagag gtggtcccgg cggagaactc cccaggggtc ctgccggcct gggccctaga 240
 cgctcccagc aaggtcctag aaaggagcca aggaagatca tcgccaccgt gctgatgacc 300
 gaggacatca agctgaacaa agctgagaag gcctggaagc ctagcagcaa gagaaccgcc 360
 gccgacaagg acagaggcga ggaggacgcc gacggatcca agaccagga cctgttcaga 420
 agagtgagaa gcacacctaa caagctgacc cctcagatgt tccagcagct gatgaagcag 480
 gtgacgcagc tcgccatcga caccgaggag agactgaagg gcgtgatcga cctgatcttt 540
 gagaaggcta tctcagagcc taacttcagc gtggcctacg ccaacatgtg ccgttgccctg 600
 atggcattga aggtgccaac caccgagaag cctactgtga ccgtcaattt ccgtaaactg 660
 ctgctgaacc ggtgccagaa agagttcgag aaggataagg acgacgacga ggtcttcgag 720
 aagaaacaga aagaaatgga cgaggccgcc accgcagagg aaaggggccg attaaaggag 780
 gagctggagg aggccagaga catcgccaga cggcgttctc tgggcaacat caagttcata 840
 ggtgagctgt tcaagctaaa gatgctcacc gaggccataa tgcacgactg cgtggtgaag 900

ctactgaaga accacgacga agaaagcctg gagtgcctgt gcagactgct gaccaccatc	960
ggcaaggacc tggacttcga gaaggcaaag cctcgaatgg accagtactt caaccagatg	1020
gagaagatta tcaaggagaa gaagaccagc agcagaatca gattcatgct gcaggacgta	1080
ctggacctgc gcggaagcaa ctgggtgccca aggagagggg accaaggacc aaagaccatc	1140
gaccagatcc acaaggaagc ggagatggag gagcacagag agcacataaa ggtgcagcag	1200
cttatggcca agggcagcga caagcgaaga ggcgggccgc ccggacctcc tatcagcaga	1260
ggccttcctc tggtagacga cggcggctgg aacaccgtgc ctatctctaa gggctccaga	1320
cctatcgaca cttcccgtct taccaagatc accaagccag gatctattga cagcaacaac	1380
cagctgttcg ccccaggagg aagacttagc tggggcaagg gaagttccgg cggatccggc	1440
gccaagcctt ccgacgccgc cagcgaggct gccagacctg ccaccagcac cttgaaccgc	1500
ttttccgctc tgcagcaagc tgtgccgacc gagagcaccg acaaccggcg agttgtgcag	1560
agaagcagcc tgagcagaga gagggcgag aaagccggcg atagaggtga ccgactggag	1620
agaagcgaga gaggaggtga tagaggcgac cggcttgaca gagccagaac ccctgccaca	1680
aagcgatcgt tcagcaagga ggtcgaagag aggtccaggg agcgccctag ccagccggaa	1740
ggactcagaa aggcagccag tctaacagag gaccgcgaca ggggaagaga cgccgtgaag	1800
agggaggccg cactgcctcc tgtgagccct ctgaaggcgg cactgagcga agaagaactt	1860
gagaagaaga gtaaggcgat catcgaggag tacctgcacc tgaacgacat gaaggaggcc	1920
gtgcagtgcg tgcaagagct cgcgagccca tctactgtgt tcatcttcgt ccggcacggc	1980
gtggagtcca cactggaaag atctgccatt gctagggagc atatgggcca gttgttgcac	2040
caattgcttt gcgccggcca cctaagtact gccagtact atcagggttt atacgagatc	2100
ctcgaactgg ccgaagacat ggagatcgac atccctcacg tgtggctgta cctcgcagaa	2160
ctggtgacct ctatcctgca ggagggcggc gttccaatgg gtgagttggt cagagaaatc	2220
acaaagccac tgcgcccact gggcaaggcg gctagtctcc ttctggagat tctcggcctg	2280
ctctgtaaga gcatgggtcc aaagaaggtg ggcaccctgt ggagggaagc tggactctcc	2340
tggaaggaat tcctccctga gggtcaggac atcggcgcct tcgtggccga acagaaggtt	2400
gagtacaccc tgggagagga atcgggaagc ccgggtcagc gggctctgcc gagtgaggag	2460
ctcaacagac aactcgagaa gctcctcaag gagggatcct caaatcagag agtgttcgac	2520

tggatcgagg ccaatctcag cgagcagcaa atcgtgagca acacgttggt tcgcgctctc	2580
atgacagccg tgtgctactc agccattatc ttcgagacgc ctctccgcgt ggacgtggca	2640
gtgctcaagg cccgcgctaa gctgttacag aagtacctgt gcgacgagca gaaggaactg	2700
caggccctgt acgccctgca agctctggtc gtcacactcg agcagcctcc taacctgctg	2760
agaatgttct tcgacgcatt gtacgacgag gacgttgtga aggaagacgc tttctacagc	2820
tgggagtcta gtaaggatcc tgcggaacaa caaggcaagg gcgttgctct aaagagcgtg	2880
accgccttct tcaagtggct gcgagaggct gaggaagagt cggaccacaa c	2931

<210> 10

<211> 1099

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 10

Met	Ala	Ser	Asn	Phe	Thr	Gln	Phe	Val	Leu	Val	Asp	Asn	Gly	Gly	Thr
1				5					10					15	

Gly	Asp	Val	Thr	Val	Ala	Pro	Ser	Asn	Phe	Ala	Asn	Gly	Ile	Ala	Glu
			20					25					30		

Trp	Ile	Ser	Ser	Asn	Ser	Arg	Ser	Gln	Ala	Tyr	Lys	Val	Thr	Cys	Ser
		35				40						45			

Val	Arg	Gln	Ser	Ser	Ala	Gln	Asn	Arg	Lys	Tyr	Thr	Ile	Lys	Val	Glu
	50					55					60				

Val	Pro	Lys	Gly	Ala	Trp	Arg	Ser	Tyr	Leu	Asn	Met	Glu	Leu	Thr	Ile
65					70					75					80

Pro	Ile	Phe	Ala	Thr	Asn	Ser	Asp	Cys	Glu	Leu	Ile	Val	Lys	Ala	Met
				85					90					95	

Gln	Gly	Leu	Leu	Lys	Asp	Gly	Asn	Pro	Ile	Pro	Ser	Ala	Ile	Ala	Ala
		100						105					110		

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Ile Phe Ala Ser Met Gln
115 120 125

Lys Pro Glu Gly Leu Pro His Ile Ser Asp Val Val Leu Asp Lys Ala
130 135 140

Asn Lys Thr Pro Leu Arg Pro Leu Asp Pro Thr Arg Leu Gln Gly Ile
145 150 155 160

Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly Arg Thr
165 170 175

Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly Glu Leu
180 185 190

Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln Gly Pro
195 200 205

Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr Glu Asp
210 215 220

Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser Lys Arg
225 230 235 240

Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly Ser Lys
245 250 255

Thr Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys Leu Thr
260 265 270

Pro Gln Met Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu Ala Ile
275 280 285

Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe Glu Lys
290 295 300

Ala Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met Cys Arg
305 310 315 320

Cys Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr Val Thr

325

330

335

Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe Glu
 340 345 350

Lys Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys Glu Met
 355 360 365

Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu Glu Leu
 370 375 380

Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn Ile Lys
 385 390 395 400

Phe Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile Met
 405 410 415

His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser Leu
 420 425 430

Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe
 435 440 445

Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys
 450 455 460

Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln
 465 470 475 480

Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg Gly Asp
 485 490 495

Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu Met Glu
 500 505 510

Glu His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys Gly Ser
 515 520 525

Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg Gly Leu
 530 535 540

Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser Lys Gly
545 550 555 560

Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys Pro Gly
565 570 575

Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly Gly Arg Leu Ser
580 585 590

Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala Lys Pro Ser Asp Ala
595 600 605

Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg Phe Ser
610 615 620

Ala Leu Gln Gln Ala Val Pro Thr Glu Ser Thr Asp Asn Arg Arg Val
625 630 635 640

Val Gln Arg Ser Ser Leu Ser Arg Glu Arg Gly Glu Lys Ala Gly Asp
645 650 655

Arg Gly Asp Arg Leu Glu Arg Ser Glu Arg Gly Gly Asp Arg Gly Asp
660 665 670

Arg Leu Asp Arg Ala Arg Thr Pro Ala Thr Lys Arg Ser Phe Ser Lys
675 680 685

Glu Val Glu Glu Arg Ser Arg Glu Arg Pro Ser Gln Pro Glu Gly Leu
690 695 700

Arg Lys Ala Ala Ser Leu Thr Glu Asp Arg Asp Arg Gly Arg Asp Ala
705 710 715 720

Val Lys Arg Glu Ala Ala Leu Pro Pro Val Ser Pro Leu Lys Ala Ala
725 730 735

Leu Ser Glu Glu Glu Leu Glu Lys Lys Ser Lys Ala Ile Ile Glu Glu
740 745 750

Tyr Leu His Leu Asn Asp Met Lys Glu Ala Val Gln Cys Val Gln Glu

755

760

765

Leu Ala Ser Pro Ser Leu Leu Phe Ile Phe Val Arg His Gly Val Glu
 770 775 780

Ser Thr Leu Glu Arg Ser Ala Ile Ala Arg Glu His Met Gly Gln Leu
 785 790 795 800

Leu His Gln Leu Leu Cys Ala Gly His Leu Ser Thr Ala Gln Tyr Tyr
 805 810 815

Gln Gly Leu Tyr Glu Ile Leu Glu Leu Ala Glu Asp Met Glu Ile Asp
 820 825 830

Ile Pro His Val Trp Leu Tyr Leu Ala Glu Leu Val Thr Pro Ile Leu
 835 840 845

Gln Glu Gly Gly Val Pro Met Gly Glu Leu Phe Arg Glu Ile Thr Lys
 850 855 860

Pro Leu Arg Pro Leu Gly Lys Ala Ala Ser Leu Leu Leu Glu Ile Leu
 865 870 875 880

Gly Leu Leu Cys Lys Ser Met Gly Pro Lys Lys Val Gly Thr Leu Trp
 885 890 895

Arg Glu Ala Gly Leu Ser Trp Lys Glu Phe Leu Pro Glu Gly Gln Asp
 900 905 910

Ile Gly Ala Phe Val Ala Glu Gln Lys Val Glu Tyr Thr Leu Gly Glu
 915 920 925

Glu Ser Glu Ala Pro Gly Gln Arg Ala Leu Pro Ser Glu Glu Leu Asn
 930 935 940

Arg Gln Leu Glu Lys Leu Leu Lys Glu Gly Ser Ser Asn Gln Arg Val
 945 950 955 960

Phe Asp Trp Ile Glu Ala Asn Leu Ser Glu Gln Gln Ile Val Ser Asn
 965 970 975

Thr Leu Val Arg Ala Leu Met Thr Ala Val Cys Tyr Ser Ala Ile Ile
980 985 990

Phe Glu Thr Pro Leu Arg Val Asp Val Ala Val Leu Lys Ala Arg Ala
995 1000 1005

Lys Leu Leu Gln Lys Tyr Leu Cys Asp Glu Gln Lys Glu Leu Gln
1010 1015 1020

Ala Leu Tyr Ala Leu Gln Ala Leu Val Val Thr Leu Glu Gln Pro
1025 1030 1035

Pro Asn Leu Leu Arg Met Phe Phe Asp Ala Leu Tyr Asp Glu Asp
1040 1045 1050

Val Val Lys Glu Asp Ala Phe Tyr Ser Trp Glu Ser Ser Lys Asp
1055 1060 1065

Pro Ala Glu Gln Gln Gly Lys Gly Val Ala Leu Lys Ser Val Thr
1070 1075 1080

Ala Phe Phe Lys Trp Leu Arg Glu Ala Glu Glu Glu Ser Asp His
1085 1090 1095

Asn

<210> 11
<211> 3297
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 11
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggcccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180

atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcatat tcgcgagcat gcagaagcct gaaggtctgc ctcacatcag cgacgtggtg	420
ctggacaagg ccaacaagac ccctcttaga cctctggacc ctaccagact gcagggcatc	480
aactgcggcc ctgacttcac tccttctttc gcaaacctgg gcagaaccac cctgagcacc	540
agaggccctc ctagagggtg tcccggcgga gaactccca ggggtcctgc cggcctgggc	600
cctagacgct cccagcaagg tcctagaaag gagccaagga agatcatcgc caccgtgctg	660
atgaccgagg acatcaagct gaacaaagct gagaaggcct ggaagcctag cagcaagaga	720
accgccgccg acaaggacag aggcgaggag gacgccgacg gatccaagac ccaggacctg	780
ttcagaagag tgagaagcat cctcaacaag ctgaccctc agatgttcca gcagctgatg	840
aagcaggtga cgcagctcgc catcgacacc gaggagagac tgaaggcggt gatcgacctg	900
atctttgaga aggctatctc agagcctaac ttcagcgtgg cctacgcaa catgtgccgt	960
tgctgatgg cattgaaggt gccaaaccacc gagaagccta ctgtgaccgt caatttcgt	1020
aaactgctgc tgaaccggtg ccagaaagag ttcgagaagg ataaggacga cgacgaggtc	1080
ttcgagaaga aacagaaaga aatggacgag gccgccaccg cagaggaaag gggccgatta	1140
aaggaggagc tggaggaggc cagagacatc gccagacggc gttctctggg caacatcaag	1200
ttcataggtg agctgttcaa gctaaagatg ctcaccgagg ccataatgca cgactgcgtg	1260
gtgaagctac tgaagaacca cgacgaagaa agcctggagt gcctgtgcag actgctgacc	1320
accatcggca aggacctgga cttcgagaag gcaaagcctc gaatggacca gtacttcaac	1380
cagatggaga agattatcaa ggagaagaag accagcagca gaatcagatt catgctgcag	1440
gacgtactgg acctgcgcgg aagcaactgg gtgccaaagga gaggggacca aggaccaaag	1500
accatcgacc agatccacaa ggaagcggag atggaggagc acagagagca cataaaggtg	1560
cagcagctta tggccaaggg cagcgacaag cgaagaggcg gcccgcccg acctcctatc	1620
agcagaggcc ttcctctggt agacgacggc ggctggaaca ccgtgcctat ctctaagggc	1680
tccagaccta tcgacacttc ccgtcttacc aagatcacca agccaggatc tattgacagc	1740
aacaaccagc tgttcgcccc aggaggaaga cttagctggg gcaagggaag ttccggcgga	1800

tccggcgcca agccttccga cgccgccagc gaggctgcca gacctgccac cagcaccttg	1860
aaccgctttt ccgctctgca gcaagctgtg ccgaccgaga gcaccgacaa ccggcgagtt	1920
gtgcagagaa gcagcctgag cagagagagg ggcgagaaag ccggcgatag aggtgaccga	1980
ctggagagaa gcgagagagg aggtgataga ggcgaccggc ttgacagagc cagaaccctt	2040
gccacaaagc gatcgttcag caaggaggtc gaagagaggt ccaggagcg ccctagccag	2100
ccggaaggac tcagaaaggc agccagtcta acagaggacc gcgacagggg aagagacgcc	2160
gtgaagaggg aggccgcact gcctcctgtg agccctctga aggcggcact gagcgaagaa	2220
gaacttgaga agaagagtaa ggcatcatc gaggagtacc tgcacctgaa cgacatgaag	2280
gaggccgtgc agtgcgtgca agagctcgcg agcccatcac tgctgttcat cttcgtccgg	2340
cacggcgtgg agtccacact ggaaagatct gccattgcta gggagcatat gggccagttg	2400
ttgcaccaat tgctttgcgc cggccaccta agtactgccc agtactatca gggtttatac	2460
gagatcctcg aactggccga agacatggag atcgacatcc ctcacgtgtg gctgtacctc	2520
gcagaactgg tgacccttat cctgcaggag ggcggcgttc caatgggtga gttgttcaga	2580
gaaatcacia agccactgcg cccactgggc aaggcggcta gtctccttct ggagattctc	2640
ggcctgctct gtaagagcat ggggtccaaag aagggtggga ccctgtggag ggaagctgga	2700
ctctcctgga aggaattcct ccctgagggt caggacatcg gcgccttcgt ggccgaacag	2760
aaggttgagt acaccctggg agaggaatcg gaagcgccgg gtcagcgggc tctgccgagt	2820
gaggagctca acagacaact cgagaagctc ctcaaggagg gatcctcaa tcagagagtg	2880
ttcgactgga tcgaggccaa tctcagcgag cagcaaatcg tgagcaacac gttggttcgc	2940
gctctcatga cagccgtgtg ctactcagcc attatcttcg agacgcctct ccgcgtggac	3000
gtggcagtgc tcaaggcccc cgctaagctg ttacagaagt acctgtgcga cgagcagaag	3060
gaactgcagg ccctgtacgc cctgcaagct ctggctgtca cactcgagca gcctcctaac	3120
ctgctgagaa tgttcttcga cgcattgtac gacgaggacg ttgtgaagga agacgctttc	3180
tacagctggg agtctagtaa ggatcctgcg gaacaacaag gcaagggcgt tgctctaaag	3240
agcgtgaccg ctttcttcaa gtggctgcga gaggtgagg aagagtcgga ccacaac	3297

<210> 12

<211> 1720

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 12

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asn Lys Ala Pro Gln Ser
115 120 125

Thr Gly Pro Pro Pro Ala Pro Ser Pro Gly Leu Pro Gln Pro Ala Phe
130 135 140

Pro Pro Gly Gln Thr Ala Pro Val Val Phe Ser Thr Pro Gln Ala Thr
145 150 155 160

Gln Met Asn Thr Pro Ser Gln Pro Arg Gln His Phe Tyr Pro Ser Arg
165 170 175

Ala Gln Pro Pro Ser Ser Ala Ala Ser Arg Val Gln Ser Ala Ala Pro
180 185 190

Ala Arg Pro Gly Pro Ala Ala His Val Tyr Pro Ala Gly Ser Gln Val
195 200 205

Met Met Ile Pro Ser Gln Ile Ser Tyr Pro Ala Ser Gln Gly Ala Tyr
210 215 220

Tyr Ile Pro Gly Gln Gly Arg Ser Thr Tyr Val Val Pro Thr Gln Gln
225 230 235 240

Tyr Pro Val Gln Pro Gly Ala Pro Gly Phe Tyr Pro Gly Ala Ser Pro
245 250 255

Thr Glu Phe Gly Thr Tyr Ala Gly Ala Tyr Tyr Pro Ala Gln Gly Val
260 265 270

Gln Gln Phe Pro Thr Gly Val Ala Pro Thr Pro Val Leu Met Asn Gln
275 280 285

Pro Pro Gln Ile Ala Pro Lys Arg Glu Arg Lys Thr Ile Arg Ile Arg
290 295 300

Asp Pro Asn Gln Gly Gly Lys Asp Ile Thr Glu Glu Ile Met Ser Gly
305 310 315 320

Ala Arg Thr Ala Ser Thr Pro Thr Pro Gln Thr Gly Gly Gly Leu
325 330 335

Glu Pro Gln Ala Asn Gly Glu Thr Pro Gln Val Ala Val Ile Val Arg
340 345 350

Pro Asp Asp Arg Ser Gln Gly Ala Ile Ile Ala Asp Arg Pro Gly Leu
355 360 365

Pro Gly Pro Glu His Ser Pro Ser Glu Ser Gln Pro Ser Ser Pro Ser
370 375 380

Pro Thr Pro Ser Pro Ser Pro Val Leu Glu Pro Gly Ser Glu Pro Asn
385 390 395 400

Leu Ala Val Leu Ser Ile Pro Gly Asp Thr Met Thr Thr Ile Gln Met
405 410 415

Ser Val Glu Glu Ser Thr Pro Ile Ser Arg Glu Thr Gly Glu Pro Tyr
420 425 430

Arg Leu Ser Pro Glu Pro Thr Pro Leu Ala Glu Pro Ile Leu Glu Val
435 440 445

Glu Val Thr Leu Ser Lys Pro Val Pro Glu Ser Glu Phe Ser Ser Ser
450 455 460

Pro Leu Gln Ala Pro Thr Pro Leu Ala Ser His Thr Val Glu Ile His
465 470 475 480

Glu Pro Asn Gly Met Val Pro Ser Glu Asp Leu Glu Pro Glu Val Glu
485 490 495

Ser Ser Pro Glu Leu Ala Pro Pro Pro Ala Cys Pro Ser Glu Ser Pro
500 505 510

Val Pro Ile Ala Pro Thr Ala Gln Pro Glu Glu Leu Leu Asn Gly Ala
515 520 525

Pro Ser Pro Pro Ala Val Asp Leu Ser Pro Val Ser Glu Pro Glu Glu
530 535 540

Gln Ala Lys Glu Val Thr Ala Ser Met Ala Pro Pro Thr Ile Pro Ser
545 550 555 560

Ala Thr Pro Ala Thr Ala Pro Ser Ala Thr Ser Pro Ala Gln Glu Glu
565 570 575

Glu Met Glu Glu Glu Glu Glu Glu Glu Glu Gly Glu Ala Gly Glu Ala
580 585 590

Gly Glu Ala Glu Ser Glu Lys Gly Gly Glu Glu Leu Leu Pro Pro Glu
595 600 605

Ser Thr Pro Ile Pro Ala Asn Leu Ser Gln Asn Leu Glu Ala Ala Ala
610 615 620

Ala Thr Gln Val Ala Val Ser Val Pro Lys Arg Arg Arg Lys Ile Lys
625 630 635 640

Glu Leu Asn Lys Lys Glu Ala Val Gly Asp Leu Leu Asp Ala Phe Lys
645 650 655

Glu Ala Asn Pro Ala Val Pro Glu Val Glu Asn Gln Pro Pro Ala Gly
660 665 670

Ser Asn Pro Gly Pro Glu Ser Glu Gly Ser Gly Val Pro Pro Arg Pro
675 680 685

Glu Glu Ala Asp Glu Thr Trp Asp Ser Lys Glu Asp Lys Ile His Asn
690 695 700

Ala Glu Asn Ile Gln Pro Gly Glu Gln Lys Tyr Glu Tyr Lys Ser Asp
705 710 715 720

Gln Trp Lys Pro Leu Asn Leu Glu Glu Lys Lys Arg Tyr Asp Arg Glu
725 730 735

Phe Leu Leu Gly Phe Gln Phe Ile Phe Ala Ser Met Gln Lys Pro Glu
740 745 750

Gly Leu Pro His Ile Ser Asp Val Val Leu Asp Lys Ala Asn Lys Thr
755 760 765

Pro Leu Arg Pro Leu Asp Pro Thr Arg Leu Gln Gly Ile Asn Cys Gly
770 775 780

Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly Arg Thr Thr Leu Ser
785 790 795 800

Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly Glu Leu Pro Arg Gly
805 810 815

Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu
820 825 830

Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr Glu Asp Ile Lys Leu
835 840 845

Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala
850 855 860

Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp
865 870 875 880

Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys Leu Thr Pro Gln Met
885 890 895

Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu Ala Ile Asp Thr Glu
900 905 910

Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser
915 920 925

Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met Cys Arg Cys Leu Met
930 935 940

Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr Val Thr Val Asn Phe
945 950 955 960

Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys
965 970 975

Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys Glu Met Asp Glu Ala
980 985 990

Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala
995 1000 1005

Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn Ile Lys Phe Ile
1010 1015 1020

Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile Met His
1025 1030 1035

Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser Leu
1040 1045 1050

Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp
1055 1060 1065

Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met
1070 1075 1080

Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe
1085 1090 1095

Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro
1100 1105 1110

Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys
1115 1120 1125

Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln
1130 1135 1140

Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly
1145 1150 1155

Pro Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp
1160 1165 1170

Asn Thr Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser
1175 1180 1185

Arg Leu Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn
1190 1195 1200

Gln Leu Phe Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser
1205 1210 1215

Ser Gly Gly Ser Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala
1220 1225 1230

Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln
1235 1240 1245

Gln	Ala	Val	Pro	Thr	Glu	Ser	Thr	Asp	Asn	Arg	Arg	Val	Val	Gln
1250						1255					1260			
Arg	Ser	Ser	Leu	Ser	Arg	Glu	Arg	Gly	Glu	Lys	Ala	Gly	Asp	Arg
1265						1270					1275			
Gly	Asp	Arg	Leu	Glu	Arg	Ser	Glu	Arg	Gly	Gly	Asp	Arg	Gly	Asp
1280						1285					1290			
Arg	Leu	Asp	Arg	Ala	Arg	Thr	Pro	Ala	Thr	Lys	Arg	Ser	Phe	Ser
1295						1300					1305			
Lys	Glu	Val	Glu	Glu	Arg	Ser	Arg	Glu	Arg	Pro	Ser	Gln	Pro	Glu
1310						1315					1320			
Gly	Leu	Arg	Lys	Ala	Ala	Ser	Leu	Thr	Glu	Asp	Arg	Asp	Arg	Gly
1325						1330					1335			
Arg	Asp	Ala	Val	Lys	Arg	Glu	Ala	Ala	Leu	Pro	Pro	Val	Ser	Pro
1340						1345					1350			
Leu	Lys	Ala	Ala	Leu	Ser	Glu	Glu	Glu	Leu	Glu	Lys	Lys	Ser	Lys
1355						1360					1365			
Ala	Ile	Ile	Glu	Glu	Tyr	Leu	His	Leu	Asn	Asp	Met	Lys	Glu	Ala
1370						1375					1380			
Val	Gln	Cys	Val	Gln	Glu	Leu	Ala	Ser	Pro	Ser	Leu	Leu	Phe	Ile
1385						1390					1395			
Phe	Val	Arg	His	Gly	Val	Glu	Ser	Thr	Leu	Glu	Arg	Ser	Ala	Ile
1400						1405					1410			
Ala	Arg	Glu	His	Met	Gly	Gln	Leu	Leu	His	Gln	Leu	Leu	Cys	Ala
1415						1420					1425			
Gly	His	Leu	Ser	Thr	Ala	Gln	Tyr	Tyr	Gln	Gly	Leu	Tyr	Glu	Ile
1430						1435					1440			
Leu	Glu	Leu	Ala	Glu	Asp	Met	Glu	Ile	Asp	Ile	Pro	His	Val	Trp
1445						1450					1455			

Leu Tyr Leu Ala Glu Leu Val Thr Pro Ile Leu Gln Glu Gly Gly
1460 1465 1470

Val Pro Met Gly Glu Leu Phe Arg Glu Ile Thr Lys Pro Leu Arg
1475 1480 1485

Pro Leu Gly Lys Ala Ala Ser Leu Leu Leu Glu Ile Leu Gly Leu
1490 1495 1500

Leu Cys Lys Ser Met Gly Pro Lys Lys Val Gly Thr Leu Trp Arg
1505 1510 1515

Glu Ala Gly Leu Ser Trp Lys Glu Phe Leu Pro Glu Gly Gln Asp
1520 1525 1530

Ile Gly Ala Phe Val Ala Glu Gln Lys Val Glu Tyr Thr Leu Gly
1535 1540 1545

Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala Leu Pro Ser Glu Glu
1550 1555 1560

Leu Asn Arg Gln Leu Glu Lys Leu Leu Lys Glu Gly Ser Ser Asn
1565 1570 1575

Gln Arg Val Phe Asp Trp Ile Glu Ala Asn Leu Ser Glu Gln Gln
1580 1585 1590

Ile Val Ser Asn Thr Leu Val Arg Ala Leu Met Thr Ala Val Cys
1595 1600 1605

Tyr Ser Ala Ile Ile Phe Glu Thr Pro Leu Arg Val Asp Val Ala
1610 1615 1620

Val Leu Lys Ala Arg Ala Lys Leu Leu Gln Lys Tyr Leu Cys Asp
1625 1630 1635

Glu Gln Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Leu Val
1640 1645 1650

Val Thr Leu Glu Gln Pro Pro Asn Leu Leu Arg Met Phe Phe Asp
 1655 1660 1665

Ala Leu Tyr Asp Glu Asp Val Val Lys Glu Asp Ala Phe Tyr Ser
 1670 1675 1680

Trp Glu Ser Ser Lys Asp Pro Ala Glu Gln Gln Gly Lys Gly Val
 1685 1690 1695

Ala Leu Lys Ser Val Thr Ala Phe Phe Lys Trp Leu Arg Glu Ala
 1700 1705 1710

Glu Glu Glu Ser Asp His Asn
 1715 1720

<210> 13
 <211> 5160
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 13
 atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
 gtggcccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
 caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
 atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
 cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
 aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt 360
 ggaagcaaca agggccctca gagcaccggc cctcctcctg cccctagccc tggcctgcct 420
 cagcctgcct tccctcctgg ccagaccgcc cctgtggtgt tcagcacccc tcaggccacc 480
 cagatgaaca cacctagcca gcctagacag cacttctacc ctagcagagc ccagcctcct 540
 agcagcgccg ccagcagagt gcagtccgct gcacctgcca gacctggccc tgccgcccac 600
 gtgtaccctg ccggcagcca ggtgatgatg atcccagatc aaatcagcta ccctgcaagc 660
 cagggcgcct actacatccc gggccagggc agaagcacct acgtggtgcc taccacagcag 720

taccctgtgc agcctggcgc ccctggtttc tatcctggcg caagccctac cgagttcggg	780
acttacgccg gcgcttacta tccagctcag ggcgtgcagc agttccctac cggcgtggcc	840
cctacccctg tgctgatgaa tcagccacct cagatcgccc ctaagcgcg	900
atcagaatcc gcgatacctaa ccagggcggc aaggacatca ccgaggagat catgagcggc	960
gccagaacag caagtactcc aaccccgcca caaacggcg gcggcctgga gcctcaagcc	1020
aacggcgaga cgccacaagt ggccgtgatc gtacgccctg acgaccggag ccaaggtgca	1080
atcatcgccg ataggcctgg cctcccaggt ccggagcaca gccctagcga gtcccagccg	1140
tcttcaccat caccaacccc tagtccatcc cctgtttctcg aaccaggcag cgagcctaac	1200
ctggccgtgc tgagcatacc aggtgacacc atgaccacca tccagatgag cgtggaggag	1260
agcaccctaa tcagcagaga aactggagag ccttacagac tgtccccaga gccgacccca	1320
ctggccgagc caatactgga ggtggagggtg accctgagca agcctgtgcc tgagagcgag	1380
ttcagctcct ctccactgca ggccccaact cctctcgcaa gccacaccgt ggagatccac	1440
gaacctaacg gcatggtacc aagcgaagat cttgagccag aggtcgaatc aagcccagaa	1500
ctggccctc cacctgcctg cccgtctgaa tctccggctc ctatcgctcc tacggcacag	1560
cctgaggagc tgctgaacgg tgccccgagc cctccagcag tggacttatc cccagtatca	1620
gagcctgaag aacaggccaa ggaggtaact gcctctatgg cgccacctac cataccttcg	1680
gcaacaccgg ctacagcacc atctgcgact agtccggctc aggaggagga gatggaggaa	1740
gaagaagagg aagaggaggg cgaggccgga gaggccggtg aagccgagtc cgagaagggc	1800
ggcgaagaac ttctccctcc agagtcaact cctatccctg ccaaccttag tcagaatctg	1860
gaggccgccg ccgctactca ggttgacgtg agcgtgccaa agagacgtcg caagatcaag	1920
gagctgaaca agaaggaggc cgtgggagac ctgctggacg ccttcaagga ggcaaaccg	1980
gcggtgcctg aagtggagaa tcagcctccg gccggatcaa accctggtcc tgagagtga	2040
ggcagcggcg tcccaccaag acctgaagag gctgacgaaa cttgggacag caaggaggac	2100
aagatccaca acgccgagaa catccagcca ggcgagcaga agtacgagta caagagcgac	2160
cagtggaagc ctcttaacct tgaagagaag aagagatacg acagagagtt cctgctgggc	2220
ttccagttca tattcgcgag catgcagaag cctgaaggtc tgcctcacat cagcgacgtg	2280
gtgctggaca aggccaacaa gaccctctt agacctctgg accctaccag actgcagggc	2340

atcaactgcg gccctgactt cactccttct ttcgcaaacc tgggcagaac caccctgagc	2400
accagaggcc ctctagagg tgggtcccggc ggagaactcc ccaggggtcc tgccggcctg	2460
ggccctagac gctcccagca aggtcctaga aaggagccaa ggaagatcat cgccaccgtg	2520
ctgatgaccg aggacatcaa gctgaacaaa gctgagaagg cctggaagcc tagcagcaag	2580
agaaccgccg ccgacaagga cagaggcgag gaggacgccg acggatccaa gacccaggac	2640
ctgttcagaa gagtgagaag catcctcaac aagctgaccc ctcatagtgt ccagcagctg	2700
atgaagcagg tgacgcagct cgccatcgac accgaggaga gactgaaggg cgtgatcgac	2760
ctgatctttg agaaggctat ctcatagcct aacttcagcg tggcctacgc caacatgtgc	2820
cgttgccctga tggcattgaa ggtgcccaacc accgagaagc ctactgtgac cgtcaatttc	2880
cgtaaactgc tgctgaaccg gtgccagaaa gagttcgaga aggataagga cgacgacgag	2940
gtcttcgaga agaaacagaa agaaatggac gaggccgcca ccgcagagga aaggggccga	3000
ttaaaggagg agctggagga ggccagagac atcgccagac ggcgttctct gggcaacatc	3060
aagttcatag gtgagctgtt caagctaaag atgctcaccg aggccataat gcacgactgc	3120
gtggtgaagc tactgaagaa ccacgacgaa gaaagcctgg agtgcctgtg cagactgctg	3180
accaccatcg gcaaggacct ggacttcgag aaggcaaagc ctgcaatgga ccagtacttc	3240
aaccagatgg agaagattat caaggagaag aagaccagca gcagaatcag attcatgctg	3300
caggacgtac tggacctgcg cggaagcaac tgggtgccaa ggagagggga ccaaggacca	3360
aagaccatcg accagatcca caaggaagcg gagatggagg agcacagaga gcacataaag	3420
gtgcagcagc ttatggccaa gggcagcgac aagcgaagag gcggcccgcc cggacctcct	3480
atcagcagag gccttcctct ggtagacgac ggcggctgga acaccgtgcc tatctctaag	3540
ggctccagac ctatcgacac ttcccgtctt accaagatca ccaagccagg atctattgac	3600
agcaacaacc agctgttcgc cccaggagga agacttagct ggggcaaggg aagttccggc	3660
ggatccggcg ccaagccttc cgacgccgcc agcgaggctg ccagacctgc caccagcacc	3720
ttgaaccgct tttccgctct gcagcaagct gtgccgaccg agagcaccga caaccggcga	3780
gttgtgcaga gaagcagcct gagcagagag aggggcgaga aagccggcga tagagggtgac	3840
cgactggaga gaagcgagag aggaggtgat agaggcgacc ggcttgacag agccagaacc	3900
cctgccacaa agcgatcggt cagcaaggag gtcgaagaga ggtccaggga gcgccctagc	3960

cagccggaag gactcagaaa ggcagccagt ctaacagagg accgcgacag gggaagagac	4020
gccgtgaaga gggaggccgc actgcctcct gtgagccctc tgaaggcggc actgagcgaa	4080
gaagaacttg agaagaagag taaggcgatc atcgaggagt acctgcacct gaacgacatg	4140
aaggaggccg tgcagtgcgt gcaagagctc gcgagcccat cactgctgtt catcttcgtc	4200
cggcacggcg tggagtccac actggaaaga tctgccattg ctagggagca tatgggccag	4260
ttgttgacc aattgctttg cgccggccac ctaagtactg cccagtacta tcagggttta	4320
tacgagatcc tcgaactggc cgaagacatg gagatcgaca tccctcacgt gtggctgtac	4380
ctcgcagaac tggtgacccc tctcctgcag gagggcggcg ttccaatggg tgagttgttc	4440
agagaaatca caaagccact gcgcccactg ggcaaggcgg ctagtctcct tctggagatt	4500
ctcggcctgc tctgtaagag catgggtcca aagaaggtgg gcaccctgtg gagggaagct	4560
ggactctcct ggaaggaatt cctccctgag ggtcaggaca tcggcgcctt cgtggccgaa	4620
cagaaggttg agtacaccct gggagaggaa tcggaagcgc cgggtcagcg ggctctgccg	4680
agtgaggagc tcaacagaca actcgagaag ctctcaagg agggatcctc aaatcagaga	4740
gtgttcgact ggatcgaggc caatctcagc gagcagcaaa tcgtgagcaa cacgttggtt	4800
cgcgctctca tgacagccgt gtgctactca gccattatct tcgagacgcc tctccgcgtg	4860
gacgtggcag tgctcaaggc ccgcgctaag ctgttacaga agtacctgtg cgacgagcag	4920
aaggaactgc aggccctgta cgccctgcaa gctctggtcg tcacactcga gcagcctcct	4980
aacctgctga gaatgttctt cgacgcattg tacgacgagg acgttggtgaa ggaagacgct	5040
ttctacagct gggagtctag taaggatcct gcggaacaac aaggcaaggg cgttgctcta	5100
aagagcgtga ccgccttctt caagtggctg cgagaggctg aggaagagtc ggaccacaac	5160

<210> 14

<211> 1099

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 14

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1	5	10	15
Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu	20	25	30
Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser	35	40	45
Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu	50	55	60
Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile	65	70	75
Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met	85	90	95
Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala	100	105	110
Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Ile Phe Ala Ser Met Gln	115	120	125
Lys Pro Glu Gly Leu Pro His Ile Ser Asp Val Val Leu Asp Lys Ala	130	135	140
Asn Lys Thr Pro Leu Arg Pro Leu Asp Pro Thr Arg Leu Gln Gly Ile	145	150	155
Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly Arg Thr	165	170	175
Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly Glu Leu	180	185	190
Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln Gly Pro	195	200	205
Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr Glu Asp	210	215	220

Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser Lys Arg
225 230 235 240

Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly Ser Lys
245 250 255

Thr Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Ala Asn Lys Ala Thr
260 265 270

Pro Gln Met Ala Gln Gln Leu Met Lys Gln Val Thr Gln Leu Ala Ile
275 280 285

Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe Glu Lys
290 295 300

Ala Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met Cys Arg
305 310 315 320

Cys Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr Val Thr
325 330 335

Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe Glu
340 345 350

Lys Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys Glu Met
355 360 365

Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu Glu Leu
370 375 380

Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn Ile Lys
385 390 395 400

Phe Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile Met
405 410 415

His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser Leu
420 425 430

Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe

435

440

445

Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys
 450 455 460

Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln
 465 470 475 480

Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg Gly Asp
 485 490 495

Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu Met Glu
 500 505 510

Glu His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys Gly Ser
 515 520 525

Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg Gly Leu
 530 535 540

Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser Lys Gly
 545 550 555 560

Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys Pro Gly
 565 570 575

Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly Gly Arg Leu Ser
 580 585 590

Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala Lys Pro Ser Asp Ala
 595 600 605

Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg Phe Ser
 610 615 620

Ala Leu Gln Gln Ala Val Pro Thr Glu Ser Thr Asp Asn Arg Arg Val
 625 630 635 640

Val Gln Arg Ser Ser Leu Ser Arg Glu Arg Gly Glu Lys Ala Gly Asp
 645 650 655

Arg Gly Asp Arg Leu Glu Arg Ser Glu Arg Gly Gly Asp Arg Gly Asp
660 665 670

Arg Leu Asp Arg Ala Arg Thr Pro Ala Thr Lys Arg Ser Phe Ser Lys
675 680 685

Glu Val Glu Glu Arg Ser Arg Glu Arg Pro Ser Gln Pro Glu Gly Leu
690 695 700

Arg Lys Ala Ala Ser Leu Thr Glu Asp Arg Asp Arg Gly Arg Asp Ala
705 710 715 720

Val Lys Arg Glu Ala Ala Leu Pro Pro Val Ser Pro Leu Lys Ala Ala
725 730 735

Leu Ser Glu Glu Glu Leu Glu Lys Lys Ser Lys Ala Ile Ile Glu Glu
740 745 750

Tyr Leu His Leu Asn Asp Met Lys Glu Ala Val Gln Cys Val Gln Glu
755 760 765

Leu Ala Ser Pro Ser Leu Leu Phe Ile Phe Val Arg His Gly Val Glu
770 775 780

Ser Thr Leu Glu Arg Ser Ala Ile Ala Arg Glu His Met Gly Gln Leu
785 790 795 800

Leu His Gln Leu Leu Cys Ala Gly His Leu Ser Thr Ala Gln Tyr Tyr
805 810 815

Gln Gly Leu Tyr Glu Ile Leu Glu Leu Ala Glu Asp Met Glu Ile Asp
820 825 830

Ile Pro His Val Trp Leu Tyr Leu Ala Glu Leu Val Thr Pro Ile Leu
835 840 845

Gln Glu Gly Gly Val Pro Met Gly Glu Leu Phe Arg Glu Ile Thr Lys
850 855 860

Pro Leu Arg Pro Leu Gly Lys Ala Ala Ser Leu Leu Leu Glu Ile Leu

865		870		875		880									
Gly	Leu	Leu	Cys	Lys	Ser	Met	Gly	Pro	Lys	Lys	Val	Gly	Thr	Leu	Trp
				885					890					895	
Arg	Glu	Ala	Gly	Leu	Ser	Trp	Lys	Glu	Phe	Leu	Pro	Glu	Gly	Gln	Asp
			900					905					910		
Ile	Gly	Ala	Phe	Val	Ala	Glu	Gln	Lys	Val	Glu	Tyr	Thr	Leu	Gly	Glu
		915					920					925			
Glu	Ser	Glu	Ala	Pro	Gly	Gln	Arg	Ala	Leu	Pro	Ser	Glu	Glu	Leu	Asn
	930					935					940				
Arg	Gln	Leu	Glu	Lys	Leu	Leu	Lys	Glu	Gly	Ser	Ser	Asn	Gln	Arg	Val
945					950					955					960
Phe	Asp	Trp	Ile	Glu	Ala	Asn	Leu	Ser	Glu	Gln	Gln	Ile	Val	Ser	Asn
				965					970					975	
Thr	Leu	Val	Arg	Ala	Leu	Met	Thr	Ala	Val	Cys	Tyr	Ser	Ala	Ile	Ile
			980					985					990		
Phe	Glu	Thr	Pro	Leu	Arg	Val	Asp	Val	Ala	Val	Leu	Lys	Ala	Arg	Ala
	995						1000					1005			
Lys	Leu	Leu	Gln	Lys	Tyr	Leu	Cys	Asp	Glu	Gln	Lys	Glu	Leu	Gln	
	1010					1015					1020				
Ala	Leu	Tyr	Ala	Leu	Gln	Ala	Leu	Val	Val	Thr	Leu	Glu	Gln	Pro	
	1025					1030					1035				
Pro	Asn	Leu	Leu	Arg	Met	Phe	Phe	Asp	Ala	Leu	Tyr	Asp	Glu	Asp	
	1040					1045					1050				
Val	Val	Lys	Glu	Asp	Ala	Phe	Tyr	Ser	Trp	Glu	Ser	Ser	Lys	Asp	
	1055					1060					1065				
Pro	Ala	Glu	Gln	Gln	Gly	Lys	Gly	Val	Ala	Leu	Lys	Ser	Val	Thr	
	1070					1075					1080				

Ala	Phe	Phe	Lys	Trp	Leu	Arg	Glu	Ala	Glu	Glu	Glu	Ser	Asp	His
1085						1090						1095		

Asn

<210> 15
 <211> 3297
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 15	
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcatat tcgcgagcat gcagaagcct gaaggtctgc ctacatcag cgacgtgggtg	420
ctggacaagg ccaacaagac ccctcttaga cctctggacc ctaccagact gcagggcatc	480
aactgcggcc ctgacttcac tccttctttc gcaaacctgg gcagaaccac cctgagcacc	540
agaggccctc ctagaggtgg tcccggcgga gaactcccca ggggtcctgc cggcctgggc	600
cctagacgct cccagcaagg tcctagaaag gagccaagga agatcatcgc caccgtgctg	660
atgaccgagg acatcaagct gaacaaagct gagaaggcct ggaagcctag cagcaagaga	720
accgccgccg acaaggacag aggcgaggag gacgccgacg gatccaagac ccaggacctg	780
ttcagaagag tgagaagcat cggaacaag gcgaccctc agatggcgca gcagctgatg	840
aagcaggtga cgcagctcgc catcgacacc gaggagagac tgaagggcgt gatcgacctg	900
atctttgaga aggctatctc agagcctaac ttcagcgtgg cctacgcaa catgtgccgt	960
tgcttgatgg cattgaaggt gccaaaccacc gagaagccta ctgtgaccgt caatttcgt	1020

aaactgctgc tgaaccggtg ccagaaagag ttcgagaagg ataaggacga cgacgaggtc	1080
ttcgagaaga aacagaaaga aatggacgag gccgccaccg cagaggaaag gggccgatta	1140
aaggaggagc tggaggaggc cagagacatc gccagacggc gttctctggg caacatcaag	1200
ttcatagggtg agctgttcaa gctaaagatg ctcaccgagg ccataatgca cgactgcgtg	1260
gtgaagctac tgaagaacca cgacgaagaa agcctggagt gcctgtgcag actgctgacc	1320
accatcggca aggacctgga cttcgagaag gcaaagcctc gaatggacca gtacttcaac	1380
cagatggaga agattatcaa ggagaagaag accagcagca gaatcagatt catgctgcag	1440
gacgtactgg acctgcgcgg aagcaactgg gtgccaagga gaggggacca aggaccaaag	1500
accatcgacc agatccacaa ggaagcggag atggaggagc acagagagca cataaaggtg	1560
cagcagctta tggccaaggg cagcgacaag cgaagaggcg gcccgcccg acctcctatc	1620
agcagaggcc ttcctctggt agacgacggc ggctggaaca ccgtgcctat ctctaagggc	1680
tccagaccta tcgacacttc ccgtcttacc aagatcacca agccaggatc tattgacagc	1740
aacaaccagc tgttcgcccc aggaggaaga cttagctggg gcaagggaag ttccggcgga	1800
tccggcgcca agccttccga cgccgccagc gaggctgcca gacctgccac cagcaccttg	1860
aaccgctttt ccgctctgca gcaagctgtg ccgaccgaga gcaccgacaa ccggcgagtt	1920
gtgcagagaa gcagcctgag cagagagagg ggcgagaaag ccggcgatag aggtgaccga	1980
ctggagagaa gcgagagagg aggtgataga ggcgaccggc ttgacagagc cagaaccctt	2040
gccacaaagc gatcgttcag caaggaggtc gaagagaggt ccagggagcg ccctagccag	2100
ccggaaggac tcagaaaggc agccagtcta acagaggacc gcgacagggg aagagacgcc	2160
gtgaagaggg aggccgcact gcctcctgtg agccctctga aggcggcact gagcgaagaa	2220
gaacttgaga agaagagtaa ggcgatcatc gaggagtacc tgcacctgaa cgacatgaag	2280
gaggccgtgc agtgcggtgca agagctcgcg agcccatcac tgctgttcat cttcgtccgg	2340
cacggcgtgg agtccacact ggaaagatct gccattgcta gggagcatat gggccagttg	2400
ttgcaccaat tgctttgcgc cggccaccta agtactgccc agtactatca gggtttatac	2460
gagatcctcg aactggccga agacatggag atcgacatcc ctcacgtgtg gctgtacctc	2520
gcagaactgg tgacccttat cctgcaggag ggcggcgttc caatgggtga gttgttcaga	2580
gaaatcacia agccactgcg cccactgggc aaggcggcta gtctccttct ggagattctc	2640

```

ggcctgctct gtaagagcat gggtcctaaag aaggtgggca ccctgtggag ggaagctgga      2700
ctctcctgga aggaattcct ccctgagggt caggacatcg ggccttcgt ggccgaacag      2760
aaggttgagt acaccctggg agaggaatcg gaagcgccgg gtcagcgggc tctgccgagt      2820
gaggagctca acagacaact cgagaagctc ctcaaggagg gatcctcaaa tcagagagtg      2880
ttcgactgga tcgaggccaa tctcagcgag cagcaaatcg tgagcaaacac gttggttcgc      2940
gctctcatga cagccgtgtg ctactcagcc attatcttcg agacgcctct ccgcgtggac      3000
gtggcagtg ctaaggcccc cgctaagctg ttacagaagt acctgtgcga cgagcagaag      3060
gaactgcagg ccctgtacgc cctgcaagct ctggtcgtca cactcgagca gcctcctaac      3120
ctgctgagaa tgttcttcga cgcattgtac gacgaggacg ttgtgaagga agacgctttc      3180
tacagctggg agtctagtaa ggatcctgcg gaacaacaag gcaagggcgt tgctctaaag      3240
agcgtgaccg ctttcttcaa gtggctgcga gaggctgagg aagagtcgga ccacaac      3297

```

<210> 16

<211> 1069

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 16

```

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1           5           10          15

```

```

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
          20          25          30

```

```

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
          35          40          45

```

```

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
          50          55          60

```

```

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65          70          75          80

```

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asp Pro Thr Arg Leu Gln
115 120 125

Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly
130 135 140

Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly
145 150 155 160

Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln
165 170 175

Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr
180 185 190

Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser
195 200 205

Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly
210 215 220

Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys
225 230 235 240

Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu
245 250 255

Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe
260 265 270

Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met
275 280 285

Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr
290 295 300

Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu
305 310 315 320

Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys
325 330 335

Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu
340 345 350

Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn
355 360 365

Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala
370 375 380

Ile Met His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu
385 390 395 400

Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu
405 410 415

Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met
420 425 430

Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met
435 440 445

Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg
450 455 460

Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu
465 470 475 480

Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys
485 490 495

Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg
500 505 510

Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser
515 520 525

Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys
530 535 540

Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly Gly Arg
545 550 555 560

Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala Lys Pro Ser
565 570 575

Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg
580 585 590

Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu Ser Thr Asp Asn Arg
595 600 605

Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu Arg Gly Glu Lys Ala
610 615 620

Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu Arg Gly Gly Asp Arg
625 630 635 640

Gly Asp Arg Leu Asp Arg Ala Arg Thr Pro Ala Thr Lys Arg Ser Phe
645 650 655

Ser Lys Glu Val Glu Glu Arg Ser Arg Glu Arg Pro Ser Gln Pro Glu
660 665 670

Gly Leu Arg Lys Ala Ala Ser Leu Thr Glu Asp Arg Asp Arg Gly Arg
675 680 685

Asp Ala Val Lys Arg Glu Ala Ala Leu Pro Pro Val Ser Pro Leu Lys
690 695 700

Ala Ala Leu Ser Glu Glu Glu Leu Glu Lys Lys Ser Lys Ala Ile Ile
705 710 715 720

Glu Glu Tyr Leu His Leu Asn Asp Met Lys Glu Ala Val Gln Cys Val
725 730 735

Gln Glu Leu Ala Ser Pro Ser Leu Leu Phe Ile Phe Val Arg His Gly
740 745 750

Val Glu Ser Thr Leu Glu Arg Ser Ala Ile Ala Arg Glu His Met Gly
755 760 765

Gln Leu Leu His Gln Leu Leu Cys Ala Gly His Leu Ser Thr Ala Gln
770 775 780

Tyr Tyr Gln Gly Leu Tyr Glu Ile Leu Glu Leu Ala Glu Asp Met Glu
785 790 795 800

Ile Asp Ile Pro His Val Trp Leu Tyr Leu Ala Glu Leu Val Thr Pro
805 810 815

Ile Leu Gln Glu Gly Gly Val Pro Met Gly Glu Leu Phe Arg Glu Ile
820 825 830

Thr Lys Pro Leu Arg Pro Leu Gly Lys Ala Ala Ser Leu Leu Leu Glu
835 840 845

Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro Lys Lys Val Gly Thr
850 855 860

Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu Phe Leu Pro Glu Gly
865 870 875 880

Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys Val Glu Tyr Thr Leu
885 890 895

Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala Leu Pro Ser Glu Glu
900 905 910

Leu Asn Arg Gln Leu Glu Lys Leu Leu Lys Glu Gly Ser Ser Asn Gln
915 920 925

Arg Val Phe Asp Trp Ile Glu Ala Asn Leu Ser Glu Gln Gln Ile Val
930 935 940

Ser Asn Thr Leu Val Arg Ala Leu Met Thr Ala Val Cys Tyr Ser Ala
945 950 955 960

Ile Ile Phe Glu Thr Pro Leu Arg Val Asp Val Ala Val Leu Lys Ala
965 970 975

Arg Ala Lys Leu Leu Gln Lys Tyr Leu Cys Asp Glu Gln Lys Glu Leu
980 985 990

Gln Ala Leu Tyr Ala Leu Gln Ala Leu Val Val Thr Leu Glu Gln Pro
995 1000 1005

Pro Asn Leu Leu Arg Met Phe Phe Asp Ala Leu Tyr Asp Glu Asp
1010 1015 1020

Val Val Lys Glu Asp Ala Phe Tyr Ser Trp Glu Ser Ser Lys Asp
1025 1030 1035

Pro Ala Glu Gln Gln Gly Lys Gly Val Ala Leu Lys Ser Val Thr
1040 1045 1050

Ala Phe Phe Lys Trp Leu Arg Glu Ala Glu Glu Glu Ser Asp His
1055 1060 1065

Asn

<210> 17
<211> 3207
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 17
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggcccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240

cctatcttcg ccaccaacag cgactgcgag ctgacgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggg	360
ggaagcgacc ctaccagact gcagggcac aactgcggcc ctgacttcac tccttctttc	420
gcaaacctgg gcagaaccac cctgagcacc agaggccctc ctagagggtg tcccggcgga	480
gaactcccca ggggtcctgc cggcctgggc cctagacgct cccagcaagg tcctagaaag	540
gagccaagga agatcatcgc caccgtgctg atgaccgagg acatcaagct gaacaaagct	600
gagaaggcct ggaagcctag cagcaagaga accgccgccg acaaggacag aggcgaggag	660
gacgccgacg gatccaagac ccaggacctg ttcagaagag tgagaagcat cctcaacaag	720
ctgaccctc agatgttcca gcagctgatg aagcagggtga cgcagctcgc catcgacacc	780
gaggagagac tgaagggcgt gatcgacctg atctttgaga aggctatctc agagcctaac	840
ttcagcgtgg cctacgcaa catgtgccgt tgcctgatgg cattgaaggt gccaacacc	900
gagaagccta ctgtgaccgt caatttccgt aaactgctgc tgaaccggtg ccagaaagag	960
ttcgagaagg ataaggacga cgacgaggtc ttcgagaaga aacagaaaga aatggacgag	1020
gccgccaccg cagaggaaag gggccgatta aaggaggagc tggaggaggc cagagacatc	1080
gccagacggc gttctctggg caacatcaag ttcataggtg agctgttcaa gctaaagatg	1140
ctcaccgagg ccataatgca cgactgcgtg gtgaagctac tgaagaacca cgacgaagaa	1200
agcctggagt gcctgtgcag actgctgacc accatcggca aggacctgga cttcgagaag	1260
gcaaagcctc gaatggacca gtacttcaac cagatggaga agattatcaa ggagaagaag	1320
accagcagca gaatcagatt catgctgcag gacgtactgg acctgcgcgg aagcaactgg	1380
gtgccaagga gaggggacca aggaccaaag accatcgacc agatccacaa ggaagcggag	1440
atggaggagc acagagagca cataaagggtg cagcagctta tggccaaggg cagcgacaag	1500
cgaagaggcg gcccggccgg acctcctatc agcagaggcc ttcctctggt agacgacggc	1560
ggctggaaca ccgtgcctat ctctaagggc tccagaccta tcgacacttc ccgtcttacc	1620
aagatcacca agccaggatc tattgacagc aacaaccagc tgttcgcccc aggaggaaga	1680
cttagctggg gcaagggaag ttccggcgga tccggcgcca agccttccga cgccgccagc	1740
gaggctgcca gacctgccac cagcaccttg aaccgctttt ccgctctgca gcaagctgtg	1800
ccgaccgaga gcaccgacaa ccggcgagtt gtgcagagaa gcagcctgag cagagagagg	1860

ggcgagaaag ccggcgatag aggtgaccga ctggagagaa gcgagagagg aggtgataga	1920
ggcgaccggc ttgacagagc cagaaccct gccacaaagc gatcggttcag caaggaggtc	1980
gaagagaggt ccaggagcg ccctagccag ccggaaggac tcagaaaggc agccagtcta	2040
acagaggacc gcgacagggg aagagacgcc gtgaagaggg aggccgcact gcctcctgtg	2100
agccctctga aggcggcact gagcgaagaa gaacttgaga agaagagtaa ggcgatcatc	2160
gaggagtacc tgcacctgaa cgacatgaag gaggccgtgc agtgcggtgca agagctcgcg	2220
agcccatcac tgctgttcat cttcgtccgg cacggcgtgg agtccacact ggaaagatct	2280
gccattgcta gggagcatat gggccagttg ttgcaccaat tgctttgcgc cggccaccta	2340
agtactgccc agtactatca gggtttatac gagatcctcg aactggccga agacatggag	2400
atcgacatcc ctacagtgtg gctgtacctc gcagaactgg tgacccttat cctgcaggag	2460
ggcggcggtt caatgggtga gttgttcaga gaaatcaca agccactgcg cccactgggc	2520
aaggcggcta gtctccttct ggagattctc ggcctgctct gtaagagcat ggggtccaaag	2580
aaggtgggca ccctgtggag ggaagctgga ctctcctgga aggaattcct ccctgagggt	2640
caggacatcg gcgccttcgt ggccgaacag aaggttgagt acaccctggg agaggaatcg	2700
gaagcgccgg gtcagcgggc tctgccgagt gaggagctca acagacaact cgagaagctc	2760
ctcaaggagg gatcctcaaa tcagagagtg ttcgactgga tcgaggccaa tctcagcgag	2820
cagcaaatcg tgagcaacac gttggttcgc gctctcatga cagccgtgtg ctactcagcc	2880
attatcttcg agacgcctct ccgcgtggac gtggcagtgct tcaaggcccg cgctaagctg	2940
ttacagaagt acctgtgcga cgagcagaag gaactgcagg ccctgtacgc cctgcaagct	3000
ctggtcgtca cactcgagca gcctcctaac ctgctgagaa tgttcttcga cgcattgtac	3060
gacgaggacg ttgtgaagga agacgctttc tacagctggg agtctagtaa ggatcctgcg	3120
gaacaacaag gcaagggcgt tgctctaaag agcgtgaccg ctttcttcaa gtggctgcga	3180
gaggctgagg aagagtcgga ccacaac	3207

<210> 18

<211> 921

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 18

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asp Pro Thr Arg Leu Gln
115 120 125

Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly
130 135 140

Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly
145 150 155 160

Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln
165 170 175

Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr
180 185 190

Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser

195

200

205

Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly
 210 215 220

Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys
 225 230 235 240

Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu
 245 250 255

Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe
 260 265 270

Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met
 275 280 285

Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr
 290 295 300

Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu
 305 310 315 320

Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys
 325 330 335

Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu
 340 345 350

Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn
 355 360 365

Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala
 370 375 380

Ile Met His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu
 385 390 395 400

Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu
 405 410 415

Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met
420 425 430

Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met
435 440 445

Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg
450 455 460

Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu
465 470 475 480

Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys
485 490 495

Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg
500 505 510

Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser
515 520 525

Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys
530 535 540

Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly Gly Arg
545 550 555 560

Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala Lys Pro Ser
565 570 575

Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg
580 585 590

Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu Ser Thr Asp Asn Arg
595 600 605

Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu Arg Gly Glu Lys Ala
610 615 620

Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu Arg Gly Gly Asp Arg

625		630		635		640									
Gly	Asp	Arg	Leu	Asp	Arg	Ala	Arg	Thr	Pro	Ala	Thr	Lys	Arg	Ser	Phe
			645						650					655	
Ser	Lys	Glu	Val	Glu	Glu	Arg	Ser	Arg	Glu	Arg	Pro	Ser	Gln	Pro	Glu
		660						665					670		
Gly	Leu	Arg	Lys	Ala	Ala	Ser	Leu	Thr	Glu	Asp	Arg	Asp	Arg	Gly	Arg
		675					680					685			
Asp	Ala	Val	Lys	Arg	Glu	Ala	Ala	Leu	Pro	Pro	Val	Ser	Pro	Leu	Lys
	690						695				700				
Ala	Ala	Leu	Ser	Glu	Glu	Glu	Leu	Glu	Lys	Lys	Ser	Lys	Ala	Ile	Ile
705					710					715					720
Glu	Glu	Tyr	Leu	His	Leu	Asn	Asp	Met	Lys	Glu	Ala	Val	Gln	Cys	Val
			725						730					735	
Gln	Glu	Leu	Ala	Ser	Pro	Ser	Leu	Leu	Phe	Ile	Phe	Val	Arg	His	Gly
			740					745					750		
Val	Glu	Ser	Thr	Leu	Glu	Arg	Ser	Ala	Ile	Ala	Arg	Glu	His	Met	Gly
		755					760					765			
Gln	Leu	Leu	His	Gln	Leu	Leu	Cys	Ala	Gly	His	Leu	Ser	Thr	Ala	Gln
	770					775					780				
Tyr	Tyr	Gln	Gly	Leu	Tyr	Glu	Ile	Leu	Glu	Leu	Ala	Glu	Asp	Met	Glu
785					790					795					800
Ile	Asp	Ile	Pro	His	Val	Trp	Leu	Tyr	Leu	Ala	Glu	Leu	Val	Thr	Pro
				805					810					815	
Ile	Leu	Gln	Glu	Gly	Gly	Val	Pro	Met	Gly	Glu	Leu	Phe	Arg	Glu	Ile
			820					825					830		
Thr	Lys	Pro	Leu	Arg	Pro	Leu	Gly	Lys	Ala	Ala	Ser	Leu	Leu	Leu	Glu
		835					840					845			

Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro Lys Lys Val Gly Thr
850 855 860

Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu Phe Leu Pro Glu Gly
865 870 875 880

Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys Val Glu Tyr Thr Leu
885 890 895

Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala Leu Pro Ser Glu Glu
900 905 910

Leu Asn Arg Gln Leu Glu Lys Leu Leu
915 920

<210> 19
<211> 2763
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 19
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccccctt ctaacttcgc caacggcatc gccgagtggg tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt 360
ggaagcgacc ctaccagact gcagggcac aactgcggcc ctgacttcac tccttctttc 420
gcaaacctgg gcagaaccac cctgagcacc agaggccctc ctagaggtgg tcccggcgga 480
gaactcccca ggggtcctgc cggcctgggc cctagacgct ccagcaagg tcctagaaag 540
gagccaagga agatcatcgc caccgtgctg atgaccgagg acatcaagct gaacaaagct 600
gagaaggcct ggaagcctag cagcaagaga accgccgccg acaaggacag aggcgaggag 660

gacgccgacg gatccaagac ccaggacctg ttcagaagag tgagaagcat cctcaacaag	720
ctgaccctc agatgttcca gcagctgatg aagcaggtga cgcagctcgc catcgacacc	780
gaggagagac tgaagggcgt gatcgacctg atctttgaga aggctatctc agagcctaac	840
ttcagcgtgg cctacgcaa catgtgccgt tgcctgatgg cattgaaggt gccaaaccacc	900
gagaagccta ctgtgaccgt caatttccgt aaactgctgc tgaaccggtg ccagaaagag	960
ttcgagaagg ataaggacga cgacgaggtc ttcgagaaga aacagaaaga aatggacgag	1020
gccgccaccg cagaggaaag gggccgatta aaggaggagc tggaggaggc cagagacatc	1080
gccagacggc gttctctggg caacatcaag ttcataggtg agctgttcaa gctaaagatg	1140
ctcaccgagg ccataatgca cgactgcgtg gtgaagctac tgaagaacca cgacgaagaa	1200
agcctggagt gcctgtgcag actgctgacc accatcggca aggacctgga cttcgagaag	1260
gcaaagcctc gaatggacca gtacttcaac cagatggaga agattatcaa ggagaagaag	1320
accagcagca gaatcagatt catgctgcag gacgtactgg acctgcgcgg aagcaactgg	1380
gtgccaagga gaggggacca aggaccaaag accatcgacc agatccacaa ggaagcggag	1440
atggaggagc acagagagca cataaaggtg cagcagctta tggccaaggg cagcgacaag	1500
cgaagaggcg gcccgcggg acctctatc agcagaggcc ttcctctggt agacgacggc	1560
ggctggaaca ccgtgcctat ctctaagggc tccagaccta tcgacacttc ccgtcttacc	1620
aagatcacca agccaggatc tattgacagc aacaaccagc tgttcgcccc aggaggaaga	1680
cttagctggg gcaagggaag ttccggcgga tccggcgcca agccttccga cgccgccagc	1740
gaggctgcca gacctgccac cagcaccttg aaccgctttt ccgctctgca gcaagctgtg	1800
ccgaccgaga gcaccgacaa ccggcgagtt gtgcagagaa gcagcctgag cagagagagg	1860
ggcgagaaag ccggcgatag aggtgaccga ctggagagaa gcgagagagg aggtgataga	1920
ggcgaccggc ttgacagagc cagaaccct gccacaaagc gatcgttcag caaggaggtc	1980
gaagagaggt ccaggagcg ccctagccag ccggaaggac tcagaaaggc agccagtcta	2040
acagaggacc gcgacagggg aagagacgcc gtgaagaggg aggccgcact gcctcctgtg	2100
agccctctga aggcggcact gagcgaagaa gaacttgaga agaagagtaa ggcgatcatc	2160
gaggagtacc tgcacctgaa cgacatgaag gaggccgtgc agtgctgca agagctcgcg	2220
agcccatcac tgctgttcat cttcgtccgg cacggcgtgg agtccacact ggaaagatct	2280

gccattgcta gggagcatat gggccagttg ttgcaccaat tgctttgcgc cggccaccta 2340
 agtactgccc agtactatca gggtttatac gagatcctcg aactggccga agacatggag 2400
 atcgacatcc ctacagtgtg gctgtacctc gcagaactgg tgacccttat cctgcaggag 2460
 ggcggcgttc caatgggtga gttgttcaga gaaatcacia agccactgcg cccactgggc 2520
 aaggcggcta gtctccttct ggagattctc ggcctgctct gtaagagcat gggtcctaaag 2580
 aaggtgggca ccctgtggag ggaagctgga ctctcctgga aggaattcct ccctgagggt 2640
 caggacatcg gcgccttcgt ggccgaacag aaggttgagt acaccctggg agaggaatcg 2700
 gaagcgccgg gtcagcgggc tctgccgagt gaggagctca acagacaact cgagaagctc 2760
 ctc 2763

<210> 20

<211> 528

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 20

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
 1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
 20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
 35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
 50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
 65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
 85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Gly Arg Thr Thr Leu Ser
115 120 125

Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly Glu Leu Pro Arg Gly
130 135 140

Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu
145 150 155 160

Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr Glu Asp Ile Lys Leu
165 170 175

Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala
180 185 190

Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp
195 200 205

Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys Leu Thr Pro Gln Met
210 215 220

Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu Ala Ile Asp Thr Glu
225 230 235 240

Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser
245 250 255

Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met Cys Arg Cys Leu Met
260 265 270

Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr Val Thr Val Asn Phe
275 280 285

Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys
290 295 300

Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys Glu Met Asp Glu Ala
305 310 315 320

Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala
325 330 335

Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn Ile Lys Phe Ile Gly
340 345 350

Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile Met His Asp Cys
355 360 365

Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser Leu Glu Cys Leu
370 375 380

Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala
385 390 395 400

Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys Ile Ile Lys
405 410 415

Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln Asp Val Leu
420 425 430

Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg Gly Asp Gln Gly Pro
435 440 445

Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu Met Glu Glu His Arg
450 455 460

Glu His Ile Lys Val Gln Gln Leu Met Ala Lys Gly Ser Asp Lys Arg
465 470 475 480

Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg Gly Leu Pro Leu Val
485 490 495

Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser Lys Gly Ser Arg Pro
500 505 510

Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp
515 520 525

<210> 21
<211> 1584
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 21
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccctt ctaacttcgc caacggcatc gccgagtga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt 360
ggaagcggca gaaccacct gagcaccaga ggccctccta gaggtggtcc cggcggagaa 420
ctcccaggg gtcctgccgg cctgggccct agacgctccc agcaaggtcc tagaaaggag 480
ccaaggaaga tcatcgccac cgtgctgatg accgaggaca tcaagctgaa caaagctgag 540
aaggcctgga agcctagcag caagagaacc gccgccgaca aggacagagg cgaggaggac 600
gccgacggat ccaagaccca ggacctgttc agaagagtga gaagcatcct caacaagctg 660
accctcaga tgttcagca gctgatgaag cagggtacgc agctcgccat cgacaccgag 720
gagagactga agggcgtgat cgacctgatc tttgagaagg ctatctcaga gcctaacttc 780
agcgtggcct acgccaacat gtgccgttgc ctgatggcat tgaagtgcc aaccaccgag 840
aagcctactg tgaccgtcaa tttccgtaaa ctgctgctga accggtgcca gaaagagttc 900
gagaaggata aggacgacga cgaggtcttc gagaagaaac agaaagaaat ggacgaggcc 960
gccaccgagc aggaaggagg ccgattaaag gaggagctgg aggaggccag agacatcgcc 1020
agacggcggt ctctgggcaa catcaagttc ataggtgagc tgttcaagct aaagatgctc 1080
accgaggcca taatgcacga ctgcgtggtg aagctactga agaaccacga cgaagaaagc 1140
ctggagtgcc tgtgcagact gctgaccacc atcggaagg acctggactt cgagaaggca 1200
aagcctcgaa tggaccagta cttcaaccag atggagaaga ttatcaagga gaagaagacc 1260
agcagcagaa tcagattcat gctgcaggac gtactggacc tgcgcggaag caactgggtg 1320

ccaaggagag gggaccaagg accaaagacc atcgaccaga tccacaagga agcggagatg	1380
gaggagcaca gagagcacat aaaggtgcag cagcttatgg ccaagggcag cgacaagcga	1440
agaggcggcc cgcccggacc tcctatcagc agaggccttc ctctggtaga cgacggcggc	1500
tggaacaccg tgcctatctc taagggtccc agacctatcg acatttcccg tcttaccaag	1560
atcaccaagc caggatctat tgac	1584

<210> 22
 <211> 600
 <212> PRT
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 22
 Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
 1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
 20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
 35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
 50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
 65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
 85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
 100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asp Pro Thr Arg Leu Gln
 115 120 125

Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly
130 135 140

Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly
145 150 155 160

Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln
165 170 175

Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr
180 185 190

Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser
195 200 205

Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly
210 215 220

Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys
225 230 235 240

Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu
245 250 255

Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe
260 265 270

Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met
275 280 285

Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr
290 295 300

Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu
305 310 315 320

Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys
325 330 335

Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu

340

345

350

Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn
 355 360 365

Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala
 370 375 380

Ile Met His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu
 385 390 395 400

Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu
 405 410 415

Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met
 420 425 430

Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met
 435 440 445

Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg
 450 455 460

Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu
 465 470 475 480

Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys
 485 490 495

Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg
 500 505 510

Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser
 515 520 525

Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys
 530 535 540

Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly Gly Arg
 545 550 555 560

Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala Lys Pro Ser
565 570 575

Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg
580 585 590

Phe Ser Ala Leu Gln Gln Ala Val
595 600

<210> 23
<211> 1800
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 23
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccccctt ctaacttcgc caacggcatc gccgagtggg tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt 360
ggaagcgacc ctaccagact gcagggcacg aactgcggcc ctgacttcac tccttctttc 420
gcaaacctgg gcagaaccac cctgagcacc agaggccctc ctagaggtgg tcccggcgga 480
gaactcccca ggggtcctgc cggcctgggc cctagacgct ccagcaagg tcctagaaag 540
gagccaagga agatcatcgc caccgtgctg atgaccgagg acatcaagct gaacaaagct 600
gagaaggcct ggaagcctag cagcaagaga accgccgccg acaaggacag aggcgaggag 660
gacgccgacg gatccaagac ccaggacctg ttcagaagag tgagaagcat cctcaacaag 720
ctgaccctc agatgttcca gcagctgatg aagcaggtga cgcagctcgc catcgacacc 780
gaggagagac tgaaggcggt gatcgacctg atctttgaga aggctatctc agagcctaac 840
ttcagcgtgg cctacgcaa catgtgccgt tgcctgatgg cattgaaggt gccaaaccacc 900

gagaagccta ctgtgaccgt caatttccgt aaactgctgc tgaaccggtg ccagaaagag 960
 ttcgagaagg ataaggacga cgacgaggtc ttcgagaaga aacagaaaga aatggacgag 1020
 gccgccaccg cagaggaaag gggccgatta aaggaggagc tggaggaggc cagagacatc 1080
 gccagacggc gttctctggg caacatcaag ttcataggtg agctgttcaa gctaaagatg 1140
 ctcaccgagg ccataatgca cgactgcgtg gtgaagctac tgaagaacca cgacgaagaa 1200
 agcctggagt gcctgtgcag actgctgacc accatcggca aggacctgga cttcgagaag 1260
 gcaaagcctc gaatggacca gtacttcaac cagatggaga agattatcaa ggagaagaag 1320
 accagcagca gaatcagatt catgctgcag gacgtactgg acctgcgcgg aagcaactgg 1380
 gtgccaagga gaggggacca aggaccaaag accatcgacc agatccacaa ggaagcggag 1440
 atggaggagc acagagagca cataaagggtg cagcagctta tggccaaggg cagcgacaag 1500
 cgaagaggcg gcccggccgg acctcctatc agcagaggcc ttcctctggt agacgacggc 1560
 ggctggaaca ccgtgcctat ctctaagggc tccagaccta tcgacacttc ccgtcttacc 1620
 aagatcacca agccaggatc tattgacagc aacaaccagc tgttcgcccc aggaggaaga 1680
 cttagctggg gcaaggggaag ttccggcgga tccggcgcca agccttccga cgccgccagc 1740
 gaggctgcca gacctgccac cagcaccttg aaccgctttt ccgctctgca gcaagctgtg 1800

<210> 24

<211> 541

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 24

Met	Ala	Ser	Asn	Phe	Thr	Gln	Phe	Val	Leu	Val	Asp	Asn	Gly	Gly	Thr
1				5					10					15	

Gly	Asp	Val	Thr	Val	Ala	Pro	Ser	Asn	Phe	Ala	Asn	Gly	Ile	Ala	Glu
			20					25					30		

Trp	Ile	Ser	Ser	Asn	Ser	Arg	Ser	Gln	Ala	Tyr	Lys	Val	Thr	Cys	Ser
		35					40					45			

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Pro Arg Lys Ile Ile Ala
115 120 125

Thr Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala
130 135 140

Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu
145 150 155 160

Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg
165 170 175

Ser Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys
180 185 190

Gln Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val
195 200 205

Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val
210 215 220

Ala Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr
225 230 235 240

Thr Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn
245 250 255

Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe
260 265 270

Glu Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg
275 280 285

Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg
290 295 300

Arg Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys
305 310 315 320

Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu Lys
325 330 335

Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr
340 345 350

Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln
355 360 365

Tyr Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser
370 375 380

Arg Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn
385 390 395 400

Trp Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile
405 410 415

His Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val Gln
420 425 430

Gln Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly
435 440 445

Pro Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn
450 455 460

Thr Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu
465 470 475 480

Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe
485 490 495

Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser
500 505 510

Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr
515 520 525

Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val
530 535 540

<210> 25

<211> 1623

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 25

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggcccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcccaa ggaagatcat cgccaccgtg ctgatgaccg aggacatcaa gctgaacaaa	420
gctgagaagg cctggaagcc tagcagcaag agaaccgccg ccgacaagga cagaggcgag	480
gaggacgccg acggatccaa gaccaggac ctgttcagaa gagtgagaag catcctcaac	540
aagctgaccc ctcatgtgtt ccagcagctg atgaagcagg tgacgcagct cgccatcgac	600
accgaggaga gactgaaggg cgtgatcgac ctgatctttg agaaggctat ctcagagcct	660
aacttcagcg tggcctacgc caacatgtgc cgttgcctga tggcattgaa ggtgccaacc	720
accgagaagc ctactgtgac cgtcaatttc cgtaaactgc tgctgaaccg gtgccagaaa	780
gagttcgaga aggataagga cgacgacgag gtcttcgaga agaaacagaa agaaatggac	840

gaggccgcca ccgcagagga aaggggccga ttaaaggagg agctggagga ggccagagac 900
 atcgccagac ggcgttctct gggcaacatc aagttcatag gtgagctgtt caagctaaag 960
 atgctcaccg aggccataat gcacgactgc gtggtgaagc tactgaagaa ccacgacgaa 1020
 gaaagcctgg agtgcctgtg cagactgctg accaccatcg gcaaggacct ggacttcgag 1080
 aaggcaaagc ctcgaatgga ccagtacttc aaccagatgg agaagattat caaggagaag 1140
 aagaccagca gcagaatcag attcatgctg caggacgtac tggacctgcg cggaagcaac 1200
 tgggtgccaa ggagagggga ccaaggacca aagaccatcg accagatcca caaggaagcg 1260
 gagatggagg agcacagaga gcacataaag gtgcagcagc ttatggccaa gggcagcgac 1320
 aagcgaagag gcggcccgcc cggacctcct atcagcagag gccttcctct ggtagacgac 1380
 ggcggctgga acaccgtgcc tatctctaag ggctccagac ctatcgacac ttcccgtctt 1440
 accaagatca ccaagccagg atctattgac agcaacaacc agctgttcgc cccaggagga 1500
 agacttagct ggggcaaggg aagttccggc ggatccggcg ccaagccttc cgacgccgcc 1560
 agcgaggctg ccagacctgc caccagcacc ttgaaccgct tttccgctct gcagcaagct 1620
 gtg 1623

<210> 26
 <211> 642
 <212> PRT
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 26
 Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
 1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
 20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
 35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu

50

55

60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Ser Asn Asn Gln Leu Phe
115 120 125

Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser
130 135 140

Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr
145 150 155 160

Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu
165 170 175

Ser Thr Asp Asn Arg Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu
180 185 190

Arg Gly Glu Lys Ala Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu
195 200 205

Arg Gly Gly Asp Arg Gly Asp Arg Leu Asp Arg Ala Arg Thr Pro Ala
210 215 220

Thr Lys Arg Ser Phe Ser Lys Glu Val Glu Glu Arg Ser Arg Glu Arg
225 230 235 240

Pro Ser Gln Pro Glu Gly Leu Arg Lys Ala Ala Ser Leu Thr Glu Asp
245 250 255

Arg Asp Arg Gly Arg Asp Ala Val Lys Arg Glu Ala Ala Leu Pro Pro
260 265 270

Val Ser Pro Leu Lys Ala Ala Leu Ser Glu Glu Glu Leu Glu Lys Lys
275 280 285

Ser Lys Ala Ile Ile Glu Glu Tyr Leu His Leu Asn Asp Met Lys Glu
290 295 300

Ala Val Gln Cys Val Gln Glu Leu Ala Ser Pro Ser Leu Leu Phe Ile
305 310 315 320

Phe Val Arg His Gly Val Glu Ser Thr Leu Glu Arg Ser Ala Ile Ala
325 330 335

Arg Glu His Met Gly Gln Leu Leu His Gln Leu Leu Cys Ala Gly His
340 345 350

Leu Ser Thr Ala Gln Tyr Tyr Gln Gly Leu Tyr Glu Ile Leu Glu Leu
355 360 365

Ala Glu Asp Met Glu Ile Asp Ile Pro His Val Trp Leu Tyr Leu Ala
370 375 380

Glu Leu Val Thr Pro Ile Leu Gln Glu Gly Gly Val Pro Met Gly Glu
385 390 395 400

Leu Phe Arg Glu Ile Thr Lys Pro Leu Arg Pro Leu Gly Lys Ala Ala
405 410 415

Ser Leu Leu Leu Glu Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro
420 425 430

Lys Lys Val Gly Thr Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu
435 440 445

Phe Leu Pro Glu Gly Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys
450 455 460

Val Glu Tyr Thr Leu Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala
465 470 475 480

Leu Pro Ser Glu Glu Leu Asn Arg Gln Leu Glu Lys Leu Leu Lys Glu

485

490

495

Gly Ser Ser Asn Gln Arg Val Phe Asp Trp Ile Glu Ala Asn Leu Ser
500 505 510

Glu Gln Gln Ile Val Ser Asn Thr Leu Val Arg Ala Leu Met Thr Ala
515 520 525

Val Cys Tyr Ser Ala Ile Ile Phe Glu Thr Pro Leu Arg Val Asp Val
530 535 540

Ala Val Leu Lys Ala Arg Ala Lys Leu Leu Gln Lys Tyr Leu Cys Asp
545 550 555 560

Glu Gln Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Leu Val Val
565 570 575

Thr Leu Glu Gln Pro Pro Asn Leu Leu Arg Met Phe Phe Asp Ala Leu
580 585 590

Tyr Asp Glu Asp Val Val Lys Glu Asp Ala Phe Tyr Ser Trp Glu Ser
595 600 605

Ser Lys Asp Pro Ala Glu Gln Gln Gly Lys Gly Val Ala Leu Lys Ser
610 615 620

Val Thr Ala Phe Phe Lys Trp Leu Arg Glu Ala Glu Glu Glu Ser Asp
625 630 635 640

His Asn

<210> 27

<211> 1926

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 27

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggcccctt ctaacttcgc caacggcatc gccgagtggg tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcgggt	360
ggaagcagca acaaccagct gttcgcccca ggaggaagac ttagctgggg caagggaagt	420
tccggcggat ccggcgccaa gccttccgac gccgccagcg aggctgccag acctgccacc	480
agcaccttga accgcttttc cgctctgcag caagctgtgc cgaccgagag caccgacaac	540
cggcgagttg tgcagagaag cagcctgagc agagagaggg gcgagaaagc cggcgataga	600
ggtgaccgac tggagagaag cgagagagga ggtgatagag gcgaccggct tgacagagcc	660
agaaccctg ccacaaagcg atcgttcagc aaggaggtcg aagagaggtc caggagcgc	720
cctagccagc cggaaggact cagaaaggca gccagtctaa cagaggaccg cgacagggga	780
agagacgccg tgaagaggga ggccgcactg cctcctgtga gccctctgaa ggcggcactg	840
agcgaagaag aacttgagaa gaagagtaag gcgatcatcg aggagtacct gcacctgaac	900
gacatgaagg aggccgtgca gtgctgcaa gagctcgcga gcccatcact gctgttcac	960
ttcgtccggc acggcgtgga gtccacactg gaaagatctg ccattgctag ggagcatatg	1020
ggccagttgt tgcaccaatt gctttgcgcc ggccacctaa gtactgcca gtactatcag	1080
ggtttatacg agatcctcga actggccgaa gacatggaga tcgacatccc tcacgtgtgg	1140
ctgtacctg cagaactggt gaccctatc ctgcaggagg gcggcgttcc aatgggtgag	1200
ttgttcagag aaatcacaaa gccactgcgc ccactgggca aggcggctag tctccttctg	1260
gagattctcg gcctgctctg taagagcatg ggtccaaaga aggtgggcac cctgtggagg	1320
gaagctggac tctcctggaa ggaattcctc cctgagggtc aggacatcgg cgccttcgtg	1380
gccgaacaga aggttgagta caccctggga gaggaatcgg aagcgccggg tcagcgggct	1440
ctgccgagtg aggagctcaa cagacaactc gagaagctcc tcaaggaggg atcctcaaat	1500
cagagagtgt tcgactggat cgaggccaat ctcagcgagc agcaaatcgt gagcaacacg	1560
ttggttcgcg ctctcatgac agccgtgtgc tactcagcca ttatcttcga gacgcctctc	1620

cgcgtggacg tggcagtgt caagggccgc gctaagctgt tacagaagta cctgtgacgac 1680
gagcagaagg aactgcaggc cctgtacgcc ctgcaagctc tggtcgtcac actcgagcag 1740
cctcctaacc tgctgagaat gttcttcgac gcattgtacg acgaggacgt tgtgaaggaa 1800
gacgctttct acagctggga gtctagtaag gatcctgcgg aacaacaagg caagggcggtt 1860
gctctaaaga gcgtgaccgc cttcttcaag tggctgcgag aggctgagga agagtcggac 1920
cacaac 1926

<210> 28
<211> 494
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 28
Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Ser Asn Asn Gln Leu Phe
115 120 125

Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser
130 135 140

Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr
145 150 155 160

Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu
165 170 175

Ser Thr Asp Asn Arg Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu
180 185 190

Arg Gly Glu Lys Ala Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu
195 200 205

Arg Gly Gly Asp Arg Gly Asp Arg Leu Asp Arg Ala Arg Thr Pro Ala
210 215 220

Thr Lys Arg Ser Phe Ser Lys Glu Val Glu Glu Arg Ser Arg Glu Arg
225 230 235 240

Pro Ser Gln Pro Glu Gly Leu Arg Lys Ala Ala Ser Leu Thr Glu Asp
245 250 255

Arg Asp Arg Gly Arg Asp Ala Val Lys Arg Glu Ala Ala Leu Pro Pro
260 265 270

Val Ser Pro Leu Lys Ala Ala Leu Ser Glu Glu Glu Leu Glu Lys Lys
275 280 285

Ser Lys Ala Ile Ile Glu Glu Tyr Leu His Leu Asn Asp Met Lys Glu
290 295 300

Ala Val Gln Cys Val Gln Glu Leu Ala Ser Pro Ser Leu Leu Phe Ile
305 310 315 320

Phe Val Arg His Gly Val Glu Ser Thr Leu Glu Arg Ser Ala Ile Ala
325 330 335

Arg Glu His Met Gly Gln Leu Leu His Gln Leu Leu Cys Ala Gly His
340 345 350

Leu Ser Thr Ala Gln Tyr Tyr Gln Gly Leu Tyr Glu Ile Leu Glu Leu
355 360 365

Ala Glu Asp Met Glu Ile Asp Ile Pro His Val Trp Leu Tyr Leu Ala
370 375 380

Glu Leu Val Thr Pro Ile Leu Gln Glu Gly Gly Val Pro Met Gly Glu
385 390 395 400

Leu Phe Arg Glu Ile Thr Lys Pro Leu Arg Pro Leu Gly Lys Ala Ala
405 410 415

Ser Leu Leu Leu Glu Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro
420 425 430

Lys Lys Val Gly Thr Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu
435 440 445

Phe Leu Pro Glu Gly Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys
450 455 460

Val Glu Tyr Thr Leu Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala
465 470 475 480

Leu Pro Ser Glu Glu Leu Asn Arg Gln Leu Glu Lys Leu Leu
485 490

<210> 29

<211> 1482

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 29

atggccagca acttcacca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60

gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120

caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgagag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcagca acaaccagct gttcgcccca ggaggaagac ttagctgggg caagggaagt	420
tccggcggat ccggcgccaa gccttccgac gccgccagcg aggctgccag acctgccacc	480
agcaccttga accgcttttc cgctctgcag caagctgtgc cgaccgagag caccgacaac	540
cggcgagttg tgcagagaag cagcctgagc agagagaggg gcgagaaagc cggcgataga	600
ggtgaccgac tggagagaag cgagagagga ggtgatagag gcgaccggct tgacagagcc	660
agaacccctg ccacaaagcg atcgttcagc aaggaggtcg aagagaggtc cagggagcgc	720
cctagccagc cggaaggact cagaaaggca gccagtctaa cagaggaccg cgacagggga	780
agagacgccg tgaagagggg ggccgcactg cctcctgtga gccctctgaa ggcggcactg	840
agcgaagaag aacttgagaa gaagagtaag gcgatcatcg aggagtacct gcacctgaac	900
gacatgaagg aggccgtgca gtgcgtgcaa gagctcgcga gcccatcact gctgttcac	960
ttcgtccggc acggcgtgga gtccacactg gaaagatctg ccattgctag ggagcatatg	1020
ggccagttgt tgcaccaatt gctttgcgcc ggccaccta gtactgcca gtactatcag	1080
ggtttatacg agatcctcga actggccgaa gacatggaga tcgacatccc tcacgtgtgg	1140
ctgtacctcg cagaactggt gaccctatc ctgcaggagg gcggcgttcc aatgggtgag	1200
ttgttcagag aaatcacaaa gccactgcgc cactgggca aggcggctag tctccttctg	1260
gagattctcg gcctgctctg taagagcatg ggtccaaaga aggtgggcac cctgtggagg	1320
gaagctggac tctcctggaa ggaattcctc cctgagggtc aggacatcgg cgccttcgtg	1380
gccgaacaga aggttgagta caccctggga gaggaatcgg aagcgccggg tcagcgggct	1440
ctgccgagtg aggagctcaa cagacaactc gagaagctcc tc	1482

<210> 30

<211> 719

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 30

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Glu Asn Ser Ser Ala Ala
115 120 125

Ser Ala Ser Ser Glu Ala Gly Ser Ser Arg Ser Gln Glu Ile Glu Glu
130 135 140

Leu Glu Arg Phe Ile Asp Ser Tyr Val Leu Glu Tyr Gln Val Gln Gly
145 150 155 160

Leu Leu Ala Asp Lys Thr Glu Gly Asp Gly Glu Ser Glu Arg Thr Gln
165 170 175

Ser His Ile Ser Gln Trp Thr Ala Asp Cys Ser Glu Pro Leu Asp Ser
180 185 190

Ser Cys Ser Phe Ser Arg Gly Arg Ala Pro Pro Gln Gln Asn Gly Ser

195

200

205

Lys Asp Asn Ser Leu Asp Met Leu Gly Thr Asp Ile Trp Ala Ala Asn
 210 215 220

Thr Phe Asp Ser Phe Ser Gly Ala Thr Trp Asp Leu Gln Pro Glu Lys
 225 230 235 240

Leu Asp Phe Thr Gln Phe His Arg Lys Val Arg His Thr Pro Lys Gln
 245 250 255

Pro Leu Pro His Ile Asp Arg Glu Gly Cys Gly Lys Gly Lys Leu Glu
 260 265 270

Asp Gly Asp Gly Ile Asn Leu Asn Asp Ile Glu Lys Val Leu Pro Ala
 275 280 285

Trp Gln Gly Tyr His Pro Met Pro His Glu Val Glu Ile Ala His Thr
 290 295 300

Lys Lys Leu Phe Arg Arg Arg Arg Asn Asp Arg Arg Arg Gln Gln Arg
 305 310 315 320

Pro Pro Gly Gly Asn Lys Pro Gln Gln His Gly Asp His Gln Pro Gly
 325 330 335

Ser Ala Lys His Asn Arg Asp His Gln Lys Ser Tyr Gln Gly Gly Ser
 340 345 350

Ala Pro His Pro Ser Gly Arg Pro Thr His His Gly Tyr Ser Gln Asn
 355 360 365

Arg Arg Trp His His Gly Asn Met Lys His Pro Pro Gly Asp Lys Gly
 370 375 380

Glu Ala Gly Ala His Arg Asn Ala Lys Glu Thr Met Thr Ile Glu Asn
 385 390 395 400

Pro Lys Leu Glu Asp Thr Ala Gly Asp Thr Gly His Ser Ser Leu Glu
 405 410 415

Ala Pro Arg Ser Pro Asp Thr Leu Ala Pro Val Ala Ser Glu Arg Leu
420 425 430

Pro Pro Gln Gln Ser Gly Gly Pro Glu Val Glu Thr Lys Arg Lys Asp
435 440 445

Ser Ile Leu Pro Glu Arg Ile Gly Glu Arg Pro Lys Ile Thr Leu Leu
450 455 460

Gln Ser Ser Lys Asp Arg Leu Arg Arg Arg Leu Lys Glu Lys Asp Glu
465 470 475 480

Val Ala Val Glu Thr Thr Thr Pro Gln Gln Asn Lys Met Asp Lys Leu
485 490 495

Ile Glu Ile Leu Asn Ser Met Arg Asn Asn Ser Ser Asp Val Asp Thr
500 505 510

Lys Leu Thr Thr Phe Met Glu Glu Ala Gln Asn Ser Thr Asn Ser Glu
515 520 525

Glu Met Leu Gly Glu Ile Val Arg Thr Ile Tyr Gln Lys Ala Val Ser
530 535 540

Asp Arg Ser Phe Ala Phe Thr Ala Ala Lys Leu Cys Asp Lys Met Ala
545 550 555 560

Leu Phe Met Val Glu Gly Thr Lys Phe Arg Ser Leu Leu Leu Asn Met
565 570 575

Leu Gln Lys Asp Phe Thr Val Arg Glu Glu Leu Gln Gln Gln Asp Val
580 585 590

Glu Arg Trp Leu Gly Phe Ile Thr Phe Leu Cys Glu Val Phe Gly Thr
595 600 605

Met Arg Ser Ser Thr Gly Glu Pro Phe Arg Val Leu Val Cys Pro Ile
610 615 620

Tyr Thr Cys Leu Arg Glu Leu Leu Gln Ser Gln Asp Val Lys Glu Asp

625		630		635		640
Ala Val Leu Cys Cys Ser Met Glu Leu Gln Ser Thr Gly Arg Leu Leu						
		645		650		655

Glu Glu Gln Leu Pro Glu Met Met Thr Glu Leu Leu Ala Ser Ala Arg
660 665 670

Asp Lys Met Leu Cys Pro Ser Glu Ser Met Leu Thr Arg Ser Leu Leu
675 680 685

Leu Glu Val Ile Glu Leu His Ala Asn Ser Trp Asn Pro Leu Thr Pro
690 695 700

Pro Ile Thr Gln Tyr Tyr Asn Arg Thr Ile Gln Lys Leu Thr Ala
705 710 715

<210> 31
 <211> 2157
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 31	
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggcccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcgaga acagtagcgc cgctagcgcg agcagcgaag ccgggagcag ccggtctcag	420
gagatcgagg agctggagcg gttcatcgac agctacgtgc tggagtacca ggtgcagggc	480
ctgctggccg acaagaccga gggcgacggc gagagcgagc ggacccagag ccacatcagc	540
cagtggaccg ccgactgcag cgagcccctg gactctagct gctcattcag tcgtggacgg	600

gccccctccac agcagaacgg cagcaaggac aacagcctgg acatgctggg caccgacatc	660
tgggcccgcca acaccttcga cagcttcagc ggcgccacct gggatctgca gcccgagaag	720
ctggacttta cccagttcca ccggaaggtg cggcacactc ccaagcagcc cctgccccac	780
atcgatcggg agggctgcgg caagggaag ctggaagacg gcgacggcat caacctgaac	840
gacatcgaga aggtgctgcc tgcctggcag ggctaccacc ccatgcccc aagaggtggag	900
atgccccaca ccaagaagct gttccggcga cgacgcaacg accggcgtag gcagcaacgg	960
ccgcctggag ggaacaagcc ccagcagcac ggagaccacc agcccggtag cgccaagcac	1020
aaccgggacc accagaagag ctaccagggc ggaagcgcac cacaccctc gggcagaccc	1080
acccaccacg gctacagcca gaaccggcgg tggcatcacg gtaacatgaa gcacccaccc	1140
ggcgacaaaag gagaggccgg cgctcacctg aacgccaagg agaccatgac catcgagaac	1200
cccaagctgg aggataccgc cggcgatacg ggtcacagca gcctggaggc accgcggtct	1260
cccgacaccc tggcacccgt ggccagcgaa cggtgccac cccaacagag cggcggccct	1320
gaggttgaga ccaagcggaa ggacagcatc ctgcccgaac ggatcggtga gcggcccaag	1380
atcaccttac tgcagagtag caaggaccgg ctgagacggc ggctgaagga gaaggacgag	1440
gtggccgtgg agacaaccac tccccagcag aacaagatgg acaagctgat cgagatcctg	1500
aacagcatgc ggaacaacag cagcgacgtg gacaccaagc tgaccacctt catggaggag	1560
gcccagaaca gcaccaacag cgaggagatg ctgggcgaga tcgtgcggac catctaccag	1620
aaggccgtga gcgaccggag cttcgccttc accgccgcca agctgtgcga caagatggcc	1680
ctgttcatgg tggagggcac caagttccgg agcttactgc tgaatatgct gcagaaggac	1740
ttcaccgtgc gggaggagct gcagcagcag gacgtggagc ggtggctggg cttcatcacc	1800
ttcctgtgcg aggtgttcgg caccatgcgg agcagcaccg gcgaaccctt ccgggtgctg	1860
gtgtgcccc a tctacacctg cctgcgggag ttgctgcaga gccaggacgt gaaggaggac	1920
gccgtgctgt gctgcagcat ggaactgcag agcactggcc ggctgctgga ggagcagctg	1980
cccgagatga tgaccgagct gctcgctagc gcccgggaca agatgctgtg ccccgagcag	2040
agcatgctga cccggagcct gcttctggag gtgatcgagc tgcacgcca cagctggaat	2100
cccctgaccc ctcccatcac ccagtactac aaccggacca tccagaagct gaccgcc	2157

<211> 323
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 32
Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Leu Asn Ser Met Arg Asn
115 120 125

Asn Ser Ser Asp Val Asp Thr Lys Leu Thr Thr Phe Met Glu Glu Ala
130 135 140

Gln Asn Ser Thr Asn Ser Glu Glu Met Leu Gly Glu Ile Val Arg Thr
145 150 155 160

Ile Tyr Gln Lys Ala Val Ser Asp Arg Ser Phe Ala Phe Thr Ala Ala
165 170 175

Lys Leu Cys Asp Lys Met Ala Leu Phe Met Val Glu Gly Thr Lys Phe
180 185 190

Arg Ser Leu Leu Leu Asn Met Leu Gln Lys Asp Phe Thr Val Arg Glu
195 200 205

Glu Leu Gln Gln Gln Asp Val Glu Arg Trp Leu Gly Phe Ile Thr Phe
210 215 220

Leu Cys Glu Val Phe Gly Thr Met Arg Ser Ser Thr Gly Glu Pro Phe
225 230 235 240

Arg Val Leu Val Cys Pro Ile Tyr Thr Cys Leu Arg Glu Leu Leu Gln
245 250 255

Ser Gln Asp Val Lys Glu Asp Ala Val Leu Cys Cys Ser Met Glu Leu
260 265 270

Gln Ser Thr Gly Arg Leu Leu Glu Glu Gln Leu Pro Glu Met Met Thr
275 280 285

Glu Leu Leu Ala Ser Ala Arg Asp Lys Met Leu Cys Pro Ser Glu Ser
290 295 300

Met Leu Thr Arg Ser Leu Leu Leu Glu Val Ile Glu Leu His Ala Asn
305 310 315 320

Ser Trp Asn

<210> 33

<211> 969

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 33

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60

gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120

caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc 180
 atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
 cctatcttcg ccaccaacag cgactgagcg ctgatcgtga aggccatgca gggcctgctg 300
 aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt 360
 ggaagcctga acagcatgag aaacaacagc agcgacgtgg acaccaagct gaccaccttc 420
 atggaggagg cccagaacag caccaacagc gaggagatgc tgggcgagat cgtgagaacc 480
 atctaccaga aggccgtgag cgacagaagc ttgccttca ccgccgcaa gctgtgcgac 540
 aagatggccc tgttcatggt ggagggcacc aagttcagaa gcctgctgct gaacatgctg 600
 cagaaggact tcaccgtgag agaggagctg cagcagcagg acgtggagcg atggctgggc 660
 ttcatcacct tcctgtgcga ggtgttcggc accatgagaa gcagcaccgg cgagcctttc 720
 agagtgtctg tgtgccctat ctacacctgc ctgagagagc tgctgcagag ccaggacgtg 780
 aaggaggacg ccgtgctgtg ctgcagcatg gagctgcaga gcaccggcag actgctggag 840
 gagcagctgc ctgagatgat gaccgagctg ctggccagcg ccagagacaa gatgctgtgc 900
 cctagcgaga gcatgctgac tagatcccta ctgttgagg tgatcgagct gcacgccaac 960
 agctggaac 969

<210> 34

<211> 356

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 34

Met	Ala	Ser	Asn	Phe	Thr	Gln	Phe	Val	Leu	Val	Asp	Asn	Gly	Gly	Thr
1				5					10					15	

Gly	Asp	Val	Thr	Val	Ala	Pro	Ser	Asn	Phe	Ala	Asn	Gly	Ile	Ala	Glu
			20					25					30		

Trp	Ile	Ser	Ser	Asn	Ser	Arg	Ser	Gln	Ala	Tyr	Lys	Val	Thr	Cys	Ser
		35					40					45			

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Thr Thr Pro Gln Gln Asn
115 120 125

Lys Met Asp Lys Leu Ile Glu Ile Leu Asn Ser Met Arg Asn Asn Ser
130 135 140

Ser Asp Val Asp Thr Lys Leu Thr Thr Phe Met Glu Glu Ala Gln Asn
145 150 155 160

Ser Thr Asn Ser Glu Glu Met Leu Gly Glu Ile Val Arg Thr Ile Tyr
165 170 175

Gln Lys Ala Val Ser Asp Arg Ser Phe Ala Phe Thr Ala Ala Lys Leu
180 185 190

Cys Asp Lys Met Ala Leu Phe Met Val Glu Gly Thr Lys Phe Arg Ser
195 200 205

Leu Leu Leu Asn Met Leu Gln Lys Asp Phe Thr Val Arg Glu Glu Leu
210 215 220

Gln Gln Gln Asp Val Glu Arg Trp Leu Gly Phe Ile Thr Phe Leu Cys
225 230 235 240

Glu Val Phe Gly Thr Met Arg Ser Ser Thr Gly Glu Pro Phe Arg Val
245 250 255

Leu Val Cys Pro Ile Tyr Thr Cys Leu Arg Glu Leu Leu Gln Ser Gln

260

265

270

Asp Val Lys Glu Asp Ala Val Leu Cys Cys Ser Met Glu Leu Gln Ser
 275 280 285

Thr Gly Arg Leu Leu Glu Glu Gln Leu Pro Glu Met Met Thr Glu Leu
 290 295 300

Leu Ala Ser Ala Arg Asp Lys Met Leu Cys Pro Ser Glu Ser Met Leu
 305 310 315 320

Thr Arg Ser Leu Leu Leu Glu Val Ile Glu Leu His Ala Asn Ser Trp
 325 330 335

Asn Pro Leu Thr Pro Pro Ile Thr Gln Tyr Tyr Asn Arg Thr Ile Gln
 340 345 350

Lys Leu Thr Ala
 355

<210> 35

<211> 1068

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 35

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggccccctt ctaacttcgc caacggcatc gccgagtggg tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgagag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcacca cccctcagca gaacaagatg gacaagctga tcgagatcct gaacagcatg	420
agaaacaaca gcagcgacgt ggacaccaag ctgaccacct tcatggagga ggcccagaac	480


```

agcaccaaca gcgaggagat gctgggagag atcgtgagaa ccatctacca gaaggccgtg      540
agcgacagaa gcttcgcctt caccgccgcc aagctgtgcg acaagatggc cctgttcatg      600
gtggaggggca ccaagttcag aagcctgctg ctgaacatgc tgcagaagga cttcaccgtg      660
agagaggagc tgcagcagca ggacgtggag aggtggctgg gcttcatcac cttcctgtgc      720
gaggtgttcg gcacatgag aagcagcacc ggcgagcctt tcagagtgtg ggtgtgacct      780
atctacacct gcctgagaga gctgctgcag agccaggacg tgaaggagga cgccgtgctg      840
tgctgcagca tggagctgca gagcaccggc agactgctgg aggagcagct gcctgagatg      900
atgaccgagc tgctggccag cgccagagac aagatgctgt gccctagcga gagcatgctg      960
acgcgcagcc tcctgctgga ggtgatcgag ctgcacgcc aacagctggaa ccctctgacc     1020
cctcctatca ccagctacta caacagaacc atccagaagc tgaccgcc                       1068

```

<210> 36

<211> 364

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polypeptide"

<400> 36

```

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1           5           10          15

```

```

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
          20          25          30

```

```

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
          35          40          45

```

```

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
          50          55          60

```

```

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65          70          75          80

```

```

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
          85          90          95

```

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Ala Asp Gly Ser Lys Thr
115 120 125

Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys Leu Thr Pro
130 135 140

Gln Met Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu Ala Ile Asp
145 150 155 160

Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe Glu Lys Ala
165 170 175

Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met Cys Arg Cys
180 185 190

Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr Val Thr Val
195 200 205

Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe Glu Lys
210 215 220

Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys Glu Met Asp
225 230 235 240

Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu Glu Leu Glu
245 250 255

Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn Ile Lys Phe
260 265 270

Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile Met His
275 280 285

Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser Leu Glu
290 295 300

Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe Glu
305 310 315 320

Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys Ile
325 330 335

Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln Asp
340 345 350

Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg
355 360

<210> 37

<211> 1092

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 37

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcgccg acggcagcaa gaccaggac ctgttcagaa gagtgagaag catcctgaac	420
aagctgaccc ctcatatggt ccagcagctg atgaagcagg tgaccagct ggccatcgac	480
accgaggaga gactgaaggg cgtgatcgac ctgatcttcg agaaggccat cagcagacct	540
aacttcagcg tggcctacgc caacatgtgc cggtgcctga tggccctgaa ggtgcctacc	600
accgagaagc ctaccgtgac cgtgaacttc agaaagctgc tgctgaaccg gtgccagaag	660
gagttcgaga aggacaagga cgacgacgag gtgttcgaga agaagcagaa ggagatggac	720
gaggccgcca ccgccgagga gagaggcaga ctgaaggagg agctggagga ggccagagac	780
atcgccagaa gaagaagcct gggcaacatc aagttcatcg gcgagctgtt caagctgaag	840

atgctgaccg aggccatcat gcacgactgc gtggtgaagc tgctgaagaa ccacgacgag 900
gagagcctgg agtgcctgtg cagactgctg accaccatcg gcaaggacct ggatttcgag 960
aaggcgaagc ctagaatgga ccagtacttc aaccagatgg agaagatcat caaggagaag 1020
aagaccagca gcagaatcag attcatgctg caggacgtgc tggacctgag aggcagcaac 1080
tgggtgccta ga 1092

<210> 38
<211> 1168
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 38
Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Ser Phe Thr Leu Thr Asn
115 120 125

Lys Asn Val Ile Phe Val Ala Gly Leu Gly Gly Ile Gly Leu Asp Thr
130 135 140

Ser Lys Glu Leu Leu Lys Arg Asp Pro Val Val Leu Gln Arg Arg Asp
145 150 155 160

Trp Glu Asn Pro Gly Val Thr Gln Leu Asn Arg Leu Ala Ala His Pro
165 170 175

Pro Phe Ala Ser Trp Arg Asn Ser Glu Glu Ala Arg Thr Asp Arg Pro
180 185 190

Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp Phe
195 200 205

Pro Ala Pro Glu Ala Val Pro Glu Ser Trp Leu Glu Cys Asp Leu Pro
210 215 220

Glu Ala Asp Thr Val Val Val Pro Ser Asn Trp Gln Met His Gly Tyr
225 230 235 240

Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr Pro Ile Thr Val Asn Pro
245 250 255

Pro Phe Val Pro Thr Glu Asn Pro Thr Gly Cys Tyr Ser Leu Thr Phe
260 265 270

Asn Val Asp Glu Ser Trp Leu Gln Glu Gly Gln Thr Arg Ile Ile Phe
275 280 285

Asp Gly Val Asn Ser Ala Phe His Leu Trp Cys Asn Gly Arg Trp Val
290 295 300

Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser Glu Phe Asp Leu Ser Ala
305 310 315 320

Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala Val Met Val Leu Arg Trp
325 330 335

Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp Met Trp Arg Met Ser Gly

340

345

350

Ile Phe Arg Asp Val Ser Leu Leu His Lys Pro Thr Thr Gln Ile Ser
 355 360 365

Asp Phe His Val Ala Thr Arg Phe Asn Asp Asp Phe Ser Arg Ala Val
 370 375 380

Leu Glu Ala Glu Val Gln Met Cys Gly Glu Leu Arg Asp Tyr Leu Arg
 385 390 395 400

Val Thr Val Ser Leu Trp Gln Gly Glu Thr Gln Val Ala Ser Gly Thr
 405 410 415

Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu Arg Gly Gly Tyr Ala Asp
 420 425 430

Arg Val Thr Leu Arg Leu Asn Val Glu Asn Pro Lys Leu Trp Ser Ala
 435 440 445

Glu Ile Pro Asn Leu Tyr Arg Ala Val Val Glu Leu His Thr Ala Asp
 450 455 460

Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp Val Gly Phe Arg Glu Val
 465 470 475 480

Arg Ile Glu Asn Gly Leu Leu Leu Leu Asn Gly Lys Pro Leu Leu Ile
 485 490 495

Arg Gly Val Asn Arg His Glu His His Pro Leu His Gly Gln Val Met
 500 505 510

Asp Glu Gln Thr Met Val Gln Asp Ile Leu Leu Met Lys Gln Asn Asn
 515 520 525

Phe Asn Ala Val Arg Cys Ser His Tyr Pro Asn His Pro Leu Trp Tyr
 530 535 540

Thr Leu Cys Asp Arg Tyr Gly Leu Tyr Val Val Asp Glu Ala Asn Ile
 545 550 555 560

Glu Thr His Gly Met Val Pro Met Asn Arg Leu Thr Asp Asp Pro Arg
565 570 575

Trp Leu Pro Ala Met Ser Glu Arg Val Thr Arg Met Val Gln Arg Asp
580 585 590

Arg Asn His Pro Ser Val Ile Ile Trp Ser Leu Gly Asn Glu Ser Gly
595 600 605

His Gly Ala Asn His Asp Ala Leu Tyr Arg Trp Ile Lys Ser Val Asp
610 615 620

Pro Ser Arg Pro Val Gln Tyr Glu Gly Gly Gly Ala Asp Thr Thr Ala
625 630 635 640

Thr Asp Ile Ile Cys Pro Met Tyr Ala Arg Val Asp Glu Asp Gln Pro
645 650 655

Phe Pro Ala Val Pro Lys Trp Ser Ile Lys Lys Trp Leu Ser Leu Pro
660 665 670

Gly Glu Thr Arg Pro Leu Ile Leu Cys Glu Tyr Ala His Ala Met Gly
675 680 685

Asn Ser Leu Gly Gly Phe Ala Lys Tyr Trp Gln Ala Phe Arg Gln Tyr
690 695 700

Pro Arg Leu Gln Gly Gly Phe Val Trp Asp Trp Val Asp Gln Ser Leu
705 710 715 720

Ile Lys Tyr Asp Glu Asn Gly Asn Pro Trp Ser Ala Tyr Gly Gly Asp
725 730 735

Phe Gly Asp Thr Pro Asn Asp Arg Gln Phe Cys Met Asn Gly Leu Val
740 745 750

Phe Ala Asp Arg Thr Pro His Pro Ala Leu Thr Glu Ala Lys His Gln
755 760 765

Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln Thr Ile Glu Val Thr

770

775

780

Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu Leu Leu His Trp Met
 785 790 795 800

Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly Glu Val Pro Leu Asp
 805 810 815

Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu Pro Glu Leu Pro Gln
 820 825 830

Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val Arg Val Val Gln Pro
 835 840 845

Asn Ala Thr Ala Trp Ser Glu Ala Gly His Ile Ser Ala Trp Gln Gln
 850 855 860

Trp Arg Leu Ala Glu Asn Leu Ser Val Thr Leu Pro Ala Ala Ser His
 865 870 875 880

Ala Ile Pro His Leu Thr Thr Ser Glu Met Asp Phe Cys Ile Glu Leu
 885 890 895

Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln Ser Gly Phe Leu Ser Gln
 900 905 910

Met Trp Ile Gly Asp Lys Lys Gln Leu Leu Thr Pro Leu Arg Asp Gln
 915 920 925

Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile Gly Val Ser Glu Ala Thr
 930 935 940

Arg Ile Asp Pro Asn Ala Trp Val Glu Arg Trp Lys Ala Ala Gly His
 945 950 955 960

Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys Thr Ala Asp Thr Leu Ala
 965 970 975

Asp Ala Val Leu Ile Thr Thr Ala His Ala Trp Gln His Gln Gly Lys
 980 985 990

Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile Asp Gly Ser Gly Gln
995 1000 1005

Met Ala Ile Thr Val Asp Val Glu Val Ala Ser Asp Thr Pro His
1010 1015 1020

Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln Val Ala Glu
1025 1030 1035

Arg Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr Pro Asp
1040 1045 1050

Arg Leu Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu Ser
1055 1060 1065

Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg
1070 1075 1080

Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly
1085 1090 1095

Asp Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln Gln Gln Leu Met
1100 1105 1110

Glu Thr Ser His Arg His Leu Leu His Ala Glu Glu Gly Thr Trp
1115 1120 1125

Leu Asn Ile Asp Gly Phe His Met Gly Ile Gly Gly Asp Asp Ser
1130 1135 1140

Trp Ser Pro Ser Val Ser Ala Glu Leu Gln Leu Ser Ala Gly Arg
1145 1150 1155

Tyr His Tyr Gln Leu Val Trp Cys Gln Lys
1160 1165

<210> 39

<211> 3504

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 39

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggg	360
ggaagctcgt tcacgctaac gaacaagaac gtcattcttc tagcgggact tggcgggtatc	420
ggcctagaca cgtcgaagga actactaaag cgtgaccgga tagtcctcca acgtcgcgat	480
tgggagaacc cgggcgtaac gcaactaaac cgtcttgctg cgacccgcc gtttgctcgc	540
tggcgtaact cggaggaggc gcgaacggat cgtccgtcgc aacaactacg ttcgctcaac	600
ggggagtggc gcttcgctg gttcccggcg ccggaggcgg taccggagtc gtggctcgag	660
tgcgatctac cggaggcgga cacggctcgtc gtaccgtcga actggcaaact gcacggttac	720
gacgcgccga tatacacgaa cgtcacgtac ccgataacgg taaacccgcc gttcgtccc	780
acggagaacc cgacgggggtg ctactcgcta acgttcaacg ttgacgagtc gtggttgcaa	840
gagggtcaaa cgcgtatcat attcgacggg gtaaaactcg cgttccacct gtggtgcaac	900
gggcgctggg tagggtacgg ccaagactcg cgtctaccgt cggagttcga cctatcggcg	960
ttcctacgag cgggtgagaa ccggctagcg gtcattggtc tacgttggtc ggacggttcg	1020
tacctcgagg accaagacat gtggcgaatg tcgggtatct tccgcgacgt atcgctccta	1080
cacaagccga cgacgcaaact ctcggacttc cacgtcgcga cgcgtttcaa cgacgatttc	1140
tcgcgggcag tcctagaggc ggaggtccaa atgtgcgggg agctacgtga ctacctcgt	1200
gtcacggtat cgctctggca aggtgagacg caagtagcgt cgggtacggc gccgttcggc	1260
ggtgagatca tcgacgagcg tgggtgggtac gcggaccgtg taacgctacg tctaaacgtc	1320
gagaacccga agctctgggtc ggcggagatc ccgaacctat accgtgcggg cgtcgagcta	1380
catacggcgg acgggacgct aatagaggcg gaagcgtgcg acgtcgggtt tcgagaggtt	1440

cgtatagaga acgggctgct acttctaaac gggaagccgt tgctcatagc tgggtgtcaac	1500
cgtcacgagc accacccgct acacgggtcaa gtaatggacg agcaaacgat ggtacaagac	1560
atcctactaa tgaagcagaa caacttcaac gcgggtacgct gttcgcatta cccgaaccat	1620
ccgttgtggt acacgctttg cgaccgatac ggtctatacg tcgtagacga ggcgaacata	1680
gagacgcacg ggatgggtacc gatgaatcgc ctaacggacg acccgcgttg gctaccggcg	1740
atgtcggagc gagtcacgcg tatgggtccaa cgggaccgta accacccgctc ggtaataatc	1800
tggtcgctag gcaacgaatc ggggcacggg gcgaaccacg acgcgctata ccgttggatc	1860
aagtcggtag acccgctcgcg tccgggtacaa tacgaagggtg gcggtgcgga cacgacggcg	1920
acggacatca tctgcccgat gtacgcgcgc gtcgacgaag accaaccgtt cccggcggta	1980
ccgaagtggc cgatcaagaa gtggctctcg ttgccgggtg aaacgcgtcc gttgatactt	2040
tgcgagtacg cgcacgcgat gggcaactcg ttgggtgggt tcgcgaagta ctggcaggcg	2100
ttccgtcaat acccgctct acaggtggg ttcgtctggg actgggtaga ccaatcgcta	2160
atcaagtacg acgagaacgg caaccgtgg tcggcgtagc gtggggactt cggggacacg	2220
ccgaacgacc gccaatctg tatgaacggc ctagtcttcg cggaccgaac gccgcacccg	2280
gcgttgacgg aggcaagca tcaacaacaa ttcttccaat tccgtctatc ggggcaaacg	2340
atcgaggtaa cgtcggagta cttgttccgg cactcggaca acgagctact acactggatg	2400
gtagcactag acggcaagcc gtagcgtcg ggagaagtcc ctttggacgt cgcgccgcaa	2460
ggtaagcaac taatcgagct accggagcta ccgcaaccgg agtcggcggg tcaactgtgg	2520
ttgacgggcc gtgtcgttca accgaacgcg acggcgtggc cggaggcggg tcacatctcg	2580
gcgtggcagc agtggcgtct agcggagaac ctctcggta cgctaccggc ggcgtcgcac	2640
gcgataccgc atctaacgac gtcggagatg gacttctgca tcgagttggg gaacaagagg	2700
tggcagttca accgtcaatc gggattccta tcgcaaatgt ggataggtga caagaagcaa	2760
ctactaacgc cgctacgtga tcagttcacg cgtgctccgc tagacaacga cataggtgtt	2820
tcggaggcga cgcgtataga cccgaacgcg tgggtggagc ggtggaaggc ggcggggcac	2880
taccaagcgg aggcggcgct actacagtgc acggcggaca cgctagcggg cgcggtattg	2940
atcacgacgg cgcacgcgtg gcaacaccag gggaagacgc tattcatctc gcgtaagacg	3000
taccgtatcg acggttcggg ccaaattggc atcacggtcg acgtagaggt agcgtcggac	3060

acgccgcatc cggcgcgcat cggctctaaac tgccaactag cgcaagtagc ggagcgtgta 3120
aactggctag ggctagggcc gcaagagaac tatccggacc gcctaacggc ggcgtgcttc 3180
gaccgttggg acctaccgct ttcggacatg tataccccgt acgtcttccc gtcggagaac 3240
gggttgaggt gcgggacgcg cgagctaaac tacgggccgc accagtggcg aggggacttc 3300
caattcaaca tatcgcgtta ctcgcaacaa caactaatgg agacgtcgca ccgtcaccta 3360
ctacacgcgg aggaggggac gtggctaaac atcgacgggt tccacatggg cataggtggg 3420
gacgactcgt ggtcgccgtc ggtctcggcg gagctccaac tctcggcggg tcgttaccat 3480
taccaactag tttggtgcca gaag 3504

<210> 40

<211> 478

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 40

Asp Pro Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro
1 5 10 15

Ser Phe Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro
20 25 30

Arg Gly Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly
35 40 45

Pro Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile
50 55 60

Ala Thr Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys
65 70 75 80

Ala Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly
85 90 95

Glu Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val
100 105 110

Arg Ser Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Gln Gln Leu Met
115 120 125

Lys Gln Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly
130 135 140

Val Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser
145 150 155 160

Val Ala Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro
165 170 175

Thr Thr Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu
180 185 190

Asn Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu Val
195 200 205

Phe Glu Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu
210 215 220

Arg Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg
225 230 235 240

Arg Arg Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu
245 250 255

Lys Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu
260 265 270

Lys Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr
275 280 285

Thr Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp
290 295 300

Gln Tyr Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser
305 310 315 320

Ser Arg Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser
325 330 335

Asn Trp Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln
340 345 350

Ile His Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val
355 360 365

Gln Gln Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro
370 375 380

Gly Pro Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp
385 390 395 400

Asn Thr Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg
405 410 415

Leu Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu
420 425 430

Phe Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly
435 440 445

Ser Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala
450 455 460

Thr Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val
465 470 475

<210> 41

<211> 1434

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 41

gacctacca gactgcaggg catcaactgc ggccctgact tcactccttc tttcgcaaac 60

ctgggcagaa ccaccctgag caccagaggc cctcctagag gtggtcccgg cggagaactc 120

cccaggggtc ctgccggcct gggccctaga cgctcccagc aaggtcctag aaaggagcca	180
aggaagatca tcgccaccgt gctgatgacc gaggacatca agctgaacaa agctgagaag	240
gcctggaagc ctagcagcaa gagaaccgcc gccgacaagg acagaggcga ggaggacgcc	300
gacggatcca agaccagga cctgttcaga agagtgagaa gcatcctcaa caagctgacc	360
cctcagatgt tccagcagct gatgaagcag gtgacgcagc tcgccatcga caccgaggag	420
agactgaagg gcgtgatcga cctgatcttt gagaaggcta tctcagagcc taacttcagc	480
gtggcctacg ccaacatgtg ccgttgccctg atggcattga aggtgccaac caccgagaag	540
cctactgtga ccgtcaatth ccgtaaactg ctgctgaacc ggtgccagaa agagttcgag	600
aaggataagg acgacgacga ggtcttcgag aagaaacaga aagaaatgga cgaggccgcc	660
accgcagagg aaagggggccg attaaaggag gagctggagg aggccagaga catcgccaga	720
cggcgttctc tgggcaacat caagttcata ggtgagctgt tcaagctaaa gatgctcacc	780
gaggccataa tgcacgactg cgtggtgaag ctactgaaga accacgacga agaaagcctg	840
gagtgcctgt gcagactgct gaccaccatc ggcaaggacc tggacttcga gaaggcaaag	900
cctcgaatgg accagtactt caaccagatg gagaagatta tcaaggagaa gaagaccagc	960
agcagaatca gattcatgct gcaggacgta ctggacctgc gcggaagcaa ctgggtgcc	1020
aggagagggg accaaggacc aaagaccatc gaccagatcc acaaggaagc ggagatggag	1080
gagcacagag agcacataaa ggtgcagcag cttatggcca agggcagcga caagcgaaga	1140
ggcggcccgcc ccggacctcc tatcagcaga ggccttcctc tggtagacga cggcggctgg	1200
aacaccgtgc ctatctctaa gggctccaga cctatcgaca cttcccgtct taccaagatc	1260
accaagccag gatctattga cagcaacaac cagctgttcg ccccaggagg aagacttagc	1320
tggggcaagg gaagttccgg cggatccggc gccaaagcctt ccgacgccgc cagcgaggct	1380
gccagacctg ccaccagcac cttgaaccgc ttttccgctc tgcagcaagc tgtg	1434

<210> 42

<211> 977

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 42

Ile Phe Ala Ser Met Gln Lys Pro Glu Gly Leu Pro His Ile Ser Asp
1 5 10 15

Val Val Leu Asp Lys Ala Asn Lys Thr Pro Leu Arg Pro Leu Asp Pro
20 25 30

Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe
35 40 45

Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly
50 55 60

Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg
65 70 75 80

Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr
85 90 95

Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp
100 105 110

Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu
115 120 125

Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser
130 135 140

Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys Gln
145 150 155 160

Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile
165 170 175

Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala
180 185 190

Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr
195 200 205

Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg
210 215 220

Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Glu Val Phe Glu
225 230 235 240

Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly
245 250 255

Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg
260 265 270

Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met
275 280 285

Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu Lys Asn
290 295 300

His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile
305 310 315 320

Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr
325 330 335

Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg
340 345 350

Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp
355 360 365

Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His
370 375 380

Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln
385 390 395 400

Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro
405 410 415

Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr

420

425

430

Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr
 435 440 445

Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala
 450 455 460

Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly
 465 470 475 480

Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser
 485 490 495

Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu Ser
 500 505 510

Thr Asp Asn Arg Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu Arg
 515 520 525

Gly Glu Lys Ala Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu Arg
 530 535 540

Gly Gly Asp Arg Gly Asp Arg Leu Asp Arg Ala Arg Thr Pro Ala Thr
 545 550 555 560

Lys Arg Ser Phe Ser Lys Glu Val Glu Glu Arg Ser Arg Glu Arg Pro
 565 570 575

Ser Gln Pro Glu Gly Leu Arg Lys Ala Ala Ser Leu Thr Glu Asp Arg
 580 585 590

Asp Arg Gly Arg Asp Ala Val Lys Arg Glu Ala Ala Leu Pro Pro Val
 595 600 605

Ser Pro Leu Lys Ala Ala Leu Ser Glu Glu Glu Leu Glu Lys Lys Ser
 610 615 620

Lys Ala Ile Ile Glu Glu Tyr Leu His Leu Asn Asp Met Lys Glu Ala
 625 630 635 640

Val Gln Cys Val Gln Glu Leu Ala Ser Pro Ser Leu Leu Phe Ile Phe
645 650 655

Val Arg His Gly Val Glu Ser Thr Leu Glu Arg Ser Ala Ile Ala Arg
660 665 670

Glu His Met Gly Gln Leu Leu His Gln Leu Leu Cys Ala Gly His Leu
675 680 685

Ser Thr Ala Gln Tyr Tyr Gln Gly Leu Tyr Glu Ile Leu Glu Leu Ala
690 695 700

Glu Asp Met Glu Ile Asp Ile Pro His Val Trp Leu Tyr Leu Ala Glu
705 710 715 720

Leu Val Thr Pro Ile Leu Gln Glu Gly Gly Val Pro Met Gly Glu Leu
725 730 735

Phe Arg Glu Ile Thr Lys Pro Leu Arg Pro Leu Gly Lys Ala Ala Ser
740 745 750

Leu Leu Leu Glu Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro Lys
755 760 765

Lys Val Gly Thr Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu Phe
770 775 780

Leu Pro Glu Gly Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys Val
785 790 795 800

Glu Tyr Thr Leu Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala Leu
805 810 815

Pro Ser Glu Glu Leu Asn Arg Gln Leu Glu Lys Leu Leu Lys Glu Gly
820 825 830

Ser Ser Asn Gln Arg Val Phe Asp Trp Ile Glu Ala Asn Leu Ser Glu
835 840 845

Gln Gln Ile Val Ser Asn Thr Leu Val Arg Ala Leu Met Thr Ala Val

850

855

860

Cys Tyr Ser Ala Ile Ile Phe Glu Thr Pro Leu Arg Val Asp Val Ala
 865 870 875 880

Val Leu Lys Ala Arg Ala Lys Leu Leu Gln Lys Tyr Leu Cys Asp Glu
 885 890 895

Gln Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Leu Val Val Thr
 900 905 910

Leu Glu Gln Pro Pro Asn Leu Leu Arg Met Phe Phe Asp Ala Leu Tyr
 915 920 925

Asp Glu Asp Val Val Lys Glu Asp Ala Phe Tyr Ser Trp Glu Ser Ser
 930 935 940

Lys Asp Pro Ala Glu Gln Gln Gly Lys Gly Val Ala Leu Lys Ser Val
 945 950 955 960

Thr Ala Phe Phe Lys Trp Leu Arg Glu Ala Glu Glu Glu Ser Asp His
 965 970 975

Asn

<210> 43

<211> 2931

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 43

atattcgcga gcatgcagaa gcctgaaggt ctgcctcaca tcagcgacgt ggtgctggac 60

aaggccaaca agaccctct tagacctctg gaccctacca gactgcaggg catcaactgc 120

ggcctgact tcactccttc tttcgcaaac ctgggcagaa ccaccctgag caccagaggc 180

cctcctagag gtgggtcccg cggagaactc cccaggggtc ctgccggcct gggcctaga 240

cgctcccagc aaggtcctag aaaggagcca aggaagatca tcgccaccgt gctgatgacc	300
gaggacatca agctgaacaa agctgagaag gcctggaagc ctagcagcaa gagaaccgcc	360
gccgacaagg acagaggcga ggaggacgcc gacggatcca agaccagga cctgttcaga	420
agagtgagaa gcatcctcaa caagctgacc cctcagatgt tccagcagct gatgaagcag	480
gtgacgcagc tcgccatcga caccgaggag agactgaagg gcgtgatcga cctgatcttt	540
gagaaggcta tctcagagcc taacttcagc gtggcctacg ccaacatgtg ccgttgcctg	600
atggcattga aggtgccaac caccgagaag cctactgtga ccgtcaattt ccgtaaactg	660
ctgctgaacc ggtgccagaa agagttcgag aaggataagg acgacgacga ggtcttcgag	720
aagaaacaga aagaaatgga cgaggccgcc accgcagagg aaaggggccg attaaaggag	780
gagctggagg aggccagaga catcgccaga cggcgttctc tgggcaacat caagttcata	840
ggtgagctgt tcaagctaaa gatgctcacc gaggccataa tgcacgactg cgtggtgaag	900
ctactgaaga accacgacga agaaagcctg gagtgcctgt gcagactgct gaccaccatc	960
ggcaaggacc tggacttcga gaaggcaaag cctcgaatgg accagtactt caaccagatg	1020
gagaagatta tcaaggagaa gaagaccagc agcagaatca gattcatgct gcaggacgta	1080
ctggacctgc gcggaagcaa ctgggtgcc aaggagagggg accaaggacc aaagaccatc	1140
gaccagatcc acaaggaagc ggagatggag gagcacagag agcacataaa ggtgcagcag	1200
cttatggcca agggcagcga caagcgaaga ggcgggccgc ccggacctcc tatcagcaga	1260
ggccttcctc tggtagacga cggcggctgg aacaccgtgc ctatctctaa gggctccaga	1320
cctatcgaca cttcccgtct taccaagatc accaagccag gatctattga cagcaacaac	1380
cagctgttcg ccccaggagg aagacttagc tggggcaagg gaagttccgg cggatccggc	1440
gccaagcctt ccgacgccgc cagcgaggct gccagacctg ccaccagcac cttgaaccgc	1500
ttttccgctc tgcagcaagc tgtgccgacc gagagcaccg acaaccggcg agttgtgcag	1560
agaagcagcc tgagcagaga gaggggagcag aaagccggcg atagaggtga ccgactggag	1620
agaagcgaga gaggaggtga tagaggcgac cggcttgaca gagccagaac ccctgccaca	1680
aagcgatcgt tcagcaagga ggtcgaagag aggtccaggg agcgccctag ccagccggaa	1740
ggactcagaa aggcagccag tctaacagag gaccgcgaca ggggaagaga cgccgtgaag	1800
agggaggccg cactgcctcc tgtgagccct ctgaaggcgg cactgagcga agaagaactt	1860

gagaagaaga gtaaggcgat catcgaggag tacctgcacc tgaacgacat gaaggaggcc 1920
 gtgcagtgcg tgcaagagct cgcgagccca tcaactgctgt tcactcttcgt ccggcacggc 1980
 gtggagtcca cactggaaag atctgccatt gctagggagc atatgggcca gttgttgcac 2040
 caattgcttt gcgccggcca cctaagtact gccagtgact atcagggttt atacgagatc 2100
 ctcgaactgg ccgaagacat ggagatcgac atccctcacg tgtggctgta cctcgcagaa 2160
 ctggtgaccc ctatcctgca ggagggcggc gttccaatgg gtgagttggt cagagaaatc 2220
 acaaagccac tgcgcccact gggcaaggcg gctagtctcc ttctggagat tctcggcctg 2280
 ctctgtaaga gcatgggtcc aaagaagggtt ggcaccctgt ggagggaagc tggactctcc 2340
 tggaaggaat tcctccctga gggtcaggac atcggcgcct tcgtggccga acagaagggtt 2400
 gagtacaccc tgggagagga atcggaagcg ccgggtcagc gggctctgcc gagtgaggag 2460
 ctcaacagac aactcgagaa gctcctcaag gagggatcct caaatcagag agtggttcgac 2520
 tggatcgagg ccaatctcag cgagcagcaa atcgtgagca acacgttggt tcgcgctctc 2580
 atgacagccg tgtgtacttc agccattatc ttcgagacgc ctctccgctt ggacgtggca 2640
 gtgctcaagg cccgcgctaa gctgttacag aagtacctgt gcgacgagca gaaggaactg 2700
 caggccctgt acgccctgca agctctggtc gtcacactcg agcagcctcc taacctgctg 2760
 agaatgttct tcgacgcatt gtacgacgag gacgttggtga aggaagacgc tttctacagc 2820
 tgggagtcta gtaaggatcc tgcggaacaa caaggcaagg gcgttgctct aaagagcgtg 2880
 accgccttct tcaagtggct gcgagaggct gaggaagagt cggaccacaa c 2931

<210> 44

<211> 977

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 44

Ile	Phe	Ala	Ser	Met	Gln	Lys	Pro	Glu	Gly	Leu	Pro	His	Ile	Ser	Asp
1				5					10					15	

Val	Val	Leu	Asp	Lys	Ala	Asn	Lys	Thr	Pro	Leu	Arg	Pro	Leu	Asp	Pro
			20					25					30		

Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe
35 40 45

Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly
50 55 60

Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg
65 70 75 80

Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr
85 90 95

Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp
100 105 110

Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu
115 120 125

Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser
130 135 140

Ile Ala Asn Lys Ala Thr Pro Gln Met Ala Gln Gln Leu Met Lys Gln
145 150 155 160

Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile
165 170 175

Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala
180 185 190

Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr
195 200 205

Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg
210 215 220

Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe Glu
225 230 235 240

Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly
245 250 255

Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg
260 265 270

Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met
275 280 285

Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu Lys Asn
290 295 300

His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile
305 310 315 320

Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr
325 330 335

Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg
340 345 350

Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp
355 360 365

Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His
370 375 380

Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln
385 390 395 400

Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro
405 410 415

Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr
420 425 430

Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr
435 440 445

Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala
450 455 460

Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly
465 470 475 480

Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser
485 490 495

Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu Ser
500 505 510

Thr Asp Asn Arg Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu Arg
515 520 525

Gly Glu Lys Ala Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu Arg
530 535 540

Gly Gly Asp Arg Gly Asp Arg Leu Asp Arg Ala Arg Thr Pro Ala Thr
545 550 555 560

Lys Arg Ser Phe Ser Lys Glu Val Glu Glu Arg Ser Arg Glu Arg Pro
565 570 575

Ser Gln Pro Glu Gly Leu Arg Lys Ala Ala Ser Leu Thr Glu Asp Arg
580 585 590

Asp Arg Gly Arg Asp Ala Val Lys Arg Glu Ala Ala Leu Pro Pro Val
595 600 605

Ser Pro Leu Lys Ala Ala Leu Ser Glu Glu Glu Leu Glu Lys Lys Ser
610 615 620

Lys Ala Ile Ile Glu Glu Tyr Leu His Leu Asn Asp Met Lys Glu Ala
625 630 635 640

Val Gln Cys Val Gln Glu Leu Ala Ser Pro Ser Leu Leu Phe Ile Phe
645 650 655

Val Arg His Gly Val Glu Ser Thr Leu Glu Arg Ser Ala Ile Ala Arg
660 665 670

Glu His Met Gly Gln Leu Leu His Gln Leu Leu Cys Ala Gly His Leu
675 680 685

Ser Thr Ala Gln Tyr Tyr Gln Gly Leu Tyr Glu Ile Leu Glu Leu Ala
690 695 700

Glu Asp Met Glu Ile Asp Ile Pro His Val Trp Leu Tyr Leu Ala Glu
705 710 715 720

Leu Val Thr Pro Ile Leu Gln Glu Gly Gly Val Pro Met Gly Glu Leu
725 730 735

Phe Arg Glu Ile Thr Lys Pro Leu Arg Pro Leu Gly Lys Ala Ala Ser
740 745 750

Leu Leu Leu Glu Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro Lys
755 760 765

Lys Val Gly Thr Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu Phe
770 775 780

Leu Pro Glu Gly Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys Val
785 790 795 800

Glu Tyr Thr Leu Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala Leu
805 810 815

Pro Ser Glu Glu Leu Asn Arg Gln Leu Glu Lys Leu Leu Lys Glu Gly
820 825 830

Ser Ser Asn Gln Arg Val Phe Asp Trp Ile Glu Ala Asn Leu Ser Glu
835 840 845

Gln Gln Ile Val Ser Asn Thr Leu Val Arg Ala Leu Met Thr Ala Val
850 855 860

Cys Tyr Ser Ala Ile Ile Phe Glu Thr Pro Leu Arg Val Asp Val Ala
865 870 875 880

Val Leu Lys Ala Arg Ala Lys Leu Leu Gln Lys Tyr Leu Cys Asp Glu
885 890 895

Gln Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Leu Val Val Thr
900 905 910

Leu Glu Gln Pro Pro Asn Leu Leu Arg Met Phe Phe Asp Ala Leu Tyr
915 920 925

Asp Glu Asp Val Val Lys Glu Asp Ala Phe Tyr Ser Trp Glu Ser Ser
930 935 940

Lys Asp Pro Ala Glu Gln Gln Gly Lys Gly Val Ala Leu Lys Ser Val
945 950 955 960

Thr Ala Phe Phe Lys Trp Leu Arg Glu Ala Glu Glu Glu Ser Asp His
965 970 975

Asn

<210> 45
<211> 2931
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 45
atattcgcga gcatgcagaa gcctgaaggt ctgcctcaca tcagcgacgt ggtgctggac 60
aaggccaaca agaccctct tagacctctg gaccctacca gactgcaggg catcaactgc 120
ggccctgact tcactccttc tttcgcaaac ctgggcagaa ccaccctgag caccagaggc 180
cctcctagag gtgggtcccg cgagagaactc cccaggggtc ctgccggcct gggccctaga 240
cgctcccagc aaggtcctag aaaggagcca aggaagatca tcgccaccgt gctgatgacc 300
gaggacatca agctgaacaa agctgagaag gcctggaagc ctagcagcaa gagaaccgcc 360
gccgacaagg acagaggcga ggaggacgcc gacggatcca agaccagga cctgttcaga 420
agagtgagaa gcatcgcgaa caaggcgacc cctcagatgg cgcagcagct gatgaagcag 480
gtgacgcagc tcgccatcga caccgaggag agactgaagg gcgtgacga cctgatcttt 540

gagaaggcta tctcagagcc taacttcagc gtggcctacg ccaacatgtg ccgttgcctg	600
atggcattga aggtgccaac caccgagaag cctactgtga ccgtcaattt ccgtaaactg	660
ctgctgaacc ggtgccagaa agagttcgag aaggataagg acgacgacga ggtcttcgag	720
aagaaacaga aagaaatgga cgaggccgcc accgcagagg aaaggggccc attaaaggag	780
gagctggagg aggccagaga catcgccaga cggcgttctc tgggcaacat caagttcata	840
ggtgagctgt tcaagctaaa gatgctcacc gaggccataa tgcacgactg cgtggtgaag	900
ctactgaaga accacgacga agaaagcctg gagtgcctgt gcagactgct gaccaccatc	960
ggcaaggacc tggacttcga gaaggcaaag cctcgaatgg accagtactt caaccagatg	1020
gagaagatta tcaaggagaa gaagaccagc agcagaatca gattcatgct gcaggacgta	1080
ctggacctgc gcggaagcaa ctgggtgccca aggagagggg accaaggacc aaagaccatc	1140
gaccagatcc acaaggaagc ggagatggag gagcacagag agcacataaa ggtgcagcag	1200
cttatggcca agggcagcga caagcgaaga ggcggcccgcc ccggacctcc tatcagcaga	1260
ggccttcctc tggtagacga cggcggctgg aacaccgtgc ctatctctaa gggctccaga	1320
cctatcgaca cttcccgtct taccaagatc accaagccag gatctattga cagcaacaac	1380
cagctgttcg ccccaggagg aagacttagc tggggcaagg gaagttccgg cggatccggc	1440
gccaagcctt ccgacgccgc cagcgaggct gccagacctg ccaccagcac cttgaaccgc	1500
ttttccgctc tgcagcaagc tgtgccgacc gagagcaccg acaaccggcg agttgtgcag	1560
agaagcagcc tgagcagaga gaggggagcga aaagccggcg atagaggtga ccgactggag	1620
agaagcgaga gaggaggtga tagaggcgac cggcttgaca gagccagaac ccctgccaca	1680
aagcgatcgt tcagcaagga ggtcgaagag aggtccaggg agcgccctag ccagccggaa	1740
ggactcagaa aggcagccag tctaacagag gaccgcgaca ggggaagaga cgccgtgaag	1800
agggaggccg cactgcctcc tgtgagccct ctgaaggcgg cactgagcga agaagaactt	1860
gagaagaaga gtaaggcgat catcgaggag tacctgcacc tgaacgacat gaaggaggcc	1920
gtgcagtgcg tgcaagagct cgcgagccca tctactgtgt tcatcttcgt ccggcacggc	1980
gtggagtcca cactggaaag atctgccatt gctagggagc atatgggcca gttgttgac	2040
caattgcttt gcgccggcca cctaagtact gccagtact atcagggttt atacgagatc	2100
ctcgaactgg ccgaagacat ggagatcgac atccctcacg tgtggctgta cctcgagaa	2160

ctggtgaccc ctatcctgca ggagggcggc gttccaatgg gtgagttggt cagagaaatc 2220
 acaaagccac tgcgcccact gggcaaggcg gctagtctcc ttctggagat tctcggcctg 2280
 ctctgtaaga gcatgggtcc aaagaagggtg ggcaccctgt ggaggggaagc tggactctcc 2340
 tggaaggaat tcctccctga gggtcaggac atcggcgcct tcgtggccga acagaaggtt 2400
 gagtacaccc tgggagagga atcggaagcg ccgggtcagc gggctctgcc gagtgaggag 2460
 ctcaacagac aactcgagaa gctcctcaag gagggatcct caaatcagag agtgttcgac 2520
 tggatcgagg ccaatctcag cgagcagcaa atcgtgagca acacgttggt tcgcgctctc 2580
 atgacagccg tgtgctactc agccattatc ttcgagacgc ctctccgcgt ggacgtggca 2640
 gtgctcaagg cccgcgctaa gctgttacag aagtacctgt gcgacgagca gaaggaactg 2700
 caggccctgt acgccctgca agctctggtc gtcacactcg agcagcctcc taacctgctg 2760
 agaatgttct tcgacgcatt gtacgacgag gacgttgtga aggaagacgc tttctacagc 2820
 tgggagtcta gtaaggatcc tgcggaacaa caaggcaagg gcgttgctct aaagagcgtg 2880
 accgccttct tcaagtggct gcgagaggct gaggaagagt cggaccacaa c 2931

<210> 46

<211> 478

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 46

Asp Pro Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro
 1 5 10 15

Ser Phe Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro
 20 25 30

Arg Gly Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly
 35 40 45

Pro Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile
 50 55 60

Ala Thr Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys
65 70 75 80

Ala Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly
85 90 95

Glu Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val
100 105 110

Arg Ser Ile Ala Asn Lys Ala Thr Pro Gln Met Ala Gln Gln Leu Met
115 120 125

Lys Gln Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly
130 135 140

Val Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser
145 150 155 160

Val Ala Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro
165 170 175

Thr Thr Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu
180 185 190

Asn Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu Val
195 200 205

Phe Glu Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu
210 215 220

Arg Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg
225 230 235 240

Arg Arg Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu
245 250 255

Lys Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu
260 265 270

Lys Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr

275

280

285

Thr Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp
 290 295 300

Gln Tyr Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser
 305 310 315 320

Ser Arg Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser
 325 330 335

Asn Trp Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln
 340 345 350

Ile His Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val
 355 360 365

Gln Gln Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro
 370 375 380

Gly Pro Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp
 385 390 395 400

Asn Thr Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg
 405 410 415

Leu Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu
 420 425 430

Phe Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly
 435 440 445

Ser Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala
 450 455 460

Thr Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val
 465 470 475

<210> 47

<211> 1434

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 47

gaccctacca gactgcaggg catcaactgc ggccctgact tcactccttc tttcgcaaac	60
ctgggcagaa ccaccctgag caccagaggc cctcctagag gtggtcccgg cggagaactc	120
cccaggggtc ctgccggcct gggccctaga cgctcccagc aaggtcctag aaaggagcca	180
aggaagatca tcgccaccgt gctgatgacc gaggacatca agctgaacaa agctgagaag	240
gcctggaagc ctagcagcaa gagaaccgcc gccgacaagg acagaggcga ggaggacgcc	300
gacggatcca agaccagga cctgttcaga agagtgagaa gcatcgccaa caaggcgacc	360
cctcagatgg cgcagcagct gatgaagcag gtgacgcagc tcgccatcga caccgaggag	420
agactgaagg gcgtgatcga cctgatcttt gagaaggcta tctcagagcc taacttcagc	480
gtggcctacg ccaacatgtg ccgttgcttg atggcattga aggtgccaac caccgagaag	540
cctactgtga ccgtcaatth ccgtaaaactg ctgctgaacc ggtgccagaa agagttcgag	600
aaggataagg acgacgacga ggtcttcgag aagaaacaga aagaaatgga cgaggccgcc	660
accgcagagg aaaggggccc attaaaggag gagctggagg aggccagaga catcgccaga	720
cggcgttctc tgggcaacat caagttcata ggtgagctgt tcaagctaaa gatgctcacc	780
gaggccataa tgcacgactg cgtggtgaag ctactgaaga accacgacga agaaagcctg	840
gagtgcctgt gcagactgct gaccaccatc ggcaaggacc tggacttcga gaaggcaaag	900
cctcgaatgg accagtactt caaccagatg gagaagatta tcaaggagaa gaagaccagc	960
agcagaatca gattcatgct gcaggacgta ctggacctgc gcggaagcaa ctgggtgcca	1020
aggagagggg accaaggacc aaagaccatc gaccagatcc acaaggaagc ggagatggag	1080
gagcacagag agcacataaa ggtgcagcag cttatggcca agggcagcga caagcgaaga	1140
ggcggcccg cccggacctc tatcagcaga ggccttcctc tggtagacga cggcggctgg	1200
aacaccgtgc ctatctctaa gggctccaga cctatcgaca cttcccgtct taccaagatc	1260
accaagccag gatctattga cagcaacaac cagctgttcg cccagaggag aagacttagc	1320
tggggcaagg gaagttccgg cggatccggc gccaaacctt ccgacgccgc cagcgaggct	1380

gccagacctg ccaccagcac cttgaaccgc ttttccgctc tgcagcaagc tgtg

1434

<210> 48

<211> 635

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 48

Asn Pro Ser Ala Pro Ser Tyr Pro Met Ala Ser Leu Tyr Val Gly Asp
1 5 10 15

Leu His Pro Asp Val Thr Glu Ala Met Leu Tyr Glu Lys Phe Ser Pro
20 25 30

Ala Gly Pro Ile Leu Ser Ile Arg Val Cys Arg Asp Met Ile Thr Arg
35 40 45

Arg Ser Leu Gly Tyr Ala Tyr Val Asn Phe Gln Gln Pro Ala Asp Ala
50 55 60

Glu Arg Ala Leu Asp Thr Met Asn Phe Asp Val Ile Lys Gly Lys Pro
65 70 75 80

Val Arg Ile Met Trp Ser Gln Arg Asp Pro Ser Leu Arg Lys Ser Gly
85 90 95

Val Gly Asn Ile Phe Ile Lys Asn Leu Asp Lys Ser Ile Asp Asn Lys
100 105 110

Ala Leu Tyr Asp Thr Phe Ser Ala Phe Gly Asn Ile Leu Ser Cys Lys
115 120 125

Val Val Cys Asp Glu Asn Gly Ser Lys Gly Tyr Gly Phe Val His Phe
130 135 140

Glu Thr Gln Glu Ala Ala Glu Arg Ala Ile Glu Lys Met Asn Gly Met
145 150 155 160

Leu Leu Asn Asp Arg Lys Val Phe Val Gly Arg Phe Lys Ser Arg Lys
165 170 175

Glu Arg Glu Ala Glu Leu Gly Ala Arg Ala Lys Glu Phe Thr Asn Val
180 185 190

Tyr Ile Lys Asn Phe Gly Glu Asp Met Asp Asp Glu Arg Leu Lys Asp
195 200 205

Leu Phe Gly Lys Phe Gly Pro Ala Leu Ser Val Lys Val Met Thr Asp
210 215 220

Glu Ser Gly Lys Ser Lys Gly Phe Gly Phe Val Ser Phe Glu Arg His
225 230 235 240

Glu Asp Ala Gln Lys Ala Val Asp Glu Met Asn Gly Lys Glu Leu Asn
245 250 255

Gly Lys Gln Ile Tyr Val Gly Arg Ala Gln Lys Lys Val Glu Arg Gln
260 265 270

Thr Glu Leu Lys Arg Lys Phe Glu Gln Met Lys Gln Asp Arg Ile Thr
275 280 285

Arg Tyr Gln Gly Val Asn Leu Tyr Val Lys Asn Leu Asp Asp Gly Ile
290 295 300

Asp Asp Glu Arg Leu Arg Lys Glu Phe Ser Pro Phe Gly Thr Ile Thr
305 310 315 320

Ser Ala Lys Val Met Met Glu Gly Gly Arg Ser Lys Gly Phe Gly Phe
325 330 335

Val Cys Phe Ser Ser Pro Glu Glu Ala Thr Lys Ala Val Thr Glu Met
340 345 350

Asn Gly Arg Ile Val Ala Thr Lys Pro Leu Tyr Val Ala Leu Ala Gln
355 360 365

Arg Lys Glu Glu Arg Gln Ala His Leu Thr Asn Gln Tyr Met Gln Arg
370 375 380

Met Ala Ser Val Arg Ala Val Pro Asn Pro Val Ile Asn Pro Tyr Gln
385 390 395 400

Pro Ala Pro Pro Ser Gly Tyr Phe Met Ala Ala Ile Pro Gln Thr Gln
405 410 415

Asn Arg Ala Ala Tyr Tyr Pro Pro Ser Gln Ile Ala Gln Leu Arg Pro
420 425 430

Ser Pro Arg Trp Thr Ala Gln Gly Ala Arg Pro His Pro Phe Gln Asn
435 440 445

Met Pro Gly Ala Ile Arg Pro Ala Ala Pro Arg Pro Pro Phe Ser Thr
450 455 460

Met Arg Pro Ala Ser Ser Gln Val Pro Arg Val Met Ser Thr Gln Arg
465 470 475 480

Val Ala Asn Thr Ser Thr Gln Thr Met Gly Pro Arg Pro Ala Ala Ala
485 490 495

Ala Ala Ala Ala Thr Pro Ala Val Arg Thr Val Pro Gln Tyr Lys Tyr
500 505 510

Ala Ala Gly Val Arg Asn Pro Gln Gln His Leu Asn Ala Gln Pro Gln
515 520 525

Val Thr Met Gln Gln Pro Ala Val His Val Gln Gly Gln Glu Pro Leu
530 535 540

Thr Ala Ser Met Leu Ala Ser Ala Pro Pro Gln Glu Gln Lys Gln Met
545 550 555 560

Leu Gly Glu Arg Leu Phe Pro Leu Ile Gln Ala Met His Pro Thr Leu
565 570 575

Ala Gly Lys Ile Thr Gly Met Leu Leu Glu Ile Asp Asn Ser Glu Leu
580 585 590

Leu His Met Leu Glu Ser Pro Glu Ser Leu Arg Ser Lys Val Asp Glu
595 600 605

Ala Val Ala Val Leu Gln Ala His Gln Ala Lys Glu Ala Ala Gln Lys
610 615 620

Ala Val Asn Ser Ala Thr Gly Val Pro Thr Val
625 630 635

<210> 49
<211> 1905
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 49
aaccacagcg cccaagcta ccccatggcc agcctgtacg tgggcgacct gcaccccgac 60
gttaccgagg ccatgctgta cgagaaatct tctcccgctg gcccacacct gagcatccgg 120
gtgtgccggg acatgattac ccggcgctct ctgggggtacg cctacgtgaa cttccagcag 180
cccgagacg ccgaacgggc cctggacacc atgaacttcg acgtgatcaa gggcaagccc 240
gtgcgggatca tgtggagcca gcgggaccca agcctgagaa agagcggcgt gggcaacatc 300
ttcatcaaga atctggacaa gagcatcgac aacaaggccc tgtacgacac cttcagcgcc 360
ttcggcaata tcttaagctg caaggtggtg tgcgacgaga acggcagcaa gggctacggc 420
ttcgtgcact tcgaaacaca agaagccgca gagcgggcca tcgagaagat gaacggcatg 480
ctgctgaacg accggaaggt gttcgtgggc cggttcaaga gccggaagga gcgggaggct 540
gaactcgggtg cccgggcca gaggttcacc aacgtgtaca tcaagaactt cggcgaggac 600
atggacgacg agcgggtgaa ggacctgttc ggcaagttcg gccccgccct gtctgtgaag 660
gtgatgaccg acgagagcgg caagtcgaag ggtttcggct ttgtgagctt cgagcggcac 720
gaggacgcac agaaggccgt ggacgaaatg aacggtaagg agcttaacgg taagcagatc 780
tacgttggcc gggcccagaa gaaggtggag cggcagaccg agctgaagcg gaagttcgag 840
cagatgaagc aggaccgat caccgggtac cagggcgtga acctatacgt gaagaacctg 900
gacgacggca tcgacgacga acggctgcgg aaggaattca gcccttcgg caccatcacc 960

agcgccaagg taatgatgga ggggtggccgg tcgaagggat tcggtttcgt gtgcttcagc 1020
 agccccgagg aggccaccaa ggccgttact gagatgaacg gccggatcgt agccacaaag 1080
 ccactttacg tcgccctggc ccaacgcaag gaggaacggc aggcccacct gaccaaccag 1140
 tacatgcagc ggatggccag cgtgcgggca gtgccaacc ccgtgatcaa cccctaccaa 1200
 cccgcaccac caagcggcta cttcatggcc gccatccctc agaccagaa ccgggccgcc 1260
 tactaccac ccagccagat cgcccagttg cggcctagtc cccgctggac ggctcaagga 1320
 gcacgacccc accccttcca gaacatgccc ggagccatca gaccagctgc ccctaggcca 1380
 cccttcagca ccatgcggcc cgcctcaagc caggtgcccc gggatgatgag caccagcgg 1440
 gtggccaaca ccagcacca gaccatgggc cctagaccag cagctgctgc ggctgctgcc 1500
 actcctgccg tccggaccgt gccacagtac aagtacgccg ccggcgttcg gaaccacag 1560
 cagcacctga acgcccagcc ccaggtgacc atgcagcagc ctgccgtgca cgtgcagggc 1620
 caggagcccc tgaccgccag catgttagca agcggccctc cacaggagca gaagcagatg 1680
 ctgggcgaaa ggctgtttcc cctgatccaa gcgatgcac ccaccctggc cggcaagatc 1740
 accggtatgc tgctggagat cgataatagc gagctgctgc acatgctgga gagccccgag 1800
 agtcttagat ctaagggtga cgaggccgtg gccgtgctcc aggtcacca ggcaaaggag 1860
 gccgcccaga aagccgtgaa ctccgccaca ggcgtgccca ccgtg 1905

<210> 50

<211> 757

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 50

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
 1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
 20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35

40

45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asn Pro Ser Ala Pro Ser
115 120 125

Tyr Pro Met Ala Ser Leu Tyr Val Gly Asp Leu His Pro Asp Val Thr
130 135 140

Glu Ala Met Leu Tyr Glu Lys Phe Ser Pro Ala Gly Pro Ile Leu Ser
145 150 155 160

Ile Arg Val Cys Arg Asp Met Ile Thr Arg Arg Ser Leu Gly Tyr Ala
165 170 175

Tyr Val Asn Phe Gln Gln Pro Ala Asp Ala Glu Arg Ala Leu Asp Thr
180 185 190

Met Asn Phe Asp Val Ile Lys Gly Lys Pro Val Arg Ile Met Trp Ser
195 200 205

Gln Arg Asp Pro Ser Leu Arg Lys Ser Gly Val Gly Asn Ile Phe Ile
210 215 220

Lys Asn Leu Asp Lys Ser Ile Asp Asn Lys Ala Leu Tyr Asp Thr Phe
225 230 235 240

Ser Ala Phe Gly Asn Ile Leu Ser Cys Lys Val Val Cys Asp Glu Asn
245 250 255

Gly Ser Lys Gly Tyr Gly Phe Val His Phe Glu Thr Gln Glu Ala Ala
260 265 270

Glu Arg Ala Ile Glu Lys Met Asn Gly Met Leu Leu Asn Asp Arg Lys
275 280 285

Val Phe Val Gly Arg Phe Lys Ser Arg Lys Glu Arg Glu Ala Glu Leu
290 295 300

Gly Ala Arg Ala Lys Glu Phe Thr Asn Val Tyr Ile Lys Asn Phe Gly
305 310 315 320

Glu Asp Met Asp Asp Glu Arg Leu Lys Asp Leu Phe Gly Lys Phe Gly
325 330 335

Pro Ala Leu Ser Val Lys Val Met Thr Asp Glu Ser Gly Lys Ser Lys
340 345 350

Gly Phe Gly Phe Val Ser Phe Glu Arg His Glu Asp Ala Gln Lys Ala
355 360 365

Val Asp Glu Met Asn Gly Lys Glu Leu Asn Gly Lys Gln Ile Tyr Val
370 375 380

Gly Arg Ala Gln Lys Lys Val Glu Arg Gln Thr Glu Leu Lys Arg Lys
385 390 395 400

Phe Glu Gln Met Lys Gln Asp Arg Ile Thr Arg Tyr Gln Gly Val Asn
405 410 415

Leu Tyr Val Lys Asn Leu Asp Asp Gly Ile Asp Asp Glu Arg Leu Arg
420 425 430

Lys Glu Phe Ser Pro Phe Gly Thr Ile Thr Ser Ala Lys Val Met Met
435 440 445

Glu Gly Gly Arg Ser Lys Gly Phe Gly Phe Val Cys Phe Ser Ser Pro
450 455 460

Glu Glu Ala Thr Lys Ala Val Thr Glu Met Asn Gly Arg Ile Val Ala

465		470		475		480									
Thr	Lys	Pro	Leu	Tyr	Val	Ala	Leu	Ala	Gln	Arg	Lys	Glu	Glu	Arg	Gln
			485						490					495	
Ala	His	Leu	Thr	Asn	Gln	Tyr	Met	Gln	Arg	Met	Ala	Ser	Val	Arg	Ala
		500						505					510		
Val	Pro	Asn	Pro	Val	Ile	Asn	Pro	Tyr	Gln	Pro	Ala	Pro	Pro	Ser	Gly
		515					520					525			
Tyr	Phe	Met	Ala	Ala	Ile	Pro	Gln	Thr	Gln	Asn	Arg	Ala	Ala	Tyr	Tyr
	530					535					540				
Pro	Pro	Ser	Gln	Ile	Ala	Gln	Leu	Arg	Pro	Ser	Pro	Arg	Trp	Thr	Ala
545					550					555					560
Gln	Gly	Ala	Arg	Pro	His	Pro	Phe	Gln	Asn	Met	Pro	Gly	Ala	Ile	Arg
				565					570					575	
Pro	Ala	Ala	Pro	Arg	Pro	Pro	Phe	Ser	Thr	Met	Arg	Pro	Ala	Ser	Ser
			580					585					590		
Gln	Val	Pro	Arg	Val	Met	Ser	Thr	Gln	Arg	Val	Ala	Asn	Thr	Ser	Thr
		595					600					605			
Gln	Thr	Met	Gly	Pro	Arg	Pro	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Thr	Pro
	610					615					620				
Ala	Val	Arg	Thr	Val	Pro	Gln	Tyr	Lys	Tyr	Ala	Ala	Gly	Val	Arg	Asn
625					630					635					640
Pro	Gln	Gln	His	Leu	Asn	Ala	Gln	Pro	Gln	Val	Thr	Met	Gln	Gln	Pro
				645					650					655	
Ala	Val	His	Val	Gln	Gly	Gln	Glu	Pro	Leu	Thr	Ala	Ser	Met	Leu	Ala
			660					665					670		
Ser	Ala	Pro	Pro	Gln	Glu	Gln	Lys	Gln	Met	Leu	Gly	Glu	Arg	Leu	Phe
		675					680					685			

Pro Leu Ile Gln Ala Met His Pro Thr Leu Ala Gly Lys Ile Thr Gly
690 695 700

Met Leu Leu Glu Ile Asp Asn Ser Glu Leu Leu His Met Leu Glu Ser
705 710 715 720

Pro Glu Ser Leu Arg Ser Lys Val Asp Glu Ala Val Ala Val Leu Gln
725 730 735

Ala His Gln Ala Lys Glu Ala Ala Gln Lys Ala Val Asn Ser Ala Thr
740 745 750

Gly Val Pro Thr Val
755

<210> 51
<211> 2271
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 51
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt 360
ggaagcaacc ccagcgcccc aagctacccc atggccagcc tgtacgtggg cgacctgcac 420
cccgacgtta ccgaggccat gctgtacgag aaattttctc ccgctggccc catcctgagc 480
atccgggtgt gccgggacat gattaccggg cgctctctgg ggtacgccta cgtgaacttc 540
cagcagcccc cagacgccga acggggccctg gacaccatga acttcgacgt gatcaagggc 600
aagcccgtgc ggatcatgtg gagccagcgg gacccaagcc tgagaaagag cggcgtgggc 660

aacatcttca tcaagaatct ggacaagagc atcgacaaca aggccctgta cgacaccttc	720
agcgccttcg gcaatatctt aagctgcaag gtgggtgtgcg acgagaacgg cagcaagggc	780
tacggcttcg tgcacttcga aacacaagaa gccgcagagc gggccatcga gaagatgaac	840
ggcatgctgc tgaacgaccg gaaggtgttc gtgggcccgt tcaagagccg gaaggagcgg	900
gaggctgaac tcggtgcccc ggccaaggag ttcaccaacg tgtacatcaa gaacttcggc	960
gaggacatgg acgacgagcg gttgaaggac ctgttcggca agttcggccc cgccctgtct	1020
gtgaagggtga tgaccgacga gagcggcaag tcgaagggtt tcggctttgt gagcttcgag	1080
cggcacgagg acgcacagaa ggccgtggac gaaatgaacg gtaaggagct taacggtgaag	1140
cagatctacg ttggccgggc ccagaagaag gtggagcggc agaccgagct gaagcgggaag	1200
ttcgagcaga tgaagcagga ccggatcacc cggtaccagg gcgtgaacct atacgtgaag	1260
aacctggacg acggcatcga cgacgaacgg ctgcggaagg aattcagccc cttcggcacc	1320
atcaccagcg ccaaggtaat gatggagggt ggccggtcga agggattcgg tttcgtgtgc	1380
ttcagcagcc ccgaggaggc caccaaggcc gttactgaga tgaacggccg gatcgtagcc	1440
acaaagccac tttacgtcgc cctggcccaa cgcaaggagg aacggcaggc ccacctgacc	1500
aaccagtaca tgcagcggat ggccagcgtg cgggcagtg ccaaccccgt gatcaacccc	1560
taccaacccg caccaccaag cggctacttc atggccgcca tccctcagac ccagaaccgg	1620
gccgcctact acccaccag ccagatcgcc cagttgcggc ctagtccccg ctggacggct	1680
caaggagcac gacccaccc cttccagaac atgcccggag ccatcagacc agctgcccct	1740
aggccaccct tcagcaccat gcggcccgcc tcaagccagg tgccccgggt gatgagcacc	1800
cagcgggtgg ccaacaccag caccagacc atggggccta gaccagcagc tgctgcggct	1860
gctgccactc ctgccgtccg gaccgtgcca cagtacaagt acgccgccgg cgttcggaac	1920
ccacagcagc acctgaacgc ccagccccag gtgaccatgc agcagcctgc cgtgcacgtg	1980
cagggccagg agcccctgac cgccagcatg ttagcaagcg cccctccaca ggagcagaag	2040
cagatgctgg gcgaaaggct gtttcccctg atccaagcga tgcattccac cctggccggc	2100
aagatcaccg gtatgctgct ggagatcgat aatagcgagc tgctgcacat gctggagagc	2160
cccgagagtc ttagatctaa ggtggacgag gccgtggccg tgctccaggc tcaccaggca	2220
aaggaggccg cccagaaaagc cgtgaactcc gccacaggcg tgcccaccgt g	2271

<210> 52
<211> 913
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 52
Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asp Pro Arg Val Ala Trp
115 120 125

Ile Gln Pro Glu Gln Lys Gly Pro Ala Asn Ala Leu Trp Met Gln Ile
130 135 140

Trp Glu Thr Ser Gln Gly Val Gly Arg Gly Gly Ser Gly Phe Ala Ser
145 150 155 160

Tyr Phe Cys Leu Asn Ser Pro Ala Leu Asp Thr Ala Ala Ala Ala Gly
165 170 175

Ala Ala Gly Arg Gly Ser Gly Gly Leu Gly Pro Ala Leu Pro Ala Ala
180 185 190

Ser Pro Pro Pro Pro Gly Pro Thr Ala Pro Ala Ala Leu Pro Pro Ala
195 200 205

Leu Leu Thr Ala Leu Gly Pro Ala Ala Glu Gly Ala Arg Arg Leu His
210 215 220

Lys Ser Pro Ser Leu Ser Ser Ser Ser Ser Ser Ser Ser Asn Ala
225 230 235 240

Glu Ser Gly Thr Glu Ser Pro Gly Cys Ser Ser Ser Ser Ser Ser Ser
245 250 255

Ala Ser Leu Gly Arg Pro Gly Gly Gly Arg Gly Gly Ala Phe Phe Asn
260 265 270

Phe Ala Asp Gly Ala Pro Ser Ala Pro Gly Thr Ala Asn Gly His Pro
275 280 285

Gly Pro Arg Gly Pro Ala Pro Ala Gly Ser Pro Ser Gln His Gln Phe
290 295 300

His Pro Gly Arg Arg Lys Arg Glu Asn Lys Ala Ser Thr Tyr Gly Leu
305 310 315 320

Asn Tyr Leu Leu Ser Gly Ser Arg Ala Ala Ala Leu Ser Gly Gly Gly
325 330 335

Gly Pro Gly Ala Gln Ala Pro Arg Pro Gly Thr Pro Trp Lys Ser Arg
340 345 350

Ala Tyr Ser Pro Gly Ile Gln Gly Leu His Glu Glu Ile Ile Asp Phe
355 360 365

Tyr Asn Phe Met Ser Pro Cys Pro Glu Glu Ala Ala Met Arg Arg Glu
370 375 380

Val Val Lys Arg Ile Glu Thr Val Val Lys Asp Leu Trp Pro Thr Ala
385 390 395 400

Asp Val Gln Ile Phe Gly Ser Phe Ser Thr Gly Leu Tyr Leu Pro Thr
405 410 415

Ser Asp Ile Asp Leu Val Val Phe Gly Lys Trp Glu Arg Pro Pro Leu
420 425 430

Gln Leu Leu Glu Gln Ala Leu Arg Lys His Asn Val Ala Glu Pro Cys
435 440 445

Ser Ile Lys Val Leu Asp Lys Ala Thr Val Pro Ile Ile Lys Leu Thr
450 455 460

Asp Gln Glu Thr Glu Val Lys Val Asp Ile Ser Phe Asn Met Glu Thr
465 470 475 480

Gly Val Arg Ala Ala Glu Phe Ile Lys Asn Tyr Met Lys Lys Tyr Ser
485 490 495

Leu Leu Pro Tyr Leu Ile Leu Val Leu Lys Gln Phe Leu Leu Gln Arg
500 505 510

Asp Leu Asn Glu Val Phe Thr Gly Gly Ile Ser Ser Tyr Ser Leu Ile
515 520 525

Leu Met Ala Ile Ser Phe Leu Gln Leu His Pro Arg Ile Asp Ala Arg
530 535 540

Arg Ala Asp Glu Asn Leu Gly Met Leu Leu Val Glu Phe Phe Glu Leu
545 550 555 560

Tyr Gly Arg Asn Phe Asn Tyr Leu Lys Thr Gly Ile Arg Ile Lys Glu
565 570 575

Gly Gly Ala Tyr Ile Ala Lys Glu Glu Ile Met Lys Ala Met Thr Ser
580 585 590

Gly Tyr Arg Pro Ser Met Leu Cys Ile Glu Asp Pro Leu Leu Pro Gly
595 600 605

Asn Asp Val Gly Arg Ser Ser Tyr Gly Ala Met Gln Val Lys Gln Val
610 615 620

Phe Asp Tyr Ala Tyr Ile Val Leu Ser His Ala Val Ser Pro Leu Ala
625 630 635 640

Arg Ser Tyr Pro Asn Arg Asp Ala Glu Ser Thr Leu Gly Arg Ile Ile
645 650 655

Lys Val Thr Gln Glu Val Ile Asp Tyr Arg Arg Trp Ile Lys Glu Lys
660 665 670

Trp Gly Ser Lys Ala His Pro Ser Pro Gly Met Asp Ser Arg Ile Lys
675 680 685

Ile Lys Glu Arg Ile Ala Thr Cys Asn Gly Glu Gln Thr Gln Asn Arg
690 695 700

Glu Pro Glu Ser Pro Tyr Gly Gln Arg Leu Thr Leu Ser Leu Ser Ser
705 710 715 720

Pro Gln Leu Leu Ser Ser Gly Ser Ser Ala Ser Ser Val Ser Ser Leu
725 730 735

Ser Gly Ser Asp Val Asp Ser Asp Thr Pro Pro Cys Thr Thr Pro Ser
740 745 750

Val Tyr Gln Phe Ser Leu Gln Ala Pro Ala Pro Leu Met Ala Gly Leu
755 760 765

Pro Thr Ala Leu Pro Met Pro Ser Gly Lys Pro Gln Pro Thr Thr Ser
770 775 780

Arg Thr Leu Ile Met Thr Thr Asn Asn Gln Thr Arg Phe Thr Ile Pro
785 790 795 800

Pro Pro Thr Leu Gly Val Ala Pro Val Pro Cys Arg Gln Ala Gly Val
805 810 815

Glu Gly Thr Ala Ser Leu Lys Ala Val His His Met Ser Ser Pro Ala
820 825 830

Ile Pro Ser Ala Ser Pro Asn Pro Leu Ser Ser Pro His Leu Tyr His
835 840 845

Lys Gln His Asn Gly Met Lys Leu Ser Met Lys Gly Ser His Gly His
850 855 860

Thr Gln Gly Gly Gly Tyr Ser Ser Val Gly Ser Gly Gly Val Arg Pro
865 870 875 880

Pro Val Gly Asn Arg Gly His His Gln Tyr Asn Arg Thr Gly Trp Arg
885 890 895

Arg Lys Lys His Thr His Thr Arg Asp Ser Leu Pro Val Ser Leu Ser
900 905 910

Arg

<210> 53

<211> 2373

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 53

gatccccggg tggcctggat ccagcccgag cagaagggcc ccgccaacgc actgtggatg	60
cagatctggg agacctctca gggcgtaggg aggggaggaa gcggttcgc cagctacttc	120
tgcctgaaca gcccgcctct ggatacggca gccgcagcag gagccgctgg gagaggctct	180
ggaggcctgg gaccagcact gcccgtgct tcaccaccac ctccaggccc taccgcaccc	240
gcagcccttc caccgcctt gctgaccgct ctgggcccag ctgcagaggg tgctcggcgg	300
ctgcacaaga gcccagcct gagcagcagc agtagcagct ccagcagcaa cgccgagagc	360
ggcacagaga gcccggctg cagcagcagc agctccagtt ccgcaagcct cggcagacca	420
ggcgggtggca gaggcggtgc attcttcaac ttcgctgacg gagcgccaag cgctccaggg	480

acagccaacg gacacccagg cccaagaggt ccggcacccg ctggaagccc cagccagcac	540
cagttccacc ccggcagacg gaagcgggag aacaaggcca gcacctacgg cctgaactac	600
ctgctgagcg gctctcgagc cgctgccctc tctggtggcg gtggccctgg tgctcaggcc	660
ccaagacctg gcaccccgtg gaagtcacgg gcctacagcc ccggtatcca gggcctgcac	720
gaggagatca tcgacttcta caacttcatg agcccctgtc ccgaggaagc cgccatgcgg	780
agggaggtgg tgaagcggat cgagaccgtg gtgaaggacc tgtggcccac cgcagacgtg	840
cagatcttcg gcagcttcag caccggcctg tacctgcccc ccagcgacat cgacctggtg	900
gtgttcggca agtgggaacg ccctcccctg cagctgctgg agcaggccct gcggaagcac	960
aacgtggccg agccctgcag catcaaggtg ctggacaagg ccaccgtgcc catcatcaag	1020
ctgaccgacc aggagaccga ggtgaaggtg gacatcagct tcaacatgga gaccggcgtg	1080
cgtgcagcag agttcatcaa gaactacatg aagaagtaca gcctgctgcc ctacctgac	1140
ctggtgctga agcagttcct gctgcagcgg gacctgaacg aggtgttcac cggcggcatc	1200
agcagctaca gcctgatcct catggccatc agcttcctgc agctgcaccc gcggatcgac	1260
gcaaggcggg ccgacgagaa cctgggcatg ctgctggtgg agttcttcga gctgtacggc	1320
cggaacttca actacctgaa gaccggcatc cggatcaagg agggcggcgc ctacatcgcc	1380
aaggaggaga tcatgaaggc catgaccagc ggctaccggc ccagcatgct gtgcatcgag	1440
gaccctctgc tgcccggcaa cgacgtgggc cggagcagct acggcgccat gcaggtgaag	1500
caggtgttcg actacgccta catcgtgctg agccacgccg tgagcccact ggcccggagc	1560
taccccaacc gggacgccga gagcaccctg ggccggatca tcaaggtgac ccaggaggtg	1620
atcgactacc ggcggtggat caaggagaag tggggcagca agggccaccc atctcccggc	1680
atggacagcc ggatcaagat caaggagcgg atcgccacct gcaacggcga gcagaccag	1740
aaccgggagc ccgaaagccc ctacggccag cgtctgacct tgagcctgag ctctccccag	1800
ctgctgagca gcggcagcag cgccagcagc gtgagcagcc tgagcggcag cgacgtggac	1860
agcgacaccc caccctgcac caccccaagc gtgtaccagt tctcactgca ggcccctgct	1920
ccgctgatgg ctggcctgcc taccgccctg cccatgcccc gcggcaagcc ccagcccacc	1980
accagccgga ccctgatcat gaccaccaac aaccagaccc ggttcaccat cccaccacca	2040
accctgggcg tggctcctgt gccttgtcgg caggccggag tggagggcac cgccagcctg	2100

aaggccgtgc accacatgag cagcccagcc atccccagcg ccagcccca cctctgagc 2160
agcccacacc tgtaccacaa gcagcacaac ggcatagaac tgagcatgaa gggcagccac 2220
ggacacaccc agggcggagg gtactcaagc gttgggagcg gaggggtacg gcctccagtg 2280
ggcaaccggg gccaccacca gtacaacagg accggctgga gacggaagaa gcacaccac 2340
acccgggact ctctgcccggt gaggctgagc cgg 2373

<210> 54
<211> 791
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 54
Asp Pro Arg Val Ala Trp Ile Gln Pro Glu Gln Lys Gly Pro Ala Asn
1 5 10 15

Ala Leu Trp Met Gln Ile Trp Glu Thr Ser Gln Gly Val Gly Arg Gly
20 25 30

Gly Ser Gly Phe Ala Ser Tyr Phe Cys Leu Asn Ser Pro Ala Leu Asp
35 40 45

Thr Ala Ala Ala Ala Gly Ala Ala Gly Arg Gly Ser Gly Gly Leu Gly
50 55 60

Pro Ala Leu Pro Ala Ala Ser Pro Pro Pro Gly Pro Thr Ala Pro
65 70 75 80

Ala Ala Leu Pro Pro Ala Leu Leu Thr Ala Leu Gly Pro Ala Ala Glu
85 90 95

Gly Ala Arg Arg Leu His Lys Ser Pro Ser Leu Ser Ser Ser Ser
100 105 110

Ser Ser Ser Ser Asn Ala Glu Ser Gly Thr Glu Ser Pro Gly Cys Ser
115 120 125

Ser Ser Ser Ser Ser Ser Ala Ser Leu Gly Arg Pro Gly Gly Gly Arg
130 135 140

Gly Gly Ala Phe Phe Asn Phe Ala Asp Gly Ala Pro Ser Ala Pro Gly
145 150 155 160

Thr Ala Asn Gly His Pro Gly Pro Arg Gly Pro Ala Pro Ala Gly Ser
165 170 175

Pro Ser Gln His Gln Phe His Pro Gly Arg Arg Lys Arg Glu Asn Lys
180 185 190

Ala Ser Thr Tyr Gly Leu Asn Tyr Leu Leu Ser Gly Ser Arg Ala Ala
195 200 205

Ala Leu Ser Gly Gly Gly Gly Pro Gly Ala Gln Ala Pro Arg Pro Gly
210 215 220

Thr Pro Trp Lys Ser Arg Ala Tyr Ser Pro Gly Ile Gln Gly Leu His
225 230 235 240

Glu Glu Ile Ile Asp Phe Tyr Asn Phe Met Ser Pro Cys Pro Glu Glu
245 250 255

Ala Ala Met Arg Arg Glu Val Val Lys Arg Ile Glu Thr Val Val Lys
260 265 270

Asp Leu Trp Pro Thr Ala Asp Val Gln Ile Phe Gly Ser Phe Ser Thr
275 280 285

Gly Leu Tyr Leu Pro Thr Ser Asp Ile Asp Leu Val Val Phe Gly Lys
290 295 300

Trp Glu Arg Pro Pro Leu Gln Leu Leu Glu Gln Ala Leu Arg Lys His
305 310 315 320

Asn Val Ala Glu Pro Cys Ser Ile Lys Val Leu Asp Lys Ala Thr Val
325 330 335

Pro Ile Ile Lys Leu Thr Asp Gln Glu Thr Glu Val Lys Val Asp Ile

340

345

350

Ser Phe Asn Met Glu Thr Gly Val Arg Ala Ala Glu Phe Ile Lys Asn
 355 360 365

Tyr Met Lys Lys Tyr Ser Leu Leu Pro Tyr Leu Ile Leu Val Leu Lys
 370 375 380

Gln Phe Leu Leu Gln Arg Asp Leu Asn Glu Val Phe Thr Gly Gly Ile
 385 390 395 400

Ser Ser Tyr Ser Leu Ile Leu Met Ala Ile Ser Phe Leu Gln Leu His
 405 410 415

Pro Arg Ile Asp Ala Arg Arg Ala Asp Glu Asn Leu Gly Met Leu Leu
 420 425 430

Val Glu Phe Phe Glu Leu Tyr Gly Arg Asn Phe Asn Tyr Leu Lys Thr
 435 440 445

Gly Ile Arg Ile Lys Glu Gly Gly Ala Tyr Ile Ala Lys Glu Glu Ile
 450 455 460

Met Lys Ala Met Thr Ser Gly Tyr Arg Pro Ser Met Leu Cys Ile Glu
 465 470 475 480

Asp Pro Leu Leu Pro Gly Asn Asp Val Gly Arg Ser Ser Tyr Gly Ala
 485 490 495

Met Gln Val Lys Gln Val Phe Asp Tyr Ala Tyr Ile Val Leu Ser His
 500 505 510

Ala Val Ser Pro Leu Ala Arg Ser Tyr Pro Asn Arg Asp Ala Glu Ser
 515 520 525

Thr Leu Gly Arg Ile Ile Lys Val Thr Gln Glu Val Ile Asp Tyr Arg
 530 535 540

Arg Trp Ile Lys Glu Lys Trp Gly Ser Lys Ala His Pro Ser Pro Gly
 545 550 555 560

Met Asp Ser Arg Ile Lys Ile Lys Glu Arg Ile Ala Thr Cys Asn Gly
565 570 575

Glu Gln Thr Gln Asn Arg Glu Pro Glu Ser Pro Tyr Gly Gln Arg Leu
580 585 590

Thr Leu Ser Leu Ser Ser Pro Gln Leu Leu Ser Ser Gly Ser Ser Ala
595 600 605

Ser Ser Val Ser Ser Leu Ser Gly Ser Asp Val Asp Ser Asp Thr Pro
610 615 620

Pro Cys Thr Thr Pro Ser Val Tyr Gln Phe Ser Leu Gln Ala Pro Ala
625 630 635 640

Pro Leu Met Ala Gly Leu Pro Thr Ala Leu Pro Met Pro Ser Gly Lys
645 650 655

Pro Gln Pro Thr Thr Ser Arg Thr Leu Ile Met Thr Thr Asn Asn Gln
660 665 670

Thr Arg Phe Thr Ile Pro Pro Pro Thr Leu Gly Val Ala Pro Val Pro
675 680 685

Cys Arg Gln Ala Gly Val Glu Gly Thr Ala Ser Leu Lys Ala Val His
690 695 700

His Met Ser Ser Pro Ala Ile Pro Ser Ala Ser Pro Asn Pro Leu Ser
705 710 715 720

Ser Pro His Leu Tyr His Lys Gln His Asn Gly Met Lys Leu Ser Met
725 730 735

Lys Gly Ser His Gly His Thr Gln Gly Gly Gly Tyr Ser Ser Val Gly
740 745 750

Ser Gly Gly Val Arg Pro Pro Val Gly Asn Arg Gly His His Gln Tyr
755 760 765

Asn Arg Thr Gly Trp Arg Arg Lys Lys His Thr His Thr Arg Asp Ser

770

775

780

Leu Pro Val Ser Leu Ser Arg
785 790

<210> 55

<211> 2739

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 55

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcgatc cccgggtggc ctggatccag cccgagcaga agggccccgc caacgcactg	420
tggatgcaga tctgggagac ctctcagggc gtagggaggg gaggaagcgg cttcgccagc	480
tacttctgcc tgaacagccc cgctctggat acggcagccg cagcaggagc cgctgggaga	540
ggctctggag gcctgggacc agcactgccc gctgcttcac caccacctcc aggccctacc	600
gcacccgcag cccttcacc cgccctgctg accgctctgg gccagctgc agagggtgct	660
cggcggctgc acaagagccc cagcctgagc agcagcagta gcagctccag cagcaacgcc	720
gagagcggca cagagagccc cggctgcagc agcagcagct ccagttccgc aagcctcggc	780
agaccaggcg gtggcagagg cggtgcattc ttcaacttcg ctgacggagc gccaaagcgt	840
ccagggacag ccaacggaca cccaggccca agaggtccgg caccgctgg aagccccagc	900
cagcaccagt tccaccccgg cagacggaag cgggagaaca aggccagcac ctacggcctg	960
aactacctgc tgagcggctc tcgagccgct gccctctctg gtggcgggtg ccctgggtgct	1020
caggcccca gacctggcac cccgtggaag tcacgggcct acagccccgg tatccagggc	1080

ctgcacgagg agatcatcga cttctacaac ttcattagacc cctgtcccga ggaagccgcc	1140
atgcggaggg aggtggtgaa gcggatcgag accgtggtga aggacctgtg gcccaccgca	1200
gacgtgcaga tcttcggcag cttcagcacc ggcctgtacc tgcccaccag cgacatcgac	1260
ctggtggtgt tcggcaagtg ggaacgccct cccctgcagc tgctggagca ggccctgcgg	1320
aagcacaacg tggccgagcc ctgcagcatc aaggtgctgg acaaggccac cgtgcccatac	1380
atcaagctga ccgaccagga gaccgaggtg aaggtggaca tcagcttcaa catggagacc	1440
ggcgtgcgtg cagcagagtt catcaagaac tacatgaaga agtacagcct gctgccttac	1500
ctgacacctg tgctgaagca gttcctgctg cagcgggacc tgaacgaggt gttcaccggc	1560
ggcatcagca gctacagcct gatcctcatg gccatcagct tcctgcagct gcacccgcgg	1620
atcgacgcaa ggcgggccga cgagaacctg ggcatgctgc tgggtggagtt cttcgagctg	1680
tacggccgga acttcaacta cctgaagacc ggcatccgga tcaaggaggg cggcgcctac	1740
atcgccaagg aggagatcat gaaggccatg accagcggct accggcccag catgctgtgc	1800
atcgaggacc ctctgctgcc cggcaacgac gtgggcccga gcagctacgg cgccatgcag	1860
gtgaagcagg tgttcgacta cgcctacatc gtgctgagcc acgccgtgag cccactggcc	1920
cggagctacc ccaaccggga cgccgagagc accctgggcc ggatcatcaa ggtgaccag	1980
gaggtgatcg actaccggcg gtggatcaag gagaagtggg gcagcaaggc ccacccatct	2040
cccggcatgg acagccggat caagatcaag gagcggatcg ccacctgcaa cggcgagcag	2100
accagaacc gggagcccga aagcccctac ggccagcgtc tgaccctgag cctgagctct	2160
ccccagctgc tgagcagcgg cagcagcgcc agcagcgtga gcagcctgag cggcagcgac	2220
gtggacagcg acaccccacc ctgcaccacc ccaagcgtgt accagttctc actgcaggcc	2280
cctgctccgc tgatggctgg cctgcctacc gccctgccc tgcccagcgg caagccccag	2340
cccaccacca gccggaccct gatcatgacc accaacaacc agaccgggtt caccatccca	2400
ccaccaaccc tgggcgtggc tcctgtgcct tgtcggcagg ccggagtgga gggcaccgcc	2460
agcctgaagg ccgtgcacca catgagcagc ccagccatcc ccagcgccag ccccaaccct	2520
ctgagcagcc cacacctgta ccacaagcag cacaacggca tgaagctgag catgaagggc	2580
agccacggac acaccaggg cgagggttac tcaagcgttg ggagcggagg ggtacggcct	2640
ccagtgggca accggggcca ccaccagtac aacaggaccg gctggagacg gaagaagcac	2700

acccacaccc gggactctct gcccgtagc ctgagccgg

2739

<210> 56

<211> 665

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 56

Asp Pro Arg Ile Ala Trp Phe Gln Pro Glu Gln Leu Gly Pro Ser Asn
1 5 10 15

Ser Leu Trp Met Gln Ile Trp Glu Thr Thr Gln Gly Leu Arg Asn Leu
20 25 30

Tyr Phe Asn His His Cys His Ser Ser Gly Gly Ala Ser Gly Gly Gly
35 40 45

Gly Ser Ser Ser Ser Ser Ser Thr Ala Thr Gly Gly Ser Gly Ser Ser
50 55 60

Thr Gly Ser Pro Gly Gly Ala Ala Ser Ala Pro Ala Pro Ala Pro Ala
65 70 75 80

Gly Met Tyr Arg Ser Gly Glu Arg Leu Leu Gly Ser His Ala Leu Pro
85 90 95

Ala Glu Gln Arg Asp Phe Leu Pro Leu Glu Thr Thr Asn Asn Asn Asn
100 105 110

Asn His His Gln Pro Gly Ala Trp Ala Arg Arg Ala Gly Ser Ser Ala
115 120 125

Ser Ser Pro Pro Ser Ala Ser Ser Ser Pro His Pro Ser Ala Ala Val
130 135 140

Pro Ala Ala Asp Pro Ala Asp Ser Ala Ser Gly Ser Ser Asn Lys Arg
145 150 155 160

Lys Arg Asp Asn Lys Ala Ser Thr Tyr Gly Leu Asn Tyr Ser Leu Leu
165 170 175

Gln Pro Ser Gly Gly Arg Ala Ala Gly Gly Gly Arg Ala Asp Gly Gly
180 185 190

Gly Val Val Tyr Ser Gly Thr Pro Trp Lys Arg Arg Asn Tyr Asn Gln
195 200 205

Gly Val Val Gly Leu His Glu Glu Ile Ser Asp Phe Tyr Glu Tyr Met
210 215 220

Ser Pro Arg Pro Glu Glu Glu Lys Met Arg Met Glu Val Val Asn Arg
225 230 235 240

Ile Glu Ser Val Ile Lys Glu Leu Trp Pro Ser Ala Asp Val Gln Ile
245 250 255

Phe Gly Ser Phe Lys Thr Gly Leu Tyr Leu Pro Thr Ser Asp Ile Asp
260 265 270

Leu Val Val Phe Gly Lys Trp Glu Asn Leu Pro Leu Trp Thr Leu Glu
275 280 285

Glu Ala Leu Arg Lys His Lys Val Ala Asp Glu Asp Ser Val Lys Val
290 295 300

Leu Asp Lys Ala Thr Val Pro Ile Ile Lys Leu Thr Asp Ser Phe Thr
305 310 315 320

Glu Val Lys Val Asp Ile Ser Phe Asn Val Gln Asn Gly Val Arg Ala
325 330 335

Ala Asp Leu Ile Lys Asp Phe Thr Lys Lys Tyr Pro Val Leu Pro Tyr
340 345 350

Leu Val Leu Val Leu Lys Gln Phe Leu Leu Gln Arg Asp Leu Asn Glu
355 360 365

Val Phe Thr Gly Gly Ile Gly Ser Tyr Ser Leu Phe Leu Met Ala Val
370 375 380

Ser Phe Leu Gln Leu His Pro Arg Glu Asp Ala Cys Ile Pro Asn Thr
385 390 395 400

Asn Tyr Gly Val Leu Leu Ile Glu Phe Phe Glu Leu Tyr Gly Arg His
405 410 415

Phe Asn Tyr Leu Lys Thr Gly Ile Arg Ile Lys Asp Gly Gly Ser Tyr
420 425 430

Val Ala Lys Asp Glu Val Gln Lys Asn Met Leu Asp Gly Tyr Arg Pro
435 440 445

Ser Met Leu Tyr Ile Glu Asp Pro Leu Gln Pro Gly Asn Asp Val Gly
450 455 460

Arg Ser Ser Tyr Gly Ala Met Gln Val Lys Gln Ala Phe Asp Tyr Ala
465 470 475 480

Tyr Val Val Leu Ser His Ala Val Ser Pro Ile Ala Lys Tyr Tyr Pro
485 490 495

Asn Asn Glu Thr Glu Ser Ile Leu Gly Arg Ile Ile Arg Val Thr Asp
500 505 510

Glu Val Ala Thr Tyr Arg Asp Trp Ile Ser Lys Gln Trp Gly Leu Lys
515 520 525

Asn Arg Pro Glu Pro Ser Cys Asn Gly Pro Val Ser Ser Ser Ser Ala
530 535 540

Thr Gln Ser Ser Ser Ser Asp Val Asp Ser Asp Ala Thr Pro Cys Lys
545 550 555 560

Thr Pro Lys Gln Leu Leu Cys Arg Pro Ser Thr Gly Asn Arg Val Gly
565 570 575

Ser Gln Asp Val Ser Leu Glu Ser Ser Gln Ala Val Gly Lys Met Gln
580 585 590

Ser Thr Gln Thr Thr Asn Thr Ser Asn Ser Thr Asn Lys Ser Gln His
595 600 605

Gly Ser Ala Arg Leu Phe Arg Ser Ser Ser Lys Gly Phe Gln Gly Thr
610 615 620

Thr Gln Thr Ser His Gly Ser Leu Met Thr Asn Lys Gln His Gln Gly
625 630 635 640

Lys Ser Asn Asn Gln Tyr Tyr His Gly Lys Lys Arg Lys His Lys Arg
645 650 655

Asp Ala Pro Leu Ser Asp Leu Cys Arg
660 665

<210> 57
<211> 1995
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 57
gatccccgga tcgcctgggt ccagcccgag cagctaggcc caagcaacag cctgtggatg 60
cagatctggg agaccacca gggcctgcgg aacctgtact tcaaccacca ctgtcacagc 120
tctggcggcg caagtggcgg aggcggcagt agcagcagct cttcaacagc aaccggaggc 180
agcggctcgt ccaccggcag tccgggtgga gccgccagtg cccagctcc tgccccagca 240
ggcatgtacc ggagcggcga acggctgctg ggcagccacg ctctgcctgc cgagcagcgg 300
gacttcctgc ccctggagac caccaataac aacaacaatc accaccaacc cggggcttgg 360
gctcgacgag ccggctccag tgcaagcagt ccacctagcg ccagcagctc tccacacccc 420
agcgccgctg ttccagcagc cgatcccgcc gacagtgcc a gcggcagcag caacaagcgg 480
aagcgggaca acaaggccag cacctacggc ctgaactaca gcctgctgca accttctgga 540
ggccgtgcag ctggcggagg tagagccgac ggtggcgggtg tgggtgtacag cggcacgccc 600
tggaagcggc ggaactacaa ccagggcgtg gtgggcctgc acgaggagat cagcgacttc 660
tacgagtaca tgagcccacg gcccgaagag gagaagatgc ggatggaggt ggtgaaccgg 720

atcgagagcg tgatcaagga gctgtggcct agcgccgacg tgcagatctt cggcagcttc	780
aagaccggcc tgtacctgcc caccagcgac atcgacctgg tgggtgttcgg caagtgggag	840
aacctgcccc tgtggaccct ggaggaggcc ctgcggaaac acaaggtggc cgacgaggac	900
agcgtgaagg tgctggacaa ggccaccgtg cccatcatca agctgaccga cagcttcacc	960
gaggtgaagg tggacatcag cttcaacgtg cagaacggcg tgagggcagc cgacctgac	1020
aaggacttca ccaagaagta ccccgtgctg ccctacctgg tgctggtgct gaagcagttc	1080
ctgctgcagc gggacctgaa cgaggtgttc accggcggca tcggcagcta cagcctgttc	1140
ctgatggccg tgagcttcct gcagctgcat ccccgtgagg acgcctgcat ccccaacacc	1200
aactacggcg tgctgctgat cgagttcttc gagctgtacg gccggcactt caactacctg	1260
aagaccggca tccggatcaa ggacggcggc agctacgtgg ccaaggacga ggtgcagaag	1320
aacatgctgg acggctaccg gcccagcatg ctgtacatcg aggaccact gcagcccggc	1380
aacgacgtgg gccggagcag ctacggcgcc atgcaggtga agcaggcctt cgactacgcc	1440
tacgtggtgc tgagccacgc cgtgagcccc atcgccaagt actaccccaa caacgagacc	1500
gagagcatcc tgggccggat catccgggtg accgacgagg tggccaccta ccgggactgg	1560
atcagcaagc agtggggcct gaagaaccgg cccgagccta gctgcaacgg ccccgtgagc	1620
tcaagcagcg ccacccagag cagcagcagc gacgtggaca gcgacgccac tccctgcaag	1680
accctaagc agctgctgtg ccgtcccagc accggcaacc gggtgggcag ccaggacgtg	1740
agcctggaga gcagccaggc cgtgggcaag atgcagagca cccagaccac caacaccagc	1800
aacagcacca acaagagcca acacggcagc gcacggctgt tccggagcag cagcaagggc	1860
ttccagggca ccacccagac cagccacggc agcctgatga ccaacaagca gcaccagggc	1920
aagagcaaca accagtacta ccacggcaag aagcggaagc acaagcggga cgctcccctg	1980
agcgacctgt gccgg	1995

<210> 58

<211> 787

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 58

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asp Pro Arg Ile Ala Trp
115 120 125

Phe Gln Pro Glu Gln Leu Gly Pro Ser Asn Ser Leu Trp Met Gln Ile
130 135 140

Trp Glu Thr Thr Gln Gly Leu Arg Asn Leu Tyr Phe Asn His His Cys
145 150 155 160

His Ser Ser Gly Gly Ala Ser Gly Gly Gly Gly Ser Ser Ser Ser Ser
165 170 175

Ser Thr Ala Thr Gly Gly Ser Gly Ser Ser Thr Gly Ser Pro Gly Gly
180 185 190

Ala Ala Ser Ala Pro Ala Pro Ala Pro Ala Gly Met Tyr Arg Ser Gly
195 200 205

Glu Arg Leu Leu Gly Ser His Ala Leu Pro Ala Glu Gln Arg Asp Phe
210 215 220

Leu Pro Leu Glu Thr Thr Asn Asn Asn Asn Asn His His Gln Pro Gly
225 230 235 240

Ala Trp Ala Arg Arg Ala Gly Ser Ser Ala Ser Ser Pro Pro Ser Ala
245 250 255

Ser Ser Ser Pro His Pro Ser Ala Ala Val Pro Ala Ala Asp Pro Ala
260 265 270

Asp Ser Ala Ser Gly Ser Ser Asn Lys Arg Lys Arg Asp Asn Lys Ala
275 280 285

Ser Thr Tyr Gly Leu Asn Tyr Ser Leu Leu Gln Pro Ser Gly Gly Arg
290 295 300

Ala Ala Gly Gly Gly Arg Ala Asp Gly Gly Gly Val Val Tyr Ser Gly
305 310 315 320

Thr Pro Trp Lys Arg Arg Asn Tyr Asn Gln Gly Val Val Gly Leu His
325 330 335

Glu Glu Ile Ser Asp Phe Tyr Glu Tyr Met Ser Pro Arg Pro Glu Glu
340 345 350

Glu Lys Met Arg Met Glu Val Val Asn Arg Ile Glu Ser Val Ile Lys
355 360 365

Glu Leu Trp Pro Ser Ala Asp Val Gln Ile Phe Gly Ser Phe Lys Thr
370 375 380

Gly Leu Tyr Leu Pro Thr Ser Asp Ile Asp Leu Val Val Phe Gly Lys
385 390 395 400

Trp Glu Asn Leu Pro Leu Trp Thr Leu Glu Glu Ala Leu Arg Lys His
405 410 415

Lys Val Ala Asp Glu Asp Ser Val Lys Val Leu Asp Lys Ala Thr Val

420

425

430

Pro Ile Ile Lys Leu Thr Asp Ser Phe Thr Glu Val Lys Val Asp Ile
 435 440 445

Ser Phe Asn Val Gln Asn Gly Val Arg Ala Ala Asp Leu Ile Lys Asp
 450 455 460

Phe Thr Lys Lys Tyr Pro Val Leu Pro Tyr Leu Val Leu Val Leu Lys
 465 470 475 480

Gln Phe Leu Leu Gln Arg Asp Leu Asn Glu Val Phe Thr Gly Gly Ile
 485 490 495

Gly Ser Tyr Ser Leu Phe Leu Met Ala Val Ser Phe Leu Gln Leu His
 500 505 510

Pro Arg Glu Asp Ala Cys Ile Pro Asn Thr Asn Tyr Gly Val Leu Leu
 515 520 525

Ile Glu Phe Phe Glu Leu Tyr Gly Arg His Phe Asn Tyr Leu Lys Thr
 530 535 540

Gly Ile Arg Ile Lys Asp Gly Gly Ser Tyr Val Ala Lys Asp Glu Val
 545 550 555 560

Gln Lys Asn Met Leu Asp Gly Tyr Arg Pro Ser Met Leu Tyr Ile Glu
 565 570 575

Asp Pro Leu Gln Pro Gly Asn Asp Val Gly Arg Ser Ser Tyr Gly Ala
 580 585 590

Met Gln Val Lys Gln Ala Phe Asp Tyr Ala Tyr Val Val Leu Ser His
 595 600 605

Ala Val Ser Pro Ile Ala Lys Tyr Tyr Pro Asn Asn Glu Thr Glu Ser
 610 615 620

Ile Leu Gly Arg Ile Ile Arg Val Thr Asp Glu Val Ala Thr Tyr Arg
 625 630 635 640

Asp Trp Ile Ser Lys Gln Trp Gly Leu Lys Asn Arg Pro Glu Pro Ser
645 650 655

Cys Asn Gly Pro Val Ser Ser Ser Ser Ala Thr Gln Ser Ser Ser Ser
660 665 670

Asp Val Asp Ser Asp Ala Thr Pro Cys Lys Thr Pro Lys Gln Leu Leu
675 680 685

Cys Arg Pro Ser Thr Gly Asn Arg Val Gly Ser Gln Asp Val Ser Leu
690 695 700

Glu Ser Ser Gln Ala Val Gly Lys Met Gln Ser Thr Gln Thr Thr Asn
705 710 715 720

Thr Ser Asn Ser Thr Asn Lys Ser Gln His Gly Ser Ala Arg Leu Phe
725 730 735

Arg Ser Ser Ser Lys Gly Phe Gln Gly Thr Thr Gln Thr Ser His Gly
740 745 750

Ser Leu Met Thr Asn Lys Gln His Gln Gly Lys Ser Asn Asn Gln Tyr
755 760 765

Tyr His Gly Lys Lys Arg Lys His Lys Arg Asp Ala Pro Leu Ser Asp
770 775 780

Leu Cys Arg
785

<210> 59
<211> 2361
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 59
atggccagca acttcacca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc

gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcgggt	360
ggaagcgatc cccggatcgc ctggttccag cccgagcagc taggcccaag caacagcctg	420
tggatgcaga tctgggagac caccagggc ctgcggaacc tgtacttcaa ccaccactgt	480
cacagctctg gcggcgcaag tggcggaggc ggcagtagca gcagctcttc aacagcaacc	540
ggaggcagcg gctcgtccac cggcagtccg ggtggagccg ccagtgcccc agctcctgcc	600
ccagcaggca tgtaccggag cggcgaacgg ctgctgggca gccacgctct gcctgccgag	660
cagcgggact tcctgcccct ggagaccacc aataacaaca acaatcacca ccaaccggg	720
gcttgggctc gacgagccgg ctccagtgc agcagtccac ctagcgccag cagctctcca	780
caccccagcg ccgctgttcc agcagccgat cccgccgaca gtgccagcgg cagcagcaac	840
aagcggaagc gggacaacaa ggccagcacc tacggcctga actacagcct gctgcaacct	900
tctggaggcc gtgcagctgg cggaggtaga gccgacgggt gcggtgtggt gtacagcggc	960
acgccctgga agcggcggaa ctacaaccag ggcgtggtgg gcctgcacga ggagatcagc	1020
gacttctacg agtacatgag cccacggccc gaagaggaga agatgcggat ggaggtggtg	1080
aaccggatcg agagcgtgat caaggagctg tggcctagcg ccgacgtgca gatcttcggc	1140
agcttcaaga ccggcctgta cctgcccacc agcgacatcg acctggtggt gttcggcaag	1200
tgggagaacc tgcccctgtg gaccctggag gaggccctgc ggaaacacaa ggtggccgac	1260
gaggacagcg tgaaggtgct ggacaaggcc accgtgcccc tcatcaagct gaccgacagc	1320
ttcaccgagg tgaaggtgga catcagcttc aacgtgcaga acggcgtgag ggcagccgac	1380
ctgatcaagg acttcaccaa gaagtacccc gtgctgccct acctggtgct ggtgctgaag	1440
cagttcctgc tgcagcggga cctgaacgag gtgttcaccg gcggcatcgg cagctacagc	1500
ctgttcctga tggccgtgag cttcctgcag ctgcatcccc gtgaggacgc ctgcatcccc	1560
aacaccaact acggcgtgct gctgatcgag ttcttcgagc tgtacggccg gcacttcaac	1620
tacctgaaga ccggcatccg gatcaaggac ggcggcagct acgtggccaa ggacgaggtg	1680

cagaagaaca tgctggacgg ctaccggccc agcatgctgt acatcgagga cccactgcag 1740
 cccggcaacg acgtgggccc gagcagctac ggcgccatgc aggtgaagca ggccttcgac 1800
 tacgcctacg tgggtgctgag ccacgccgtg agcccatcg ccaagtacta cccaacaac 1860
 gagaccgaga gcatcctggg ccggatcatc cgggtgaccg acgaggtggc cacctaccgg 1920
 gactggatca gcaagcagtg gggcctgaag aaccggcccc agcctagctg caacggcccc 1980
 gtgagctcaa gcagcgccac ccagagcagc agcagcgacg tggacagcga cgccactccc 2040
 tgcaagaccc ctaagcagct gctgtgccgt cccagcaccg gcaaccgggt gggcagccag 2100
 gacgtgagcc tggagagcag ccaggccgtg ggcaagatgc agagcaccca gaccaccaac 2160
 accagcaaca gcaccaaaa gagccaacac ggcagcgcac ggctgttccg gagcagcagc 2220
 aagggttcc agggcaccac ccagaccagc cacggcagcc tgatgaccaa caagcagcac 2280
 cagggaaga gcaacaacca gtactaccac ggcaagaagc ggaagcaca gcgggacgct 2340
 cccctgagcg acctgtgccg g 2361

<210> 60

<211> 1584

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 60

Asn Ser Gln Pro Gln Thr Arg Ser Pro Phe Phe Gln Arg Pro Gln Ile
 1 5 10 15

Gln Pro Pro Arg Ala Thr Ile Pro Asn Ser Ser Pro Ser Ile Arg Pro
 20 25 30

Gly Ala Gln Thr Pro Thr Ala Val Tyr Gln Ala Asn Gln His Ile Met
 35 40 45

Met Val Asn His Leu Pro Met Pro Tyr Pro Val Pro Gln Gly Pro Gln
 50 55 60

Tyr Cys Ile Pro Gln Tyr Arg His Ser Gly Pro Pro Tyr Val Gly Pro
 65 70 75 80

Pro Gln Gln Tyr Pro Val Gln Pro Pro Gly Pro Gly Pro Phe Tyr Pro
85 90 95

Gly Pro Gly Pro Gly Asp Phe Pro Asn Ala Tyr Gly Thr Pro Phe Tyr
100 105 110

Pro Ser Gln Pro Val Tyr Gln Ser Ala Pro Ile Ile Val Pro Thr Gln
115 120 125

Gln Gln Pro Pro Pro Ala Lys Arg Glu Lys Lys Thr Ile Arg Ile Arg
130 135 140

Asp Pro Asn Gln Gly Gly Lys Asp Ile Thr Glu Glu Ile Met Ser Gly
145 150 155 160

Gly Gly Ser Arg Asn Pro Thr Pro Pro Ile Gly Arg Pro Thr Ser Thr
165 170 175

Pro Thr Pro Pro Gln Gln Leu Pro Ser Gln Val Pro Glu His Ser Pro
180 185 190

Val Val Tyr Gly Thr Val Glu Ser Ala His Leu Ala Ala Ser Thr Pro
195 200 205

Val Thr Ala Ala Ser Asp Gln Lys Gln Glu Glu Lys Pro Lys Pro Asp
210 215 220

Pro Val Leu Lys Ser Pro Ser Pro Val Leu Arg Leu Val Leu Ser Gly
225 230 235 240

Glu Lys Lys Glu Gln Glu Gly Gln Thr Ser Glu Thr Thr Ala Ile Val
245 250 255

Ser Ile Ala Glu Leu Pro Leu Pro Pro Ser Pro Thr Thr Val Ser Ser
260 265 270

Val Ala Arg Ser Thr Ile Ala Ala Pro Thr Ser Ser Ala Leu Ser Ser
275 280 285

Gln Pro Ile Phe Thr Thr Ala Ile Asp Asp Arg Cys Glu Leu Ser Ser
290 295 300

Pro Arg Glu Asp Thr Ile Pro Ile Pro Ser Leu Thr Ser Cys Thr Glu
305 310 315 320

Thr Ser Asp Pro Leu Pro Thr Asn Glu Asn Asp Asp Asp Ile Cys Lys
325 330 335

Lys Pro Cys Ser Val Ala Pro Asn Asp Ile Pro Leu Val Ser Ser Thr
340 345 350

Asn Leu Ile Asn Glu Ile Asn Gly Val Ser Glu Lys Leu Ser Ala Thr
355 360 365

Glu Ser Ile Val Glu Ile Val Lys Gln Glu Val Leu Pro Leu Thr Leu
370 375 380

Glu Leu Glu Ile Leu Glu Asn Pro Pro Glu Glu Met Lys Leu Glu Cys
385 390 395 400

Ile Pro Ala Pro Ile Thr Pro Ser Thr Val Pro Ser Phe Pro Pro Thr
405 410 415

Pro Pro Thr Pro Pro Ala Ser Pro Pro His Thr Pro Val Ile Val Pro
420 425 430

Ala Ala Ala Thr Thr Val Ser Ser Pro Ser Ala Ala Ile Thr Val Gln
435 440 445

Arg Val Leu Glu Glu Asp Glu Ser Ile Arg Thr Cys Leu Ser Glu Asp
450 455 460

Ala Lys Glu Ile Gln Asn Lys Ile Glu Val Glu Ala Asp Gly Gln Thr
465 470 475 480

Glu Glu Ile Leu Asp Ser Gln Asn Leu Asn Ser Arg Arg Ser Pro Val
485 490 495

Pro Ala Gln Ile Ala Ile Thr Val Pro Lys Thr Trp Lys Lys Pro Lys
500 505 510

Asp Arg Thr Arg Thr Thr Glu Glu Met Leu Glu Ala Glu Leu Glu Leu
515 520 525

Lys Ala Glu Glu Glu Leu Ser Ile Asp Lys Val Leu Glu Ser Glu Gln
530 535 540

Asp Lys Met Ser Gln Gly Phe His Pro Glu Arg Asp Pro Ser Asp Leu
545 550 555 560

Lys Lys Val Lys Ala Val Glu Glu Asn Gly Glu Glu Ala Glu Pro Val
565 570 575

Arg Asn Gly Ala Glu Ser Val Ser Glu Gly Glu Gly Ile Asp Ala Asn
580 585 590

Ser Gly Ser Thr Asp Ser Ser Gly Asp Gly Val Thr Phe Pro Phe Lys
595 600 605

Pro Glu Ser Trp Lys Pro Thr Asp Thr Glu Gly Lys Lys Gln Tyr Asp
610 615 620

Arg Glu Phe Leu Leu Asp Phe Gln Phe Met Pro Ala Cys Ile Gln Lys
625 630 635 640

Pro Glu Gly Leu Pro Pro Ile Ser Asp Val Val Leu Asp Lys Ile Asn
645 650 655

Gln Pro Lys Leu Pro Met Arg Thr Leu Asp Pro Arg Ile Leu Pro Arg
660 665 670

Gly Pro Asp Phe Thr Pro Ala Phe Ala Asp Phe Gly Arg Gln Thr Pro
675 680 685

Gly Gly Arg Gly Val Pro Leu Leu Asn Val Gly Ser Arg Arg Ser Gln
690 695 700

Pro Gly Gln Arg Arg Glu Pro Arg Lys Ile Ile Thr Val Ser Val Lys
705 710 715 720

Glu Asp Val His Leu Lys Lys Ala Glu Asn Ala Trp Lys Pro Ser Gln
725 730 735

Lys Arg Asp Ser Gln Ala Asp Asp Pro Glu Asn Ile Lys Thr Gln Glu
740 745 750

Leu Phe Arg Lys Val Arg Ser Ile Leu Asn Lys Leu Thr Pro Gln Met
755 760 765

Phe Asn Gln Leu Met Lys Gln Val Ser Gly Leu Thr Val Asp Thr Glu
770 775 780

Glu Arg Leu Lys Gly Val Ile Asp Leu Val Phe Glu Lys Ala Ile Asp
785 790 795 800

Glu Pro Ser Phe Ser Val Ala Tyr Ala Asn Met Cys Arg Cys Leu Val
805 810 815

Thr Leu Lys Val Pro Met Ala Asp Lys Pro Gly Asn Thr Val Asn Phe
820 825 830

Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys
835 840 845

Ala Asp Asp Asp Val Phe Glu Lys Lys Gln Lys Glu Leu Glu Ala Ala
850 855 860

Ser Ala Pro Glu Glu Arg Thr Arg Leu His Asp Glu Leu Glu Glu Ala
865 870 875 880

Lys Asp Lys Ala Arg Arg Arg Ser Ile Gly Asn Ile Lys Phe Ile Gly
885 890 895

Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile Met His Asp Cys
900 905 910

Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser Leu Glu Cys Leu
915 920 925

Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala
930 935 940

Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys Ile Val Lys
945 950 955 960

Glu Arg Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln Asp Val Ile
965 970 975

Asp Leu Arg Leu Cys Asn Trp Val Ser Arg Arg Ala Asp Gln Gly Pro
980 985 990

Lys Thr Ile Glu Gln Ile His Lys Glu Ala Lys Ile Glu Glu Gln Glu
995 1000 1005

Glu Gln Arg Lys Val Gln Gln Leu Met Thr Lys Glu Lys Arg Arg
1010 1015 1020

Pro Gly Val Gln Arg Val Asp Glu Gly Gly Trp Asn Thr Val Gln
1025 1030 1035

Gly Ala Lys Asn Ser Arg Val Leu Asp Pro Ser Lys Phe Leu Lys
1040 1045 1050

Ile Thr Lys Pro Thr Ile Asp Glu Lys Ile Gln Leu Val Pro Lys
1055 1060 1065

Ala Gln Leu Gly Ser Trp Gly Lys Gly Ser Ser Gly Gly Ala Lys
1070 1075 1080

Ala Ser Glu Thr Asp Ala Leu Arg Ser Ser Ala Ser Ser Leu Asn
1085 1090 1095

Arg Phe Ser Ala Leu Gln Pro Pro Ala Pro Ser Gly Ser Thr Pro
1100 1105 1110

Ser Thr Pro Val Glu Phe Asp Ser Arg Arg Thr Leu Thr Ser Arg
1115 1120 1125

Gly Ser Met Gly Arg Glu Lys Asn Asp Lys Pro Leu Pro Ser Ala
1130 1135 1140

Thr	Ala	Arg	Pro	Asn	Thr	Phe	Met	Arg	Gly	Gly	Ser	Ser	Lys	Asp
1145						1150					1155			
Leu	Leu	Asp	Asn	Gln	Ser	Gln	Glu	Glu	Gln	Arg	Arg	Glu	Met	Leu
1160						1165					1170			
Glu	Thr	Val	Lys	Gln	Leu	Thr	Gly	Gly	Val	Asp	Val	Glu	Arg	Asn
1175						1180					1185			
Ser	Thr	Glu	Ala	Glu	Arg	Asn	Lys	Thr	Arg	Glu	Ser	Ala	Lys	Pro
1190						1195					1200			
Glu	Ile	Ser	Ala	Met	Ser	Ala	His	Asp	Lys	Ala	Ala	Leu	Ser	Glu
1205						1210					1215			
Glu	Glu	Leu	Glu	Arg	Lys	Ser	Lys	Ser	Ile	Ile	Asp	Glu	Phe	Leu
1220						1225					1230			
His	Ile	Asn	Asp	Phe	Lys	Glu	Ala	Met	Gln	Cys	Val	Glu	Glu	Leu
1235						1240					1245			
Asn	Ala	Gln	Gly	Leu	Leu	His	Val	Phe	Val	Arg	Val	Gly	Val	Glu
1250						1255					1260			
Ser	Thr	Leu	Glu	Arg	Ser	Gln	Ile	Thr	Arg	Asp	His	Met	Gly	Gln
1265						1270					1275			
Leu	Leu	Tyr	Gln	Leu	Val	Gln	Ser	Glu	Lys	Leu	Ser	Lys	Gln	Asp
1280						1285					1290			
Phe	Phe	Lys	Gly	Phe	Ser	Glu	Thr	Leu	Glu	Leu	Ala	Asp	Asp	Met
1295						1300					1305			
Ala	Ile	Asp	Ile	Pro	His	Ile	Trp	Leu	Tyr	Leu	Ala	Glu	Leu	Val
1310						1315					1320			
Thr	Pro	Met	Leu	Lys	Glu	Gly	Gly	Ile	Ser	Met	Arg	Glu	Leu	Thr
1325						1330					1335			
Ile	Glu	Phe	Ser	Lys	Pro	Leu	Leu	Pro	Val	Gly	Arg	Ala	Gly	Val
1340						1345					1350			

Leu Leu Ser Glu Ile Leu His Leu Leu Cys Lys Gln Met Ser His
1355 1360 1365

Lys Lys Val Gly Ala Leu Trp Arg Glu Ala Asp Leu Ser Trp Lys
1370 1375 1380

Asp Phe Leu Pro Glu Gly Glu Asp Val His Asn Phe Leu Leu Glu
1385 1390 1395

Gln Lys Leu Asp Phe Ile Glu Ser Asp Ser Pro Cys Ser Ser Glu
1400 1405 1410

Ala Leu Ser Lys Lys Glu Leu Ser Ala Glu Glu Leu Tyr Lys Arg
1415 1420 1425

Leu Glu Lys Leu Ile Ile Glu Asp Lys Ala Asn Asp Glu Gln Ile
1430 1435 1440

Phe Asp Trp Val Glu Ala Asn Leu Asp Glu Ile Gln Met Ser Ser
1445 1450 1455

Pro Thr Phe Leu Arg Ala Leu Met Thr Ala Val Cys Lys Ala Ala
1460 1465 1470

Ile Ile Ala Asp Ser Ser Thr Phe Arg Val Asp Thr Ala Val Ile
1475 1480 1485

Lys Gln Arg Val Pro Ile Leu Leu Lys Tyr Leu Asp Ser Asp Thr
1490 1495 1500

Glu Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Ser Ile Val
1505 1510 1515

Lys Leu Asp Gln Pro Ala Asn Leu Leu Arg Met Phe Phe Asp Cys
1520 1525 1530

Leu Tyr Asp Glu Glu Val Ile Ser Glu Asp Ala Phe Tyr Lys Trp
1535 1540 1545

Glu Ser Ser Lys Asp Pro Ala Glu Gln Asn Gly Lys Gly Val Ala
1550 1555 1560

Leu Lys Ser Val Thr Ala Phe Phe Thr Trp Leu Arg Glu Ala Glu
1565 1570 1575

Glu Glu Ser Glu Asp Asn
1580

<210> 61
<211> 4752
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 61
aacagccaac cccagacacg gagccccttc ttccagcggc cccagatcca gccaccacgg 60
gcaaccatcc ccaacagcag ccccagcatc cgacccggcg cacagacacc caccgccgtg 120
taccaggcca accagcacat catgatggtg aaccacctgc ccatgcccta ccccgtagcca 180
caaggccctc agtactgcat cccacagtac cggcacagcg gtcctcccta cgtggggccg 240
ccacagcagt accccgttca gcctcctgga ccgggaccct tttatccggg acctggaccc 300
ggagacttcc ccaacgccta cggcacaccc ttctacccca gccagccagt gtaccagagc 360
gcgcccatca tcgtgcccac ccagcagcag cctccacccg ccaagcggga gaagaagacc 420
atccggatcc gggaccccaa ccagggcggc aaggacatca ccgaggagat catgagcggg 480
ggcggcagtc ggaacccac accgcccac ggagaccca ccagcaccac aactccacca 540
cagcagctgc ccagccaggt gcccgagcac agccccgtgg tgtacggcac cgtggaaagc 600
gcccacctgg ccgctagcac ccctgtgacc gccgccagcg accagaagca ggaggagaag 660
cccaaaccg accccgtgct gaagtctccc agccccgtgc tgcgtctggt gctgagcggc 720
gagaagaagg agcaggaggg ccagaccagc gagacaaccg ccatcgtgag catcgccgag 780
ctgccactgc ctcccagtc caccaccgtg agcagcgtgg cccggagcac catcgagcgc 840
cccacaagca gcgccctgag cagccagccc atcttcacca ccgccatcga cgaccggtgc 900
gagctgagca gtccccggga ggacaccatc cccatcccca gcctgaccag ctgcaccgag 960

accagcgacc cactgcccac caacgagaac gacgacgaca tctgcaagaa gccctgcagc	1020
gtggcaccca acgacatccc gctgggtgagc agcaccaacc tgatcaacga gatcaacggc	1080
gtagcgaga aactgtctgc caccgagagc atcgtggaga tcgtcaagca ggaggtgctg	1140
cccctgaccc tggagctaga gatcctggag aaccctccag aggagatgaa gctggagtgc	1200
atccccgccc caatcacccc gagcaccgtg cccagctttc ctccaacccc acccaccca	1260
ccagccagtc cacctcacac tcccgtgatc gtgcctgccg ccgctaccac cgtgtctagc	1320
cccagcgccg ccatcacctg ccagcgggta ctggaggagg acgagagcat ccggacctgc	1380
ctgagcgaag acgccaagga gatccagaac aagatcgagg tggaggccga cggccagacc	1440
gaggaaatcc tggacagcca gaacctgaat tcccggcgaa gccctgtgcc cgcacagatc	1500
gccatcacag tgccaagac ctggaagaag cccaaggacc ggacccggac cacagaggaa	1560
atgctggagg ccgagctgga gctgaaggcc gaggaagagc tgagcatcga caaggtgctg	1620
gagagcgagc aggacaagat gagccagggc ttccaccctg agcgggaccc cagcgacctg	1680
aagaaggtga aggccgtgga ggagaacggc gaggaggcag agcctgtgcg gaacggggcc	1740
gagagcgtga gcgagggcga gggcatcgac gccaacagcg gcagcactga cagcagcggc	1800
gacggcgtga ctttcccctt caagcccgaa agctggaagc ccaccgacac cgaaggaaag	1860
aagcagtacg accgggagtt cctgctggac ttccagttca tgcccgcctg catccagaaa	1920
cccgaggggc tgccaccaat cagcgacgtg gtgctggaca agatcaacca gcccaagctg	1980
cccatgcgga ccctggatcc tcggatcctt cccagaggcc ccgacttcac tcccgccttc	2040
gccgacttcg gcaggcagac cccaggtggc cggggtgtgc ccctgctgaa cgtgggcagc	2100
cgtcggagcc agcctggcca acggcgggaa ccccggaaga tcatcacctg gagcgtgaag	2160
gaggacgtgc atctgaagaa ggccgagaac gcctggaaac ccagccagaa gcgggacagc	2220
caagccgacg accccgagaa catcaagacc caggagctgt tccggaaggt tcggagcatc	2280
ctgaacaagc tgacaccca gatgttcaac cagctgatga agcaggtgag cggcctgacc	2340
gtggacaccg aggagcggct gaagggcgtg atcgacctgg tgtttgagaa agccattgac	2400
gagcccagct tcagcgtggc ctacgccaac atgtgccggt gcctggtgac cctgaaggtg	2460
cccatggccg acaagcccgg caacaccgtg aacttccgga agctgctgct gaaccggtgc	2520
cagaaggagt tcgagaagga caaggcagac gacgacgtgt tcgagaagaa gcagaaagag	2580

ttagaggccg caagcgcccc tgaggagcgg acccggtgc acgacgagct ggaggaggcc	2640
aaagacaagg cccggcggag gagcatcggc aacatcaagt tcatcggcga gctgttcaag	2700
ctgaagatgc tgaccgaggc catcatgcac gactgcgtgg tgaagctgct gaagaaccac	2760
gacgaggaga gcctggagtg cctgtgccgg ctgctgacca ccatcggcaa ggacctggac	2820
ttcgagaagg ccaagccccg gatggaccag tacttcaacc agatggagaa gatcgtgaag	2880
gagcgggaaga ccagcagccg gatccggttc atgctgcagg acgttatcga cctgcggctg	2940
tgcaactggg tgagccggag ggctgaccag ggccccaaga ccatcgagca gatccacaag	3000
gaggctaaga tcgaggagca ggaagagcag cggaaggtgc agcagctgat gaccaaagag	3060
aagcggaggc ctggcgtgca gcgagtggac gagggcggct ggaacaccgt gcagggcgcc	3120
aagaacagcc gggtgctgga cccagcaag ttcctgaaga tcaccaagcc caccatcgac	3180
gagaagatcc agctggtgcc caaggcccag ctcggcagct ggggcaaggg cagctcaggt	3240
ggcgccaagg ccagcgagac cgacgccctg agaagcagcg ccagcagcct gaaccggttc	3300
tctgccctgc agcctcctgc ccctagcggc tccaccccaa gcacaccgt ggagttcgac	3360
agccggcgga ccctgaccag ccggggctcc atgggtcgcg agaagaacga caagccctta	3420
cccagcgcaa ccgctcggcc caacaccttc atgcggggcg gcagctcaa ggacctgctg	3480
gacaaccaga gccaggagga gcagcggcgg gagatgctgg agaccgtgaa gcagctgacc	3540
ggaggcgtgg acgtggagcg gaacagcacc gaggccgagc ggaacaagac ccgggagagc	3600
gccaagcccg agatcagcgc catgagcgcc cacgacaagg ccgccctgag cgaggaggag	3660
ctggaacgga agagcaagag catcatcgac gaggttcctgc acatcaacga cttcaaggag	3720
gccatgcagt gcgtggagga gctgaacgcc cagggcctgc tgcacgtgtt cgtgcgggtg	3780
ggcgtggaga gcaccctgga gcggagccag atcaccggg accacatggg ccagctgctg	3840
taccaactgg tgcaagcgga gaagctgagc aagcaggact tcttcaaggg cttcagcgag	3900
actctggagc tggccgacga catggccatc gacattcccc acatctggct gtacctggcc	3960
gagcttgtga cccaatgct gaaggagggc ggcatcagca tgcgggagct gaccatcgag	4020
ttcagcaagc ccctcttgcc agtcggcaga gcaggagtgc tgctgagcga gatcctgcac	4080
ctgctgtgca agcagatgag ccacaagaag gtgggcgccc tgtggaggga ggccgacctg	4140
agctggaagg acttcctgcc cgaggcgaa gacgtgcaca acttcttgct ggagcagaag	4200

ctggacttca tcgagagcga cagcccctgc agcagcgagg ccctgagcaa gaaggaactg 4260
 agcgccgagg agctgtacaa gcggctggag aagctgatca tcgaggacaa ggccaacgac 4320
 gagcagatct tcgactgggt ggaggccaac ctggacgaga tccagatgag cagccccacc 4380
 ttcttgcggg ccctgatgac cgccgtgtgc aaggccgcca tcatcgccga cagcagcacc 4440
 ttccgggtgg acaccgccgt gatcaagcag cgggtgcccc tcctgctgaa gtacctggac 4500
 agcgacaccg agaaggagct gcaggccctg tacgccctgc aggccagcat cgtgaagctg 4560
 gaccagcccg ccaacctgct gcggatgttc ttcgactgcc tgtacgacga agaggtgac 4620
 agcgaggacg ctttctacaa gtgggagagc agcaaggacc ccgccgagca gaacggcaag 4680
 ggctgtggccc tgaagagcgt gaccgccttc ttcacctggc tgcgagaggc cgaggaggag 4740
 agcgaggaca ac 4752

<210> 62

<211> 540

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 62

Gly Glu Gln Lys Tyr Glu Tyr Lys Ser Asp Gln Trp Lys Pro Leu Asn
 1 5 10 15

Leu Glu Glu Lys Lys Arg Tyr Asp Arg Glu Phe Leu Leu Gly Phe Gln
 20 25 30

Phe Ile Phe Ala Ser Met Gln Lys Pro Glu Gly Leu Pro His Ile Ser
 35 40 45

Asp Val Val Leu Asp Lys Ala Asn Lys Thr Pro Leu Arg Pro Leu Asp
 50 55 60

Pro Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser
 65 70 75 80

Phe Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg

85

90

95

Gly Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro
 100 105 110

Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala
 115 120 125

Thr Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala
 130 135 140

Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu
 145 150 155 160

Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg
 165 170 175

Ser Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys
 180 185 190

Gln Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val
 195 200 205

Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val
 210 215 220

Ala Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr
 225 230 235 240

Thr Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn
 245 250 255

Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe
 260 265 270

Glu Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg
 275 280 285

Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg
 290 295 300

Arg Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys
305 310 315 320

Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu Lys
325 330 335

Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr
340 345 350

Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln
355 360 365

Tyr Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser
370 375 380

Arg Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn
385 390 395 400

Trp Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile
405 410 415

His Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val Gln
420 425 430

Gln Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly
435 440 445

Pro Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn
450 455 460

Thr Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu
465 470 475 480

Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe
485 490 495

Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser
500 505 510

Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr

515

520

525

Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala
 530 535 540

<210> 63

<211> 1620

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 63

```

ggcgagcaga agtacgagta caagagcgac cagtggaagc ctctgaacct ggaggagaag      60
aagagatacg acagagagtt cctgctgggc ttccagttca tattcgcgag catgcagaag      120
cctgaaggtc tgcctcacat cagcgacgtg gtgctggaca aggccaacaa gaccctcttt      180
agacctctgg accctaccag actgcagggc atcaactgcg gccctgactt cactccttct      240
ttcgcaaacc tgggcagaac caccctgagc accagaggcc ctcttagagg tgggtcccggc      300
ggagaactcc ccaggggtcc tgccggcctg ggccctagac gctcccagca aggtcctaga      360
aaggagccaa ggaagatcat cgccaccgtg ctgatgaccg aggacatcaa gctgaacaaa      420
gctgagaagg cctggaagcc tagcagcaag agaaccgccg ccgacaagga cagaggcgag      480
gaggacgccg acggatccaa gaccaggac ctgttcagaa gagtgagaag catcctcaac      540
aagctgaccc ctcatgtgtt ccagcagctg atgaagcagg tgacgcagct cgccatcgac      600
accgaggaga gactgaaggg cgtgatcgac ctgatctttg agaaggctat ctcagagcct      660
aacttcagcg tggcctacgc caacatgtgc cgttgcctga tggcattgaa ggtgccaacc      720
accgagaagc ctactgtgac cgtcaatttc cgtaaactgc tgctgaaccg gtgccagaaa      780
gagttcgaga aggataagga cgacgacgag gtcttcgaga agaaacagaa agaatggac      840
gaggccgcca ccgcagagga aaggggcca ttaaaggagg agctggagga ggccagagac      900
atcgccagac ggcgttctct gggcaacatc aagttcatag gtgagctgtt caagctaaag      960
atgctcaccg aggccataat gcacgactgc gtggtgaagc tactgaagaa ccacgacgaa     1020
gaaagcctgg agtgcctgtg cagactgctg accaccatcg gcaaggacct ggacttcgag     1080

```

aaggcaaagc ctcgaatgga ccagtacttc aaccagatgg agaagattat caaggagaag 1140
aagaccagca gcagaatcag attcatgctg caggacgtac tggacctgcg cggaagcaac 1200
tgggtgccaa ggagagggga ccaaggacca aagaccatcg accagatcca caaggaagcg 1260
gagatggagg agcacagaga gcacataaag gtgcagcagc ttatggccaa gggcagcgac 1320
aagcgaagag gcggcccgcc cggacctcct atcagcagag gccttcctct ggtagacgac 1380
ggcggctgga acaccgtgcc tatctctaag ggctccagac ctatcgacac ttcccgtctt 1440
accaagatca ccaagccagg atctattgac agcaacaacc agctgttcgc cccaggagga 1500
agacttagct ggggcaaggg aagttccggc ggatccggcg ccaagccttc cgacgccgcc 1560
agcgaggctg ccagacctgc caccagcacc ttgaaccgct tttccgctct gcagcaagct 1620

<210> 64

<211> 1584

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 64

Asn Ser Gln Pro Gln Thr Arg Ser Pro Phe Phe Gln Arg Pro Gln Ile
1 5 10 15

Gln Pro Pro Arg Ala Thr Ile Pro Asn Ser Ser Pro Ser Ile Arg Pro
20 25 30

Gly Ala Gln Thr Pro Thr Ala Val Tyr Gln Ala Asn Gln His Ile Met
35 40 45

Met Val Asn His Leu Pro Met Pro Tyr Pro Val Pro Gln Gly Pro Gln
50 55 60

Tyr Cys Ile Pro Gln Tyr Arg His Ser Gly Pro Pro Tyr Val Gly Pro
65 70 75 80

Pro Gln Gln Tyr Pro Val Gln Pro Pro Gly Pro Gly Pro Phe Tyr Pro
85 90 95

Gly Pro Gly Pro Gly Asp Phe Pro Asn Ala Tyr Gly Thr Pro Phe Tyr
100 105 110

Pro Ser Gln Pro Val Tyr Gln Ser Ala Pro Ile Ile Val Pro Thr Gln
115 120 125

Gln Gln Pro Pro Pro Ala Lys Arg Glu Lys Lys Thr Ile Arg Ile Arg
130 135 140

Asp Pro Asn Gln Gly Gly Lys Asp Ile Thr Glu Glu Ile Met Ser Gly
145 150 155 160

Gly Gly Ser Arg Asn Pro Thr Pro Pro Ile Gly Arg Pro Thr Ser Thr
165 170 175

Pro Thr Pro Pro Gln Gln Leu Pro Ser Gln Val Pro Glu His Ser Pro
180 185 190

Val Val Tyr Gly Thr Val Glu Ser Ala His Leu Ala Ala Ser Thr Pro
195 200 205

Val Thr Ala Ala Ser Asp Gln Lys Gln Glu Glu Lys Pro Lys Pro Asp
210 215 220

Pro Val Leu Lys Ser Pro Ser Pro Val Leu Arg Leu Val Leu Ser Gly
225 230 235 240

Glu Lys Lys Glu Gln Glu Gly Gln Thr Ser Glu Thr Thr Ala Ile Val
245 250 255

Ser Ile Ala Glu Leu Pro Leu Pro Pro Ser Pro Thr Thr Val Ser Ser
260 265 270

Val Ala Arg Ser Thr Ile Ala Ala Pro Thr Ser Ser Ala Leu Ser Ser
275 280 285

Gln Pro Ile Phe Thr Thr Ala Ile Asp Asp Arg Cys Glu Leu Ser Ser
290 295 300

Pro Arg Glu Asp Thr Ile Pro Ile Pro Ser Leu Thr Ser Cys Thr Glu
305 310 315 320

Thr Ser Asp Pro Leu Pro Thr Asn Glu Asn Asp Asp Asp Ile Cys Lys
325 330 335

Lys Pro Cys Ser Val Ala Pro Asn Asp Ile Pro Leu Val Ser Ser Thr
340 345 350

Asn Leu Ile Asn Glu Ile Asn Gly Val Ser Glu Lys Leu Ser Ala Thr
355 360 365

Glu Ser Ile Val Glu Ile Val Lys Gln Glu Val Leu Pro Leu Thr Leu
370 375 380

Glu Leu Glu Ile Leu Glu Asn Pro Pro Glu Glu Met Lys Leu Glu Cys
385 390 395 400

Ile Pro Ala Pro Ile Thr Pro Ser Thr Val Pro Ser Phe Pro Pro Thr
405 410 415

Pro Pro Thr Pro Pro Ala Ser Pro Pro His Thr Pro Val Ile Val Pro
420 425 430

Ala Ala Ala Thr Thr Val Ser Ser Pro Ser Ala Ala Ile Thr Val Gln
435 440 445

Arg Val Leu Glu Glu Asp Glu Ser Ile Arg Thr Cys Leu Ser Glu Asp
450 455 460

Ala Lys Glu Ile Gln Asn Lys Ile Glu Val Glu Ala Asp Gly Gln Thr
465 470 475 480

Glu Glu Ile Leu Asp Ser Gln Asn Leu Asn Ser Arg Arg Ser Pro Val
485 490 495

Pro Ala Gln Ile Ala Ile Thr Val Pro Lys Thr Trp Lys Lys Pro Lys
500 505 510

Asp Arg Thr Arg Thr Thr Glu Glu Met Leu Glu Ala Glu Leu Glu Leu
515 520 525

Lys Ala Glu Glu Glu Leu Ser Ile Asp Lys Val Leu Glu Ser Glu Gln
530 535 540

Asp Lys Met Ser Gln Gly Phe His Pro Glu Arg Asp Pro Ser Asp Leu
545 550 555 560

Lys Lys Val Lys Ala Val Glu Glu Asn Gly Glu Glu Ala Glu Pro Val
565 570 575

Arg Asn Gly Ala Glu Ser Val Ser Glu Gly Glu Gly Ile Asp Ala Asn
580 585 590

Ser Gly Ser Thr Asp Ser Ser Gly Asp Gly Val Thr Phe Pro Phe Lys
595 600 605

Pro Glu Ser Trp Lys Pro Thr Asp Thr Glu Gly Lys Lys Gln Tyr Asp
610 615 620

Arg Glu Phe Leu Leu Asp Phe Gln Phe Met Pro Ala Cys Ile Gln Lys
625 630 635 640

Pro Glu Gly Leu Pro Pro Ile Ser Asp Val Val Leu Asp Lys Ile Asn
645 650 655

Gln Pro Lys Leu Pro Met Arg Thr Leu Asp Pro Arg Ile Leu Pro Arg
660 665 670

Gly Pro Asp Phe Thr Pro Ala Phe Ala Asp Phe Gly Arg Gln Thr Pro
675 680 685

Gly Gly Arg Gly Val Pro Leu Leu Asn Val Gly Ser Arg Arg Ser Gln
690 695 700

Pro Gly Gln Arg Arg Glu Pro Arg Lys Ile Ile Thr Val Ser Val Lys
705 710 715 720

Glu Asp Val His Leu Lys Lys Ala Glu Asn Ala Trp Lys Pro Ser Gln
725 730 735

Lys Arg Asp Ser Gln Ala Asp Asp Pro Glu Asn Ile Lys Thr Gln Glu
740 745 750

Leu Phe Arg Lys Val Arg Ser Ile Leu Asn Lys Leu Thr Pro Gln Met
755 760 765

Phe Asn Gln Leu Met Lys Gln Val Ser Gly Leu Thr Val Asp Thr Glu
770 775 780

Glu Arg Leu Lys Gly Val Ile Asp Leu Val Phe Glu Lys Ala Ile Asp
785 790 795 800

Glu Pro Ser Phe Ser Val Ala Tyr Ala Asn Met Cys Arg Cys Leu Val
805 810 815

Thr Leu Lys Val Pro Met Ala Asp Lys Pro Gly Asn Thr Val Asn Phe
820 825 830

Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys
835 840 845

Ala Asp Asp Asp Val Phe Glu Lys Lys Gln Lys Glu Leu Glu Ala Ala
850 855 860

Ser Ala Pro Glu Glu Arg Thr Arg Leu His Asp Glu Leu Glu Glu Ala
865 870 875 880

Lys Asp Lys Ala Arg Arg Arg Ser Ile Gly Asn Ile Lys Phe Ile Gly
885 890 895

Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile Met His Asp Cys
900 905 910

Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser Leu Glu Cys Leu
915 920 925

Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala
930 935 940

Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys Ile Val Lys
945 950 955 960

Glu Arg Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln Asp Val Ile
965 970 975

Asp Leu Arg Leu Cys Asn Trp Val Ser Arg Arg Ala Asp Gln Gly Pro
980 985 990

Lys Thr Ile Glu Gln Ile His Lys Glu Ala Lys Ile Glu Glu Gln Glu
995 1000 1005

Glu Gln Arg Lys Val Gln Gln Leu Met Thr Lys Glu Lys Arg Arg
1010 1015 1020

Pro Gly Val Gln Arg Val Asp Glu Gly Gly Trp Asn Thr Val Gln
1025 1030 1035

Gly Ala Lys Asn Ser Arg Val Leu Asp Pro Ser Lys Phe Leu Lys
1040 1045 1050

Ile Thr Lys Pro Thr Ile Asp Glu Lys Ile Gln Leu Val Pro Lys
1055 1060 1065

Ala Gln Leu Gly Ser Trp Gly Lys Gly Ser Ser Gly Gly Ala Lys
1070 1075 1080

Ala Ser Glu Thr Asp Ala Leu Arg Ser Ser Ala Ser Ser Leu Asn
1085 1090 1095

Arg Phe Ser Ala Leu Gln Pro Pro Ala Pro Ser Gly Ser Thr Pro
1100 1105 1110

Ser Thr Pro Val Glu Phe Asp Ser Arg Arg Thr Leu Thr Ser Arg
1115 1120 1125

Gly Ser Met Gly Arg Glu Lys Asn Asp Lys Pro Leu Pro Ser Ala
1130 1135 1140

Thr Ala Arg Pro Asn Thr Phe Met Arg Gly Gly Ser Ser Lys Asp
1145 1150 1155

Leu Leu Asp Asn Gln Ser Gln Glu Glu Gln Arg Arg Glu Met Leu
1160 1165 1170

Glu Thr Val Lys Gln Leu Thr Gly Gly Val Asp Val Glu Arg Asn
1175 1180 1185

Ser Thr Glu Ala Glu Arg Asn Lys Thr Arg Glu Ser Ala Lys Pro
1190 1195 1200

Glu Ile Ser Ala Met Ser Ala His Asp Lys Ala Ala Leu Ser Glu
1205 1210 1215

Glu Glu Leu Glu Arg Lys Ser Lys Ser Ile Ile Asp Glu Phe Leu
1220 1225 1230

His Ile Asn Asp Phe Lys Glu Ala Met Gln Cys Val Glu Glu Leu
1235 1240 1245

Asn Ala Gln Gly Leu Leu His Val Phe Val Arg Val Gly Val Glu
1250 1255 1260

Ser Thr Leu Glu Arg Ser Gln Ile Thr Arg Asp His Met Gly Gln
1265 1270 1275

Leu Leu Tyr Gln Leu Val Gln Ser Glu Lys Leu Ser Lys Gln Asp
1280 1285 1290

Phe Phe Lys Gly Phe Ser Glu Thr Leu Glu Leu Ala Asp Asp Met
1295 1300 1305

Ala Ile Asp Ile Pro His Ile Trp Leu Tyr Leu Ala Glu Leu Val
1310 1315 1320

Thr Pro Met Leu Lys Glu Gly Gly Ile Ser Met Arg Glu Leu Thr
1325 1330 1335

Ile Glu Phe Ser Lys Pro Leu Leu Pro Val Gly Arg Ala Gly Val
1340 1345 1350

Leu Leu Ser Glu Ile Leu His Leu Leu Cys Lys Gln Met Ser His
1355 1360 1365

Lys Lys Val Gly Ala Leu Trp Arg Glu Ala Asp Leu Ser Trp Lys
1370 1375 1380

Asp Phe Leu Pro Glu Gly Glu Asp Val His Asn Phe Leu Leu Glu
1385 1390 1395

Gln Lys Leu Asp Phe Ile Glu Ser Asp Ser Pro Cys Ser Ser Glu
1400 1405 1410

Ala Leu Ser Lys Lys Glu Leu Ser Ala Glu Glu Leu Tyr Lys Arg
1415 1420 1425

Leu Glu Lys Leu Ile Ile Glu Asp Lys Ala Asn Asp Glu Gln Ile
1430 1435 1440

Phe Asp Trp Val Glu Ala Asn Leu Asp Glu Ile Gln Met Ser Ser
1445 1450 1455

Pro Thr Phe Leu Arg Ala Leu Met Thr Ala Val Cys Lys Ala Ala
1460 1465 1470

Ile Ile Ala Asp Ser Ser Thr Phe Arg Val Asp Thr Ala Val Ile
1475 1480 1485

Lys Gln Arg Val Pro Ile Leu Leu Lys Tyr Leu Asp Ser Asp Thr
1490 1495 1500

Glu Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Ser Ile Val
1505 1510 1515

Lys Leu Asp Gln Pro Ala Asn Leu Leu Arg Met Phe Phe Asp Cys
1520 1525 1530

Leu Tyr Asp Glu Glu Val Ile Ser Glu Asp Ala Phe Tyr Lys Trp
1535 1540 1545

Glu Ser Ser Lys Asp Pro Ala Glu Gln Asn Gly Lys Gly Val Ala
1550 1555 1560

Leu Lys Ser Val Thr Ala Phe Phe Thr Trp Leu Arg Glu Ala Glu
1565 1570 1575

Glu Glu Ser Glu Asp Asn
1580

<210> 65
<211> 4755
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 65
agcaacagcc aaccccagac acggagcccc ttcttccagc ggccccagat ccagccacca 60
cgggcaacca tccccaacag cagccccagc atccgacccg gcgcacagac acccaccgcc 120
gtgtaccagg ccaaccagca catcatgatg gtgaaccacc tgcccatgcc ctaccccgtg 180
ccacaaggcc ctcagtactg catcccacag taccggcaca gcggtcctcc ctacgtgggg 240
ccgccacagc agtaccccgt tcagcctcct ggaccgggac ctttttatcc gggacctgga 300
cccggagact tccccaacgc ctacggcaca cccttctacc ccagccagcc agtgtaccag 360
agcgcgcca tcatcgtgcc caccagcag cagcctccac ccgccaagcg ggagaagaag 420
accatccgga tccgggaccc caaccagggc ggcaaggaca tcaccgagga gatcatgagc 480
ggaggcggca gtcggaaccc cacaccgccc atcggcagac ccaccagcac cccaactcca 540
ccacagcagc tgcccagcca ggtgcccag cagagccccg tgggtgtacgg caccgtggaa 600
agcggccacc tggccgctag caccctgtg accgccgcca gcgaccagaa gcaggaggag 660
aagcccaaac ccgaccccgt gctgaagtct ccagccccg tgctgcgtct ggtgctgagc 720
ggcgagaaga aggagcagga gggccagacc agcgagacaa ccgccatcgt gagcatcgcc 780
gagctgccac tgcctcccag tcccaccacc gtgagcagcg tggcccggag caccatcgca 840
gcgcccacaa gcagcgccct gagcagccag cccatcttca ccaccgccat cgacgaccgg 900
tgcgagctga gcagtccccg ggaggacacc atccccatcc ccagcctgac cagctgcacc 960
gagaccagcg acccactgcc caccaacgag aacgacgacg acatctgcaa gaagccctgc 1020
agcgtggcac ccaacgacat cccgctgggtg agcagcacca acctgatcaa cgagatcaac 1080
ggcgtttagc agaaactgtc tgccaccgag agcatcgtgg agatcgtcaa gcaggaggtg 1140

ctgcccctga ccctggagct agagatcctg gagaaccctc cagaggagat gaagctggag	1200
tgcattccccg cccaatcac cccgagcacc gtgcccagct ttcttccaac cccaccacc	1260
ccaccagcca gtccacctca cactcccgtg atcgtgcctg ccgccgtac caccgtgtct	1320
agccccagcg ccgccatcac cgtccagcgg gtactggagg aggacgagag catccggacc	1380
tgcttgagcg aagacgcaa ggagatccag aacaagatcg aggtggaggc cgacggccag	1440
accgaggaaa tcctggacag ccagaacctg aattcccggc gaagccctgt gcccgcacag	1500
atgccatca cagtgccaa gacctggaag aagcccaagg accggacccg gaccacagag	1560
gaaatgctgg aggccgagct ggagctgaag gccgaggaag agctgagcat cgacaagggtg	1620
ctggagagcg agcaggacaa gatgagccag ggcttccacc ctgagcggga cccagcgac	1680
ctgaagaagg tgaaggccgt ggaggagaac ggcgaggagg cagagcctgt gcggaacggg	1740
gccgagagcg tgagcgaggg cgagggcatc gacgccaaca gcggcagcac tgacagcagc	1800
ggcgacggcg tgaccttccc cttcaagccc gaaagctgga agcccaccga caccgaagga	1860
aagaagcagt acgaccggga gttcctgctg gacttccagt tcatgcccgc ctgcatccag	1920
aaaccgagg ggctgccacc aatcagcgac gtggtgctgg acaagatcaa ccagcccaag	1980
ctgcccattgc ggaccctgga tcctcgatc cttcccagag gccccgactt cactcccgcc	2040
ttcgccgact tcggcaggca gaccccaggt ggtcgggggtg tgcccctgct gaacgtgggc	2100
agccgtcggga gccagcctgg ccaacggcgg gaaccccgga agatcatcac cgtgagcgtg	2160
aaggaggacg tgcatctgaa gaaggccgag aacgcctgga aaccagcca gaagcgggac	2220
agccaagccg acgaccccga gaacatcaag acccaggagc tgttcggaa ggttcggagc	2280
atcctgaaca agctgacacc ccagatgttc aaccagctga tgaagcaggt gagcggcctg	2340
accgtggaca ccgaggagcg gctgaagggc gtgatcgacc tggtgtttga gaaagccatt	2400
gacgagcca gcttcagcgt ggcctacgcc aacatgtgcc ggtgcctggt gaccctgaag	2460
gtgcccattgg ccgacaagcc cggcaacacc gtgaacttcc ggaagctgct gctgaaccgg	2520
tgccagaagg agttcgagaa ggacaaggca gacgacgacg tgttcgagaa gaagcagaaa	2580
gagttagagg ccgcaagcgc ccctgaggag cggacccggc tgcacgacga gctggaggag	2640
gccaaagaca agggccggcg gaggagcatc ggcaacatca agttcatcgg cgagctgttc	2700
aagctgaaga tgctgaccga ggccatcatg cacgactgcg tggtgaagct gctgaagaac	2760

cacgacgagg agagcctgga gtgcctgtgc cggctgtctga ccaccatcgg caaggacctg	2820
gacttcgaga aggccaagcc ccggatggac cagtacttca accagatgga gaagatcgtg	2880
aaggagcggga agaccagcag ccggatccgg ttcattgtgc aggacgttat cgacctgcgg	2940
ctgtgcaact ggggtgagccg gagggctgac cagggcccca agaccatcga gcagatccac	3000
aaggaggcta agatcgagga gcaggaagag cagcgggaagg tgcagcagct gatgacaaaa	3060
gagaagcggga ggcctggcgt gcagcgagtgc gacgagggcg gctggaacac cgtgcagggc	3120
gccaagaaca gccgggtgct ggaccccagc aagttcctga agatcaccaa gccaccatc	3180
gacgagaaga tccagctggt gcccgaaggcc cagctcggca gctggggcaa gggcagctca	3240
ggtggcgcca aggccagcga gaccgacgcc ctgagaagca gcgccagcag cctgaaccgg	3300
ttctctgccc tgcagcctcc tgcccctagc ggctccaccc caagcacacc cgtggagttc	3360
gacagccggc ggacctgac cagccggggc tccatgggtc gcgagaagaa cgacaagccc	3420
ttaccagcg caaccgctcg gcccacacc ttcattgcggg gcggcagctc caaggacctg	3480
ctggacaacc agagccagga ggagcagcgg cgggagatgc tggagaccgt gaagcagctg	3540
accggaggcg tggacgtgga gcggaacagc accgaggccg agcgggaaca gaccggggag	3600
agcgccaagc ccgagatcag cgccatgagc gcccacgaca aggccgccct gagcaggag	3660
gagctggaac ggaagagcaa gagcatcatc gacgagttcc tgcacatcaa cgacttcaag	3720
gaggccatgc agtgcgtgga ggagctgaac gccagggcc tgctgcacgt gttcgtgcgg	3780
gtgggcgtgg agagcacctt ggagcggagc cagatcaccc gggaccacat gggccagctg	3840
ctgtaccaac tggtgcagag cgagaagctg agcaagcagg acttcttcaa gggcttcagc	3900
gagactctgg agctggccga cgacatggcc atcgacattc cccacatctg gctgtacctg	3960
gccgagcttg tgaccccaat gctgaaggag ggcggcatca gcatgcggga gctgaccatc	4020
gagttcagca agcccctctt gccagtcggc agagcaggag tgctgctgag cgagatcctg	4080
cacctgctgt gcaagcagat gagccacaag aaggtgggcg ccctgtggag ggaggccgac	4140
ctgagctgga aggacttcct gcccgagggc gaagacgtgc acaacttctt gctggagcag	4200
aagctggact tcatcgagag cgacagcccc tgcagcagcg aggccctgag caagaaggaa	4260
ctgagcgccg aggagctgta caagcggctg gagaagctga tcatcgagga caaggccaac	4320
gacgagcaga tcttcgactg ggtggaggcc aacctggacg agatccagat gagcagcccc	4380

accttcctgc gggccctgat gaccgccgtg tgcaaggccg ccatcatcgc cgacagcagc 4440
accttcgagg tggacaccgc cgtgatcaag cagcgggtgc ccatcctgct gaagtacctg 4500
gacagcgaca ccgagaagga gctgcaggcc ctgtacgcc tgcaggccag catcgtgaag 4560
ctggaccagc ccgccaacct gctgcggatg ttcttcgact gcctgtacga cgaagaggtg 4620
atcagcgagg acgccttcta caagtgggag agcagcaagg accccgccga gcagaacggc 4680
aagggcgtgg ccctgaagag cgtgaccgcc ttcttcacct ggctgcgaga ggccgaggag 4740
gagagcgagg acaac 4755

<210> 66

<211> 418

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 66

Arg Thr Leu Asp Pro Arg Ile Leu Pro Arg Gly Pro Asp Phe Thr Pro
1 5 10 15

Ala Phe Ala Asp Phe Gly Arg Gln Thr Pro Gly Gly Arg Gly Val Pro
20 25 30

Leu Leu Asn Val Gly Ser Arg Arg Ser Gln Pro Gly Gln Arg Arg Glu
35 40 45

Pro Arg Lys Ile Ile Thr Val Ser Val Lys Glu Asp Val His Leu Lys
50 55 60

Lys Ala Glu Asn Ala Trp Lys Pro Ser Gln Lys Arg Asp Ser Gln Ala
65 70 75 80

Asp Asp Pro Glu Asn Ile Lys Thr Gln Glu Leu Phe Arg Lys Val Arg
85 90 95

Ser Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Asn Gln Leu Met Lys
100 105 110

Gln Val Ser Gly Leu Thr Val Asp Thr Glu Glu Arg Leu Lys Gly Val
115 120 125

Ile Asp Leu Val Phe Glu Lys Ala Ile Asp Glu Pro Ser Phe Ser Val
130 135 140

Ala Tyr Ala Asn Met Cys Arg Cys Leu Val Thr Leu Lys Val Pro Met
145 150 155 160

Ala Asp Lys Pro Gly Asn Thr Val Asn Phe Arg Lys Leu Leu Leu Asn
165 170 175

Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Ala Asp Asp Asp Val Phe
180 185 190

Glu Lys Lys Gln Lys Glu Leu Glu Ala Ala Ser Ala Pro Glu Glu Arg
195 200 205

Thr Arg Leu His Asp Glu Leu Glu Glu Ala Lys Asp Lys Ala Arg Arg
210 215 220

Arg Ser Ile Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys
225 230 235 240

Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu Lys
245 250 255

Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr
260 265 270

Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln
275 280 285

Tyr Phe Asn Gln Met Glu Lys Ile Val Lys Glu Arg Lys Thr Ser Ser
290 295 300

Arg Ile Arg Phe Met Leu Gln Asp Val Ile Asp Leu Arg Leu Cys Asn
305 310 315 320

Trp Val Ser Arg Arg Ala Asp Gln Gly Pro Lys Thr Ile Glu Gln Ile

325

330

335

His Lys Glu Ala Lys Ile Glu Glu Gln Glu Glu Gln Arg Lys Val Gln
 340 345 350

Gln Leu Met Thr Lys Glu Lys Arg Arg Pro Gly Val Gln Arg Val Asp
 355 360 365

Glu Gly Gly Trp Asn Thr Val Gln Gly Ala Lys Asn Ser Arg Val Leu
 370 375 380

Asp Pro Ser Lys Phe Leu Lys Ile Thr Lys Pro Thr Ile Asp Glu Lys
 385 390 395 400

Ile Gln Leu Val Pro Lys Ala Gln Leu Gly Ser Trp Gly Lys Gly Ser
 405 410 415

Ser Gly

<210> 67

<211> 1257

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 67

agccggaccc tggatcctcg gatccttccc agaggccccc acttcactcc cgccttcgcc	60
gacttcggca ggcagacccc aggtggtcgg ggtgtgcccc tgctgaacgt gggcagccgt	120
cggagccagc ctggccaacg gcgggaaccc cggaagatca tcaccgtgag cgtgaaggag	180
gacgtgcatc tgaagaaggc cgagaacgcc tggaaaccca gccagaagcg ggacagccaa	240
gccgacgacc ccgagaacat caagaccag gagctgttcc ggaaggttcg gagcatcctg	300
aacaagctga caccacagat gttcaaccag ctgatgaagc aggtgagcgg cctgaccgtg	360
gacaccgagg agcggctgaa gggcgtgatc gacctggtgt ttgagaaagc cattgacgag	420
cccagcttca gcgtggccta cgccaacatg tgccggtgcc tggtgaccct gaaggtgccc	480

atggccgaca agccccggcaa caccgtgaac ttccggaagc tgctgctgaa ccggtgccag 540
 aaggagtctg agaaggacaa ggcagacgac gacgtgttcg agaagaagca gaaagagtta 600
 gaggccgcaa gcgcccctga ggagcggacc cggctgcacg acgagctgga ggaggccaaa 660
 gacaaggccc ggcggaggag catcggcaac atcaagttca tcggcgagct gttcaagctg 720
 aagatgctga ccgaggccat catgcacgac tgcgtggtga agctgctgaa gaaccacgac 780
 gaggagagcc tggagtgcct gtgccggctg ctgaccacca tcggcaagga cctggacttc 840
 gagaaggcca agccccggat ggaccagtac ttcaaccaga tggagaagat cgtgaaggag 900
 cggaagacca gcagccggat ccggttcatg ctgcaggacg ttatcgacct gcggctgtgc 960
 aactgggtga gccggagggc tgaccagggc cccaagacca tcgagcagat ccacaaggag 1020
 gctaagatcg aggagcagga agagcagcgg aaggtgcagc agctgatgac caaagagaag 1080
 cggaggcctg gcgtgcagcg agtggacgag ggcggctgga acaccgtgca gggcgccaag 1140
 aacagccggg tgctggaccc cagcaagttc ctgaagatca ccaagcccac catcgacgag 1200
 aagatccagc tggtgcccaa ggcccagctc ggcagctggg gcaagggcag ctcaggt 1257

<210> 68

<211> 600

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 68

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
 1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
 20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
 35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
 50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asp Pro Thr Arg Leu Gln
115 120 125

Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly
130 135 140

Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly
145 150 155 160

Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln
165 170 175

Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr
180 185 190

Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser
195 200 205

Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly
210 215 220

Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys
225 230 235 240

Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu
245 250 255

Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe
260 265 270

Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met
275 280 285

Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr
290 295 300

Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu
305 310 315 320

Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys
325 330 335

Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu
340 345 350

Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn
355 360 365

Ile Lys Ala Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala
370 375 380

Ile Met His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu
385 390 395 400

Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu
405 410 415

Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met
420 425 430

Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met
435 440 445

Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg
450 455 460

Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu
465 470 475 480

Met Glu Glu His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys
485 490 495

Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg
500 505 510

Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser
515 520 525

Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys
530 535 540

Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly Gly Arg
545 550 555 560

Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala Lys Pro Ser
565 570 575

Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg
580 585 590

Phe Ser Ala Leu Gln Gln Ala Val
595 600

<210> 69

<211> 1800

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 69

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggcccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagcgacc ctaccagact gcagggcac aactgcggcc ctgacttcac tccttctttc	420
gcaaacctgg gcagaaccac cctgagcacc agaggccctc ctagaggtgg tcccggcgga	480

gaactcccca ggggtcctgc cggcctgggc cctagacgct cccagcaagg tcctagaaag	540
gagccaagga agatcatcgc caccgtgctg atgaccgagg acatcaagct gaacaaagct	600
gagaaggcct ggaagcctag cagcaagaga accgccgccg acaaggacag aggcgaggag	660
gacgccgacg gatccaagac ccaggacctg ttcagaagag tgagaagcat cctcaacaag	720
ctgacccttc agatgttcca gcagctgatg aagcaggtga cgcagctcgc catcgacacc	780
gaggagagac tgaagggcgt gatcgacctg atctttgaga aggctatctc agagcctaac	840
ttcagcgtgg cctacgcaa catgtgccgt tgcctgatgg cattgaaggt gccaaccacc	900
gagaagccta ctgtgaccgt caatttccgt aaactgctgc tgaaccggtg ccagaaagag	960
ttcgagaagg ataaggacga cgacgaggtc ttcgagaaga aacagaaaga aatggacgag	1020
gccgccaccg cagaggaaag gggccgatta aaggaggagc tggaggaggc cagagacatc	1080
gccagacggc gttctctggg caacatcaag gcgataggtg agctgttcaa gctaaagatg	1140
ctcaccgagg ccataatgca cgactgcgtg gtgaagctac tgaagaacca cgacgaagaa	1200
agcctggagt gcctgtgcag actgctgacc accatcgga aggacctgga cttcgagaag	1260
gcaaagcctc gaatggacca gtacttcaac cagatggaga agattatcaa ggagaagaag	1320
accagcagca gaatcagatt catgctgcag gacgtactgg acctgcgcgg aagcaactgg	1380
gtgccaagga gaggggacca aggaccaaag accatcgacc agatccacaa ggaagcggag	1440
atggaggagc acagagagca cataaagggtg cagcagctta tggccaaggg cagcgacaag	1500
cgaagaggcg gcccggccgg acctcctatc agcagaggcc ttcctctggt agacgacggc	1560
ggctggaaca ccgtgcctat ctctaagggtc tccagaccta tcgacacttc ccgtcttacc	1620
aagatcacca agccaggatc tattgacagc aacaaccagc tgttcgcccc aggaggaaga	1680
cttagctggg gcaagggaag ttccggcgga tccggcgcca agccttccga cgccgccagc	1740
gaggctgcca gacctgccac cagcaccttg aaccgctttt ccgctctgca gcaagctgtg	1800

<210> 70

<211> 662

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic

polypeptide"

<400> 70

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Gly Glu Gln Lys Tyr Glu
115 120 125

Tyr Lys Ser Asp Gln Trp Lys Pro Leu Asn Leu Glu Glu Lys Lys Arg
130 135 140

Tyr Asp Arg Glu Phe Leu Leu Gly Phe Gln Phe Ile Phe Ala Ser Met
145 150 155 160

Gln Lys Pro Glu Gly Leu Pro His Ile Ser Asp Val Val Leu Asp Lys
165 170 175

Ala Asn Lys Thr Pro Leu Arg Pro Leu Asp Pro Thr Arg Leu Gln Gly
180 185 190

Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe Ala Asn Leu Gly Arg
195 200 205

Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly Gly Pro Gly Gly Glu
210 215 220

Leu Pro Arg Gly Pro Ala Gly Leu Gly Pro Arg Arg Ser Gln Gln Gly
225 230 235 240

Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala Thr Val Leu Met Thr Glu
245 250 255

Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala Trp Lys Pro Ser Ser Lys
260 265 270

Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu Glu Asp Ala Asp Gly Ser
275 280 285

Lys Thr Gln Asp Leu Phe Arg Arg Val Arg Ser Ile Leu Asn Lys Leu
290 295 300

Thr Pro Gln Met Phe Gln Gln Leu Met Lys Gln Val Thr Gln Leu Ala
305 310 315 320

Ile Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Ile Phe Glu
325 330 335

Lys Ala Ile Ser Glu Pro Asn Phe Ser Val Ala Tyr Ala Asn Met Cys
340 345 350

Arg Cys Leu Met Ala Leu Lys Val Pro Thr Thr Glu Lys Pro Thr Val
355 360 365

Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe
370 375 380

Glu Lys Asp Lys Asp Asp Asp Glu Val Phe Glu Lys Lys Gln Lys Glu
385 390 395 400

Met Asp Glu Ala Ala Thr Ala Glu Glu Arg Gly Arg Leu Lys Glu Glu
405 410 415

Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg Arg Ser Leu Gly Asn Ile

420

425

430

Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile
 435 440 445

Met His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser
 450 455 460

Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp
 465 470 475 480

Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu
 485 490 495

Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser Arg Ile Arg Phe Met Leu
 500 505 510

Gln Asp Val Leu Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg Gly
 515 520 525

Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu Met
 530 535 540

Glu Glu His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys Gly
 545 550 555 560

Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg Gly
 565 570 575

Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser Lys
 580 585 590

Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr Lys Pro
 595 600 605

Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly Gly Arg Leu
 610 615 620

Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala Lys Pro Ser Asp
 625 630 635 640

Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser Thr Leu Asn Arg Phe
645 650 655

Ser Ala Leu Gln Gln Ala
660

<210> 71
<211> 1986
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 71
atggccagca acttcacca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcgggt 360
ggaagcggcg agcagaagta cgagtacaag agcgaccagt ggaagcctct gaacctggag 420
gagaagaaga gatacgacag agagttcctg ctgggcttcc agttcatatt cgcgagcatg 480
cagaagcctg aaggtctgcc tcacatcagc gacgtggtgc tggacaaggc caacaagacc 540
cctcttagac ctctggacct taccagactg cagggcatac actgcggccc tgacttcact 600
ccttcttttcg caaacctggg cagaaccacc ctgagcacca gaggccctcc tagaggtggt 660
cccgcgaggag aactccccag gggctcctgcc ggcctgggcc ctagacgctc ccagcaaggt 720
cctagaaagg agccaaggaa gatcatcgcc accgtgctga tgaccgagga catcaagctg 780
aaciaaagctg agaaggcctg gaagcctagc agcaagagaa ccgccgccga caaggacaga 840
ggcgaggagg acgccgacgg atccaagacc caggacctgt tcagaagagt gagaagcatc 900
ctcaacaagc tgaccctca gatgttccag cagctgatga agcaggtgac gcagctcgcc 960
atcgacaccg aggagagact gaaggcgctg atcgacctga tctttgagaa ggctatctca 1020

gagcctaact tcagcgtggc ctacgccaac atgtgccgtt gcctgatggc attgaaggtg	1080
ccaaccaccg agaagcctac tgtgaccgtc aatttccgta aactgctgct gaaccggtgc	1140
cagaaagagt tcgagaagga taaggacgac gacgaggtct tcgagaagaa acagaaagaa	1200
atggacgagg ccgccaccgc agaggaaagg ggccgattaa aggaggagct ggaggaggcc	1260
agagacatcg ccagacggcg ttctctgggc aacatcaagt tcataggtga gctgttcaag	1320
ctaaagatgc tcaccgaggc cataatgcac gactgcgtgg tgaagctact gaagaaccac	1380
gacgaagaaa gcctggagtg cctgtgcaga ctgctgacca ccatcggcaa ggacctggac	1440
ttcgagaagg caaagcctcg aatggaccag tacttcaacc agatggagaa gattatcaag	1500
gagaagaaga ccagcagcag aatcagattc atgctgcagg acgtactgga cctgcgcgga	1560
agcaactggg tgccaaggag aggggaccaa ggaccaaaga ccatcgacca gatccacaag	1620
gaagcggaga tggaggagca cagagagcac ataaaggtgc agcagcttat ggccaagggc	1680
agcgacaagc gaagaggcgg cccgcccgga cctcctatca gcagaggcct tcctctggtg	1740
gacgacggcg gctggaacac cgtgcctatc tctaagggtc ccagacctat cgacacttcc	1800
cgtcttacca agatcaccaa gccaggatct attgacagca acaaccagct gttcgcccc	1860
ggaggaagac ttagctgggg caagggaagt tccggcggat ccggcgccaa gccttccgac	1920
gccgccagcg aggctgccag acctgccacc agcaccttga accgcttttc cgctctgcag	1980
caagct	1986

<210> 72

<211> 1706

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 72

Met	Ala	Ser	Asn	Phe	Thr	Gln	Phe	Val	Leu	Val	Asp	Asn	Gly	Gly	Thr
1				5					10					15	

Gly	Asp	Val	Thr	Val	Ala	Pro	Ser	Asn	Phe	Ala	Asn	Gly	Ile	Ala	Glu
			20					25				30			

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Asn Ser Gln Pro Gln Thr
115 120 125

Arg Ser Pro Phe Phe Gln Arg Pro Gln Ile Gln Pro Pro Arg Ala Thr
130 135 140

Ile Pro Asn Ser Ser Pro Ser Ile Arg Pro Gly Ala Gln Thr Pro Thr
145 150 155 160

Ala Val Tyr Gln Ala Asn Gln His Ile Met Met Val Asn His Leu Pro
165 170 175

Met Pro Tyr Pro Val Pro Gln Gly Pro Gln Tyr Cys Ile Pro Gln Tyr
180 185 190

Arg His Ser Gly Pro Pro Tyr Val Gly Pro Pro Gln Gln Tyr Pro Val
195 200 205

Gln Pro Pro Gly Pro Gly Pro Phe Tyr Pro Gly Pro Gly Pro Gly Asp
210 215 220

Phe Pro Asn Ala Tyr Gly Thr Pro Phe Tyr Pro Ser Gln Pro Val Tyr
225 230 235 240

Gln Ser Ala Pro Ile Ile Val Pro Thr Gln Gln Gln Pro Pro Pro Ala
245 250 255

Lys Arg Glu Lys Lys Thr Ile Arg Ile Arg Asp Pro Asn Gln Gly Gly
260 265 270

Lys Asp Ile Thr Glu Glu Ile Met Ser Gly Gly Gly Ser Arg Asn Pro
275 280 285

Thr Pro Pro Ile Gly Arg Pro Thr Ser Thr Pro Thr Pro Pro Gln Gln
290 295 300

Leu Pro Ser Gln Val Pro Glu His Ser Pro Val Val Tyr Gly Thr Val
305 310 315 320

Glu Ser Ala His Leu Ala Ala Ser Thr Pro Val Thr Ala Ala Ser Asp
325 330 335

Gln Lys Gln Glu Glu Lys Pro Lys Pro Asp Pro Val Leu Lys Ser Pro
340 345 350

Ser Pro Val Leu Arg Leu Val Leu Ser Gly Glu Lys Lys Glu Gln Glu
355 360 365

Gly Gln Thr Ser Glu Thr Thr Ala Ile Val Ser Ile Ala Glu Leu Pro
370 375 380

Leu Pro Pro Ser Pro Thr Thr Val Ser Ser Val Ala Arg Ser Thr Ile
385 390 395 400

Ala Ala Pro Thr Ser Ser Ala Leu Ser Ser Gln Pro Ile Phe Thr Thr
405 410 415

Ala Ile Asp Asp Arg Cys Glu Leu Ser Ser Pro Arg Glu Asp Thr Ile
420 425 430

Pro Ile Pro Ser Leu Thr Ser Cys Thr Glu Thr Ser Asp Pro Leu Pro
435 440 445

Thr Asn Glu Asn Asp Asp Asp Ile Cys Lys Lys Pro Cys Ser Val Ala
450 455 460

Pro Asn Asp Ile Pro Leu Val Ser Ser Thr Asn Leu Ile Asn Glu Ile
465 470 475 480

Asn Gly Val Ser Glu Lys Leu Ser Ala Thr Glu Ser Ile Val Glu Ile
485 490 495

Val Lys Gln Glu Val Leu Pro Leu Thr Leu Glu Leu Glu Ile Leu Glu
500 505 510

Asn Pro Pro Glu Glu Met Lys Leu Glu Cys Ile Pro Ala Pro Ile Thr
515 520 525

Pro Ser Thr Val Pro Ser Phe Pro Pro Thr Pro Pro Thr Pro Pro Ala
530 535 540

Ser Pro Pro His Thr Pro Val Ile Val Pro Ala Ala Ala Thr Thr Val
545 550 555 560

Ser Ser Pro Ser Ala Ala Ile Thr Val Gln Arg Val Leu Glu Glu Asp
565 570 575

Glu Ser Ile Arg Thr Cys Leu Ser Glu Asp Ala Lys Glu Ile Gln Asn
580 585 590

Lys Ile Glu Val Glu Ala Asp Gly Gln Thr Glu Glu Ile Leu Asp Ser
595 600 605

Gln Asn Leu Asn Ser Arg Arg Ser Pro Val Pro Ala Gln Ile Ala Ile
610 615 620

Thr Val Pro Lys Thr Trp Lys Lys Pro Lys Asp Arg Thr Arg Thr Thr
625 630 635 640

Glu Glu Met Leu Glu Ala Glu Leu Glu Leu Lys Ala Glu Glu Glu Leu
645 650 655

Ser Ile Asp Lys Val Leu Glu Ser Glu Gln Asp Lys Met Ser Gln Gly
660 665 670

Phe His Pro Glu Arg Asp Pro Ser Asp Leu Lys Lys Val Lys Ala Val
675 680 685

Glu Glu Asn Gly Glu Glu Ala Glu Pro Val Arg Asn Gly Ala Glu Ser
690 695 700

Val Ser Glu Gly Glu Gly Ile Asp Ala Asn Ser Gly Ser Thr Asp Ser
705 710 715 720

Ser Gly Asp Gly Val Thr Phe Pro Phe Lys Pro Glu Ser Trp Lys Pro
725 730 735

Thr Asp Thr Glu Gly Lys Lys Gln Tyr Asp Arg Glu Phe Leu Leu Asp
740 745 750

Phe Gln Phe Met Pro Ala Cys Ile Gln Lys Pro Glu Gly Leu Pro Pro
755 760 765

Ile Ser Asp Val Val Leu Asp Lys Ile Asn Gln Pro Lys Leu Pro Met
770 775 780

Arg Thr Leu Asp Pro Arg Ile Leu Pro Arg Gly Pro Asp Phe Thr Pro
785 790 795 800

Ala Phe Ala Asp Phe Gly Arg Gln Thr Pro Gly Gly Arg Gly Val Pro
805 810 815

Leu Leu Asn Val Gly Ser Arg Arg Ser Gln Pro Gly Gln Arg Arg Glu
820 825 830

Pro Arg Lys Ile Ile Thr Val Ser Val Lys Glu Asp Val His Leu Lys
835 840 845

Lys Ala Glu Asn Ala Trp Lys Pro Ser Gln Lys Arg Asp Ser Gln Ala
850 855 860

Asp Asp Pro Glu Asn Ile Lys Thr Gln Glu Leu Phe Arg Lys Val Arg
865 870 875 880

Ser Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Asn Gln Leu Met Lys
885 890 895

Gln Val Ser Gly Leu Thr Val Asp Thr Glu Glu Arg Leu Lys Gly Val
900 905 910

Ile Asp Leu Val Phe Glu Lys Ala Ile Asp Glu Pro Ser Phe Ser Val
915 920 925

Ala Tyr Ala Asn Met Cys Arg Cys Leu Val Thr Leu Lys Val Pro Met
930 935 940

Ala Asp Lys Pro Gly Asn Thr Val Asn Phe Arg Lys Leu Leu Leu Asn
945 950 955 960

Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Ala Asp Asp Asp Val Phe
965 970 975

Glu Lys Lys Gln Lys Glu Leu Glu Ala Ala Ser Ala Pro Glu Glu Arg
980 985 990

Thr Arg Leu His Asp Glu Leu Glu Glu Ala Lys Asp Lys Ala Arg Arg
995 1000 1005

Arg Ser Ile Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu
1010 1015 1020

Lys Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu
1025 1030 1035

Leu Lys Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu
1040 1045 1050

Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro
1055 1060 1065

Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys Ile Val Lys Glu
1070 1075 1080

Arg Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln Asp Val Ile
1085 1090 1095

Asp Leu Arg Leu Cys Asn Trp Val Ser Arg Arg Ala Asp Gln Gly
1100 1105 1110

Pro	Lys	Thr	Ile	Glu	Gln	Ile	His	Lys	Glu	Ala	Lys	Ile	Glu	Glu
	1115					1120					1125			
Gln	Glu	Glu	Gln	Arg	Lys	Val	Gln	Gln	Leu	Met	Thr	Lys	Glu	Lys
	1130					1135					1140			
Arg	Arg	Pro	Gly	Val	Gln	Arg	Val	Asp	Glu	Gly	Gly	Trp	Asn	Thr
	1145					1150					1155			
Val	Gln	Gly	Ala	Lys	Asn	Ser	Arg	Val	Leu	Asp	Pro	Ser	Lys	Phe
	1160					1165					1170			
Leu	Lys	Ile	Thr	Lys	Pro	Thr	Ile	Asp	Glu	Lys	Ile	Gln	Leu	Val
	1175					1180					1185			
Pro	Lys	Ala	Gln	Leu	Gly	Ser	Trp	Gly	Lys	Gly	Ser	Ser	Gly	Gly
	1190					1195					1200			
Ala	Lys	Ala	Ser	Glu	Thr	Asp	Ala	Leu	Arg	Ser	Ser	Ala	Ser	Ser
	1205					1210					1215			
Leu	Asn	Arg	Phe	Ser	Ala	Leu	Gln	Pro	Pro	Ala	Pro	Ser	Gly	Ser
	1220					1225					1230			
Thr	Pro	Ser	Thr	Pro	Val	Glu	Phe	Asp	Ser	Arg	Arg	Thr	Leu	Thr
	1235					1240					1245			
Ser	Arg	Gly	Ser	Met	Gly	Arg	Glu	Lys	Asn	Asp	Lys	Pro	Leu	Pro
	1250					1255					1260			
Ser	Ala	Thr	Ala	Arg	Pro	Asn	Thr	Phe	Met	Arg	Gly	Gly	Ser	Ser
	1265					1270					1275			
Lys	Asp	Leu	Leu	Asp	Asn	Gln	Ser	Gln	Glu	Glu	Gln	Arg	Arg	Glu
	1280					1285					1290			
Met	Leu	Glu	Thr	Val	Lys	Gln	Leu	Thr	Gly	Gly	Val	Asp	Val	Glu
	1295					1300					1305			

Arg	Asn	Ser	Thr	Glu	Ala	Glu	Arg	Asn	Lys	Thr	Arg	Glu	Ser	Ala
1310						1315					1320			
Lys	Pro	Glu	Ile	Ser	Ala	Met	Ser	Ala	His	Asp	Lys	Ala	Ala	Leu
1325						1330					1335			
Ser	Glu	Glu	Glu	Leu	Glu	Arg	Lys	Ser	Lys	Ser	Ile	Ile	Asp	Glu
1340						1345					1350			
Phe	Leu	His	Ile	Asn	Asp	Phe	Lys	Glu	Ala	Met	Gln	Cys	Val	Glu
1355						1360					1365			
Glu	Leu	Asn	Ala	Gln	Gly	Leu	Leu	His	Val	Phe	Val	Arg	Val	Gly
1370						1375					1380			
Val	Glu	Ser	Thr	Leu	Glu	Arg	Ser	Gln	Ile	Thr	Arg	Asp	His	Met
1385						1390					1395			
Gly	Gln	Leu	Leu	Tyr	Gln	Leu	Val	Gln	Ser	Glu	Lys	Leu	Ser	Lys
1400						1405					1410			
Gln	Asp	Phe	Phe	Lys	Gly	Phe	Ser	Glu	Thr	Leu	Glu	Leu	Ala	Asp
1415						1420					1425			
Asp	Met	Ala	Ile	Asp	Ile	Pro	His	Ile	Trp	Leu	Tyr	Leu	Ala	Glu
1430						1435					1440			
Leu	Val	Thr	Pro	Met	Leu	Lys	Glu	Gly	Gly	Ile	Ser	Met	Arg	Glu
1445						1450					1455			
Leu	Thr	Ile	Glu	Phe	Ser	Lys	Pro	Leu	Leu	Pro	Val	Gly	Arg	Ala
1460						1465					1470			
Gly	Val	Leu	Leu	Ser	Glu	Ile	Leu	His	Leu	Leu	Cys	Lys	Gln	Met
1475						1480					1485			
Ser	His	Lys	Lys	Val	Gly	Ala	Leu	Trp	Arg	Glu	Ala	Asp	Leu	Ser
1490						1495					1500			
Trp	Lys	Asp	Phe	Leu	Pro	Glu	Gly	Glu	Asp	Val	His	Asn	Phe	Leu
1505						1510					1515			

<210> 73
<211> 5118
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 73
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt 360
ggaagcaaca gccaaaccca gacacggagc cccttcttcc agcggcccca gatccagcca 420
ccacgggcaa ccatcccaa cagcagcccc agcatccgac ccggcgcaca gacaccacc 480
gccgtgtacc aggccaacca gcacatcatg atggtgaacc acctgcccac gccctacccc 540
gtgccacaag gccctcagta ctgcatcca cagtaccggc acagcgggtc tccctacgtg 600
gggccgccac agcagtaccc cgttcagcct cctggaccgg gaccctttta tccgggacct 660
ggacccggag acttcccaa cgcctacggc acacccttct accccagcca gccagtgtac 720
cagagcgcgc ccatcatcgt gccacccag cagcagcctc caccgcga gcgggagaag 780
aagaccatcc ggatccggga cccaaccag ggcggcaagg acatcaccga ggagatcatg 840
agcggaggcg gcagtcggaa cccacaccg cccatcggca gaccaccag caccacaact 900
ccaccacagc agctgcccag ccaggtgccc gagcacagcc ccgtggtgta cggcaccgtg 960
gaaagcgccc acctggccgc tagcaccct gtgaccgccg ccagcgacca gaagcaggag 1020
gagaagccca aaccgaccc cgtgctgaag tctcccagcc ccgtgctgcg tctggtgctg 1080
agcggcgaga agaaggagca ggagggccag accagcgaga caaccgcat cgtgagcatc 1140
gccgagctgc cactgcctcc cagtcccacc accgtgagca gcgtggcccg gagcaccatc 1200
gcagcgcca caagcagcg cctgagcagc cagcccatct tcaccaccgc catcgacgac 1260
cggcgcgagc tgagcagtcc ccgggaggac accatccca tccccagcct gaccagctgc 1320

accgagacca	gcgacccact	gcccaccaac	gagaacgacg	acgacatctg	caagaagccc	1380
tgcagcgtgg	cacccaacga	catcccgtg	gtgagcagca	ccaacctgat	caacgagatc	1440
aacggcgtta	gcgagaaact	gtctgccacc	gagagcatcg	tggagatcgt	caagcaggag	1500
gtgctgcccc	tgaccctgga	gctagagatc	ctggagaacc	ctccagagga	gatgaagctg	1560
gagtgcattc	ccgccccaat	caccccgagc	accgtgcccc	gctttcctcc	aaccccaccc	1620
accccaccag	ccagtccacc	tcacactccc	gtgatcgtgc	ctgccgccgc	taccaccgtg	1680
tctagcccca	gcgccgccat	caccgtccag	cgggtactgg	aggaggacga	gagcatccgg	1740
acctgcctga	gcgaagacgc	caaggagatc	cagaacaaga	tcgaggtgga	ggccgacggc	1800
cagaccgagg	aaatcctgga	cagccagaac	ctgaattccc	ggcgaagccc	tgtgcccgca	1860
cagatcgcca	tcacagtgcc	caagacctgg	aagaagcccc	aggaccggac	ccggaccaca	1920
gaggaaatgc	tggaggccga	gctggagctg	aaggccgagg	aagagctgag	catcgacaag	1980
gtgctggaga	gcgagcagga	caagatgagc	cagggcttcc	accctgagcg	ggaccccagc	2040
gacctgaaga	aggtgaaggc	cgtggaggag	aacggcgagg	aggcagagcc	tgtgcggaac	2100
ggggccgaga	gcgtgagcga	gggcgagggc	atcgacgcca	acagcggcag	cactgacagc	2160
agcggcgacg	gcgtgacctt	ccccttcaag	cccgaagct	ggaagcccac	cgacaccgaa	2220
ggaaagaagc	agtacgaccg	ggagttcctg	ctggacttcc	agttcatgcc	cgcttgcatt	2280
cagaaacccg	aggggctgcc	accaatcagc	gacgtggtgc	tggacaagat	caaccagccc	2340
aagctgcccc	tgcggaccct	ggatcctcgg	atccttccca	gaggccccga	cttactccc	2400
gccttcgccg	acttcggcag	gcagaccccc	ggtggtcggg	gtgtgcccct	gctgaacgtg	2460
ggcagccgtc	ggagccagcc	tggccaacgg	cgggaacccc	ggaagatcat	caccgtgagc	2520
gtgaaggagg	acgtgcatct	gaagaaggcc	gagaacgcct	ggaaaccag	ccagaagcgg	2580
gacagccaag	ccgacgacct	cgagaacatc	aagaccaggg	agctgttccg	gaaggttcgg	2640
agcatcctga	acaagctgac	accccagatg	ttcaaccagc	tgatgaagca	ggtgagcggc	2700
ctgaccgtgg	acaccgagga	gcggctgaag	ggcgtgatcg	acctggtgtt	tgagaaagcc	2760
attgacgagc	ccagcttcag	cgtggcctac	gccaacatgt	gccggtgcct	ggtgaccctg	2820
aaggtgcccc	tggccgacaa	gcccggcaac	accgtgaact	tccggaagct	gctgctgaac	2880
cggtgccaga	aggagtctga	gaaggacaag	gcagacgacg	acgtgttcga	gaagaagcag	2940

aaagagttag	aggccgcaag	cgcccctgag	gagcggaccc	ggctgcacga	cgagctggag	3000
gaggccaaag	acaaggcccc	gcggaggagc	atcggaaca	tcaagttcat	cggcgagctg	3060
ttcaagctga	agatgctgac	cgaggccatc	atgcacgact	gcgtggtgaa	gctgctgaag	3120
aaccacgacg	aggagagcct	ggagtgcctg	tgccggctgc	tgaccaccat	cggcaaggac	3180
ctggacttcg	agaaggccaa	gccccggatg	gaccagtact	tcaaccagat	ggagaagatc	3240
gtgaaggagc	ggaagaccag	cagccggatc	cggttcatgc	tgaggagcgt	tatcgacctg	3300
cggctgtgca	actgggtgag	ccggagggct	gaccagggcc	ccaagaccat	cgagcagatc	3360
cacaaggagg	ctaagatcga	ggagcaggaa	gagcagcgga	aggtgcagca	gctgatgacc	3420
aaagagaagc	ggaggcctgg	cgtgcagcga	gtggacgagg	gcggctggaa	caccgtgcag	3480
ggcgccaaga	acagccgggt	gctggacccc	agcaagttcc	tgaagatcac	caagcccacc	3540
atcgacgaga	agatccagct	ggtgcccga	gcccagctcg	gcagctgggg	caagggcagc	3600
tcaggtggcg	ccaaggccag	cgagaccgac	gccctgagaa	gcagcgccag	cagcctgaac	3660
cggttctctg	ccctgcagcc	tcctgcccct	agcggctcca	ccccaagcac	acccgtggag	3720
ttcgacagcc	ggcggaccct	gaccagccgg	ggctccatgg	gtcgcgagaa	gaacgacaag	3780
cccttaccca	gcgcaaccgc	tcggcccaac	accttcatgc	ggggcggcag	ctccaaggac	3840
ctgctggaca	accagagcca	ggaggagcag	cggcgggaga	tgctggagac	cgtgaagcag	3900
ctgaccggag	gcgtggacgt	ggagcggaac	agcaccgagg	ccgagcgga	caagaccggg	3960
gagagcgcca	agcccagat	cagcgccatg	agcgcccacg	acaaggccgc	cctgagcgag	4020
gaggagctgg	aacggaagag	caagagcatc	atcgacgagt	tcctgcacat	caacgacttc	4080
aaggaggcca	tgagtgctg	ggaggagctg	aacgcccagg	gcctgctgca	cgtgttcgtg	4140
cgggtgggcg	tggagagcac	cctggagcgg	agccagatca	cccgggacca	catgggccag	4200
ctgctgtacc	aactggtgca	gagcgagaag	ctgagcaagc	aggacttctt	caagggcttc	4260
agcgagactc	tggagctggc	cgacgacatg	gccatcgaca	ttccccacat	ctggctgtac	4320
ctggccgagc	ttgtgacccc	aatgctgaag	gagggcggca	tcagcatgcg	ggagctgacc	4380
atcgagttca	gcaagcccct	cttgccagtc	ggcagagcag	gagtgctgct	gagcgagatc	4440
ctgcacctgc	tgtgcaagca	gatgagccac	aagaaggtgg	gcgccctgtg	gagggaggcc	4500
gacctgagct	ggaaggactt	cctgcccag	ggcgaagacg	tgcaaacctt	cttgctggag	4560

cagaagctgg acttcatcga gagcgacagc ccctgcagca gcgaggccct gagcaagaag 4620
 gaactgagcg ccgaggagct gtacaagcgg ctggagaagc tgatcatcga ggacaaggcc 4680
 aacgacgagc agatcttcga ctgggtggag gccaacctgg acgagatcca gatgagcagc 4740
 cccaccttcc tgcggggccct gatgaccgcc gtgtgcaagg ccgccatcat cgccgacagc 4800
 agcaccttcc ggggtggacac cgccgtgatc aagcagcggg tgcccatcct gctgaagtac 4860
 ctggacagcg acaccgagaa ggagctgcag gccctgtacg ccctgcaggc cagcatcgtg 4920
 aagctggacc agcccgccaa cctgctgcgg atgttcttcg actgcctgta cgacgaagag 4980
 gtgatcagcg aggacgcctt ctacaagtgg gagagcagca aggaccccg c gagcagaac 5040
 ggcaagggcg tggccctgaa gagcgtgacc gccttcttca cctggctgcg agaggccgag 5100
 gaggagagcg aggacaac 5118

<210> 74

<211> 540

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 74

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
 1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
 20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
 35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
 50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
 65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met

85

90

95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
 100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Arg Thr Leu Asp Pro Arg
 115 120 125

Ile Leu Pro Arg Gly Pro Asp Phe Thr Pro Ala Phe Ala Asp Phe Gly
 130 135 140

Arg Gln Thr Pro Gly Gly Arg Gly Val Pro Leu Leu Asn Val Gly Ser
 145 150 155 160

Arg Arg Ser Gln Pro Gly Gln Arg Arg Glu Pro Arg Lys Ile Ile Thr
 165 170 175

Val Ser Val Lys Glu Asp Val His Leu Lys Lys Ala Glu Asn Ala Trp
 180 185 190

Lys Pro Ser Gln Lys Arg Asp Ser Gln Ala Asp Asp Pro Glu Asn Ile
 195 200 205

Lys Thr Gln Glu Leu Phe Arg Lys Val Arg Ser Ile Leu Asn Lys Leu
 210 215 220

Thr Pro Gln Met Phe Asn Gln Leu Met Lys Gln Val Ser Gly Leu Thr
 225 230 235 240

Val Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Val Phe Glu
 245 250 255

Lys Ala Ile Asp Glu Pro Ser Phe Ser Val Ala Tyr Ala Asn Met Cys
 260 265 270

Arg Cys Leu Val Thr Leu Lys Val Pro Met Ala Asp Lys Pro Gly Asn
 275 280 285

Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys Gln Lys Glu Phe
 290 295 300

Glu Lys Asp Lys Ala Asp Asp Asp Val Phe Glu Lys Lys Gln Lys Glu
305 310 315 320

Leu Glu Ala Ala Ser Ala Pro Glu Glu Arg Thr Arg Leu His Asp Glu
325 330 335

Leu Glu Glu Ala Lys Asp Lys Ala Arg Arg Arg Ser Ile Gly Asn Ile
340 345 350

Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys Met Leu Thr Glu Ala Ile
355 360 365

Met His Asp Cys Val Val Lys Leu Leu Lys Asn His Asp Glu Glu Ser
370 375 380

Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr Ile Gly Lys Asp Leu Asp
385 390 395 400

Phe Glu Lys Ala Lys Pro Arg Met Asp Gln Tyr Phe Asn Gln Met Glu
405 410 415

Lys Ile Val Lys Glu Arg Lys Thr Ser Ser Arg Ile Arg Phe Met Leu
420 425 430

Gln Asp Val Ile Asp Leu Arg Leu Cys Asn Trp Val Ser Arg Arg Ala
435 440 445

Asp Gln Gly Pro Lys Thr Ile Glu Gln Ile His Lys Glu Ala Lys Ile
450 455 460

Glu Glu Gln Glu Glu Gln Arg Lys Val Gln Gln Leu Met Thr Lys Glu
465 470 475 480

Lys Arg Arg Pro Gly Val Gln Arg Val Asp Glu Gly Gly Trp Asn Thr
485 490 495

Val Gln Gly Ala Lys Asn Ser Arg Val Leu Asp Pro Ser Lys Phe Leu
500 505 510

Lys Ile Thr Lys Pro Thr Ile Asp Glu Lys Ile Gln Leu Val Pro Lys

515

520

525

Ala Gln Leu Gly Ser Trp Gly Lys Gly Ser Ser Gly
530 535 540

<210> 75

<211> 1620

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 75

atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc	60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc	120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg cccagaacag aaagtacacc	180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc	240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg	300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcggt	360
ggaagccgga ccctggatcc tcggatcctt cccagaggcc ccgacttcac tcccgccttc	420
gccgacttcg gcaggcagac cccaggtggg cgggggtgtgc ccctgctgaa cgtgggcagc	480
cgtcggagcc agcctggcca acggcgggaa ccccggaaga tcatcaccgt gagcgtgaag	540
gaggacgtgc atctgaagaa ggccgagaac gcctggaaac ccagccagaa gcgggacagc	600
caagccgacg accccgagaa catcaagacc caggagctgt tccggaaggt tcggagcatc	660
ctgaacaagc tgacacccca gatgttcaac cagctgatga agcaggtgag cggcctgacc	720
gtggacaccg aggagcggct gaagggcgtg atcgacctgg tgtttgagaa agccattgac	780
gagcccagct tcagcgtggc ctacgccaac atgtgccggt gcctggtgac cctgaaggtg	840
cccatggccg acaagcccgg caacaccgtg aacttccgga agctgctgct gaaccggtgc	900
cagaaggagt tcgagaagga caaggcagac gacgacgtgt tcgagaagaa gcagaaagag	960
ttagaggccg caagcgcccc tgaggagcgg acccggtgc acgacgagct ggaggaggcc	1020
aaagacaagg cccggcggag gagcatcggc aacatcaagt tcatcggcga gctgttcaag	1080

ctgaagatgc tgaccgaggc catcatgcac gactgcgtgg tgaagctgct gaagaaccac 1140
 gacgaggaga gcctggagtg cctgtgccgg ctgctgacca ccatcgga gacactggac 1200
 ttcgagaagg ccaagccccg gatggaccag tacttcaacc agatggagaa gatcgtgaag 1260
 gagcgggaaga ccagcagccg gatccgggtt atgctgcagg acgttatcga cctgcggctg 1320
 tgcaactggg tgagccggag ggctgaccag ggccccaaaga ccatcgagca gatccacaag 1380
 gaggctaaga tcgaggagca ggaagagcag cggaaggtgc agcagctgat gaccaaagag 1440
 aagcggaggc ctggcgtgca gcgagtggac gagggcggct ggaacaccgt gcaggcgcc 1500
 aagaacagcc gggtgctgga cccagcaag ttcctgaaga tcaccaagcc caccatcgac 1560
 gagaagatcc agctggtgcc caaggcccag ctcggcagct ggggcaaggg cagctcaggt 1620

<210> 76

<211> 483

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 76

Phe	Pro	Asn	Ser	Ile	Leu	Gly	Arg	Pro	Pro	Phe	Thr	Pro	Asn	His	Gln
1				5				10						15	

Gln	His	Asn	Asn	Phe	Phe	Thr	Leu	Ser	Pro	Thr	Val	Tyr	Ser	His	Gln
			20					25					30		

Gln	Leu	Ile	Asp	Ala	Gln	Phe	Asn	Phe	Gln	Asn	Ala	Asp	Leu	Ser	Arg
		35					40					45			

Ala	Val	Ser	Leu	Gln	Gln	Leu	Thr	Tyr	Gly	Asn	Val	Ser	Pro	Ile	Gln
	50					55					60				

Thr	Ser	Ala	Ser	Pro	Leu	Phe	Arg	Gly	Arg	Lys	Arg	Leu	Ser	Asp	Glu
65					70					75				80	

Lys	Asn	Leu	Pro	Leu	Asp	Gly	Lys	Arg	Gln	Arg	Phe	His	Ser	Pro	His
				85					90					95	

Gln Glu Pro Thr Val Val Asn Gln Ile Val Pro Leu Ser Gly Glu Arg
100 105 110

Arg Tyr Ser Met Pro Pro Leu Phe His Thr His Tyr Val Pro Asp Ile
115 120 125

Val Arg Cys Val Pro Pro Phe Arg Glu Ile Ala Phe Leu Glu Pro Arg
130 135 140

Glu Ile Thr Leu Pro Glu Ala Lys Asp Lys Leu Ser Gln Gln Ile Leu
145 150 155 160

Glu Leu Phe Glu Thr Cys Gln Gln Gln Ile Ser Asp Leu Lys Lys Lys
165 170 175

Glu Leu Cys Arg Thr Gln Leu Gln Arg Glu Ile Gln Leu Leu Phe Pro
180 185 190

Gln Ser Arg Leu Phe Leu Val Gly Ser Ser Leu Asn Gly Phe Gly Thr
195 200 205

Arg Ser Ser Asp Gly Asp Leu Cys Leu Val Val Lys Glu Glu Pro Cys
210 215 220

Phe Phe Gln Val Asn Gln Lys Thr Glu Ala Arg His Ile Leu Thr Leu
225 230 235 240

Val His Lys His Phe Cys Thr Arg Leu Ser Gly Tyr Ile Glu Arg Pro
245 250 255

Gln Leu Ile Arg Ala Lys Val Pro Ile Val Lys Phe Arg Asp Lys Val
260 265 270

Ser Cys Val Glu Phe Asp Leu Asn Val Asn Asn Ile Val Gly Ile Arg
275 280 285

Asn Thr Phe Leu Leu Arg Thr Tyr Ala Tyr Leu Glu Asn Arg Val Arg
290 295 300

Pro Leu Val Leu Val Ile Lys Lys Trp Ala Ser His His Gln Ile Asn
305 310 315 320

Asp Ala Ser Arg Gly Thr Leu Ser Ser Tyr Ser Leu Val Leu Met Val
325 330 335

Leu His Tyr Leu Gln Thr Leu Pro Glu Pro Ile Leu Pro Ser Leu Gln
340 345 350

Lys Ile Tyr Pro Glu Ser Phe Ser Pro Ala Ile Gln Leu His Leu Val
355 360 365

His Gln Ala Pro Cys Asn Val Pro Pro Tyr Leu Ser Lys Asn Glu Ser
370 375 380

Asn Leu Gly Asp Leu Leu Leu Gly Phe Leu Lys Tyr Tyr Ala Thr Glu
385 390 395 400

Phe Asp Trp Asn Ser Gln Met Ile Ser Val Arg Glu Ala Lys Ala Ile
405 410 415

Pro Arg Pro Asp Gly Ile Glu Trp Arg Asn Lys Tyr Ile Cys Val Glu
420 425 430

Glu Pro Phe Asp Gly Thr Asn Thr Ala Arg Ala Val His Glu Lys Gln
435 440 445

Lys Phe Asp Met Ile Lys Asp Gln Phe Leu Lys Ser Trp His Arg Leu
450 455 460

Lys Asn Lys Arg Asp Leu Asn Ser Ile Leu Pro Val Arg Ala Ala Val
465 470 475 480

Leu Lys Arg

<210> 77

<211> 1449

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 77

ttccctaaca gcatcctggg cagacctcct ttcaccccta accaccagca gcacaacaac	60
ttcttcaccc tgagccctac cgtgtacagc catcagcagc tgatcgacgc ccagttcaac	120
ttccagaacg ccgacctgag cagagccgtg agcctgcagc aactcaccta cggcaacgtt	180
agcccaatcc agaccagcgc cagccctctg ttcagaggca gaaagagact gagcgacgag	240
aagaacctgc ctctggacgg caagagacag agattccaca gccctcacca ggagccaaca	300
gtggtgaacc agatcgtgcc tctgagcggc gagagaagat acagcatgcc tccactgttc	360
cacaccact acgtgcctga catcgtgagg tgcgtgccac ctttcagaga gatcgccttc	420
ctggagcctc gcgagattac cctgcctgag gccaaaggaca agctgagcca gcagattctg	480
gaactgttcg agacttgcca acaacaaatt agcgacctga agaagaagga gctgtgcaga	540
accagctgc agcgcgagat ccagctgctg ttccctcaga gcagactgtt cctggtgggc	600
agcagcctga acggcttcgg caccagaagc agcgacggcg acctgtgcct ggtggtgaag	660
gaggagcctt gcttcttcca agtaaatacag aagaccgagg ccagacacat cctgaccctg	720
gtgcacaagc acttctgcac cagactgtcc ggctacatcg agcgccctca gcttataaga	780
gccaaaggtgc ctatcgtgaa gttcagagac aaggtgagct gcgtggagtt cgatctgaac	840
gtgaacaaca tcgtgggcat cagaaacacc ttctgtctga gaacctacgc ctacctggag	900
aacagagtca gacctctggt gctggtgatc aagaagtggg cttctcatca ccagatcaac	960
gacgcctccc gcggtactct gtctagctac agcctcgtac tgatggtgct gcactacctg	1020
cagactctcc ctgaacctat cctgcctagt ctgcagaaga tctaccctga gagcttcagc	1080
cctgccatcc aattgcatct cgttcaccag gcccttgca acgttcctcc ttatttgagc	1140
aagaacgaga gcaacctggg agatctcctg ctgggcttcc tgaagtacta cgccaccgaa	1200
ttcgattgga acagccagat gatcagcgtg agagaagcaa aggccatccc tagacctgac	1260
ggcatcgagt ggagaaacaa gtacatctgt gtggaggaac ctttcgacgg caccaacacc	1320
gcccgcgcag tacacgagaa gcagaagttc gacatgatca aggaccagtt cctcaagagc	1380
tggcacagac tgaagaacaa gagagacttg aatagcatct tgccggtgag agccgccgtg	1440
ctgaagaga	1449

<210> 78
<211> 605
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 78
Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr
1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu
20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser
35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60

Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80

Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95

Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110

Asn Ser Gly Ile Tyr Gly Gly Gly Gly Ser Phe Pro Asn Ser Ile Leu
115 120 125

Gly Arg Pro Pro Phe Thr Pro Asn His Gln Gln His Asn Asn Phe Phe
130 135 140

Thr Leu Ser Pro Thr Val Tyr Ser His Gln Gln Leu Ile Asp Ala Gln
145 150 155 160

Phe Asn Phe Gln Asn Ala Asp Leu Ser Arg Ala Val Ser Leu Gln Gln
165 170 175

Leu Thr Tyr Gly Asn Val Ser Pro Ile Gln Thr Ser Ala Ser Pro Leu
180 185 190

Phe Arg Gly Arg Lys Arg Leu Ser Asp Glu Lys Asn Leu Pro Leu Asp
195 200 205

Gly Lys Arg Gln Arg Phe His Ser Pro His Gln Glu Pro Thr Val Val
210 215 220

Asn Gln Ile Val Pro Leu Ser Gly Glu Arg Arg Tyr Ser Met Pro Pro
225 230 235 240

Leu Phe His Thr His Tyr Val Pro Asp Ile Val Arg Cys Val Pro Pro
245 250 255

Phe Arg Glu Ile Ala Phe Leu Glu Pro Arg Glu Ile Thr Leu Pro Glu
260 265 270

Ala Lys Asp Lys Leu Ser Gln Gln Ile Leu Glu Leu Phe Glu Thr Cys
275 280 285

Gln Gln Gln Ile Ser Asp Leu Lys Lys Lys Glu Leu Cys Arg Thr Gln
290 295 300

Leu Gln Arg Glu Ile Gln Leu Leu Phe Pro Gln Ser Arg Leu Phe Leu
305 310 315 320

Val Gly Ser Ser Leu Asn Gly Phe Gly Thr Arg Ser Ser Asp Gly Asp
325 330 335

Leu Cys Leu Val Val Lys Glu Glu Pro Cys Phe Phe Gln Val Asn Gln
340 345 350

Lys Thr Glu Ala Arg His Ile Leu Thr Leu Val His Lys His Phe Cys
355 360 365

Thr Arg Leu Ser Gly Tyr Ile Glu Arg Pro Gln Leu Ile Arg Ala Lys
370 375 380

Val Pro Ile Val Lys Phe Arg Asp Lys Val Ser Cys Val Glu Phe Asp

385		390		395		400									
Leu	Asn	Val	Asn	Asn	Ile	Val	Gly	Ile	Arg	Asn	Thr	Phe	Leu	Leu	Arg
			405						410					415	
Thr	Tyr	Ala	Tyr	Leu	Glu	Asn	Arg	Val	Arg	Pro	Leu	Val	Leu	Val	Ile
			420					425					430		
Lys	Lys	Trp	Ala	Ser	His	His	Gln	Ile	Asn	Asp	Ala	Ser	Arg	Gly	Thr
		435					440					445			
Leu	Ser	Ser	Tyr	Ser	Leu	Val	Leu	Met	Val	Leu	His	Tyr	Leu	Gln	Thr
	450					455					460				
Leu	Pro	Glu	Pro	Ile	Leu	Pro	Ser	Leu	Gln	Lys	Ile	Tyr	Pro	Glu	Ser
465					470					475					480
Phe	Ser	Pro	Ala	Ile	Gln	Leu	His	Leu	Val	His	Gln	Ala	Pro	Cys	Asn
			485						490					495	
Val	Pro	Pro	Tyr	Leu	Ser	Lys	Asn	Glu	Ser	Asn	Leu	Gly	Asp	Leu	Leu
			500					505					510		
Leu	Gly	Phe	Leu	Lys	Tyr	Tyr	Ala	Thr	Glu	Phe	Asp	Trp	Asn	Ser	Gln
	515						520					525			
Met	Ile	Ser	Val	Arg	Glu	Ala	Lys	Ala	Ile	Pro	Arg	Pro	Asp	Gly	Ile
	530					535					540				
Glu	Trp	Arg	Asn	Lys	Tyr	Ile	Cys	Val	Glu	Glu	Pro	Phe	Asp	Gly	Thr
545				550					555						560
Asn	Thr	Ala	Arg	Ala	Val	His	Glu	Lys	Gln	Lys	Phe	Asp	Met	Ile	Lys
			565						570					575	
Asp	Gln	Phe	Leu	Lys	Ser	Trp	His	Arg	Leu	Lys	Asn	Lys	Arg	Asp	Leu
			580					585					590		
Asn	Ser	Ile	Leu	Pro	Val	Arg	Ala	Ala	Val	Leu	Lys	Arg			
		595					600					605			

<210> 79
<211> 1815
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 79
atggccagca acttcaccca gttcgtgctg gtggacaacg gcggtaccgg agacgtgacc 60
gtggccccctt ctaacttcgc caacggcatc gccgagtgga tcagcagcaa cagcagaagc 120
caggcctaca aggtgacctg cagcgtgaga cagagcagcg ccagaaacag aaagtacacc 180
atcaaggtgg aggtgcctaa gggcgcctgg agaagctacc tgaacatgga gctgaccatc 240
cctatcttcg ccaccaacag cgactgcgag ctgatcgtga aggccatgca gggcctgctg 300
aaggacggca accctatccc tagcgccatc gccgccaatt caggcatcta cggaggcgggt 360
ggaagcttcc ctaacagcat cctgggcaga cctcctttca cccctaacca ccagcagcac 420
aacaacttct tcaccctgag ccctaccgtg tacagccatc agcagctgat cgacgccag 480
ttcaacttcc agaacgccga cctgagcaga gccgtgagcc tgcagcaact cacctacggc 540
aacgttagcc caatccagac cagcgccagc cctctgttca gaggcagaaa gagactgagc 600
gacgagaaga acctgcctct ggacggcaag agacagagat tccacagccc tcaccaggag 660
ccaacagtgg tgaaccagat cgtgcctctg agcggcgaga gaagatacag catgcctcca 720
ctgttccaca cccactacgt gcctgacatc gtgaggtgcg tgccaccttt cagagagatc 780
gccttcctgg agcctcgca gattaccctg cctgaggcca aggacaagct gagccagcag 840
attctggaac tgttcgagac ttgccaacaa caaattagcg acctgaagaa gaaggagctg 900
tgcagaaccc agctgcagcg cgagatccag ctgctgttcc ctgagagcag actgttcctg 960
gtgggcagca gcctgaacgg cttcggcacc agaagcagcg acggcgacct gtgcctggtg 1020
gtgaaggagg agccttgctt cttccaagta aatcagaaga ccgaggccag acacatcctg 1080
accctggtgc acaagcactt ctgcaccaga ctgtccggct acatcgagcg ccctcagctt 1140
ataagagcca aggtgcctat cgtgaagttc agagacaagg tgagctgcgt ggagttcgat 1200
ctgaacgtga acaacatcgt gggcatcaga aacaccttcc tgctgagaac ctacgcctac 1260

ctggagaaca gagtcaagacc tctggtgctg gtgatcaaga agtgggcttc tcatcaccag	1320
atcaacgacg cctcccgcgg tactctgtct agctacagcc tcgtactgat ggtgctgcac	1380
tacctgcaga ctctccctga acctatcctg cctagtctgc agaagatcta ccctgagagc	1440
ttcagccctg ccatccaatt gcatctcggt caccaggccc cttgcaacgt tcctccttat	1500
ttgagcaaga acgagagcaa cctgggagat ctctgtctgg gcttcctgaa gtactacgcc	1560
accgaattcg attggaacag ccagatgata agcgtgagag aagcaaaggc catccctaga	1620
cctgacggca tcgagtggag aaacaagtac atctgtgtgg aggaaccttt cgacggcacc	1680
aacaccgccc gcgcagtaca cgagaagcag aagttcgaca tgatcaagga ccagttcctc	1740
aagagctggc acagactgaa gaacaagaga gacttgaata gcatcttgcc ggtgagagcc	1800
gccgtgctga agaga	1815

<210> 80

<211> 193

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 80

Met	Gly	Val	His	Glu	Cys	Pro	Ala	Trp	Leu	Trp	Leu	Leu	Leu	Ser	Leu
1				5				10						15	

Leu	Ser	Leu	Pro	Leu	Gly	Leu	Pro	Val	Leu	Gly	Ala	Pro	Pro	Arg	Leu
			20					25						30	

Ile	Cys	Asp	Ser	Arg	Val	Leu	Glu	Arg	Tyr	Leu	Leu	Glu	Ala	Lys	Glu
		35					40					45			

Ala	Glu	Asn	Ile	Thr	Thr	Gly	Cys	Ala	Glu	His	Cys	Ser	Leu	Asn	Glu
	50					55					60				

Asn	Ile	Thr	Val	Pro	Asp	Thr	Lys	Val	Asn	Phe	Tyr	Ala	Trp	Lys	Arg
65					70					75					80

Met	Glu	Val	Gly	Gln	Gln	Ala	Val	Glu	Val	Trp	Gln	Gly	Leu	Ala	Leu
				85					90					95	

Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser
100 105 110

Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly
115 120 125

Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu
130 135 140

Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile
145 150 155 160

Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu
165 170 175

Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp
180 185 190

Arg

<210> 81

<211> 579

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 81

atgggagtgac acgagtggtc cgcgtgggtg tgggtgctgc tgtcgctctt gagcctccca 60

ctgggactgc ctgtgctggg ggcaccaccc agattgatct gcgactcacg ggtacttgag 120

aggtaccttc ttgaagccaa agaagccgaa aacatcacaa ccgatgcgc cgagcactgc 180

tccctcaatg agaacattac tgtaccgatg acaaagggtc atttctatgc atggaagaga 240

atggaagtag gacagcaggc cgtcgaagtg tggcaggggc tcgcgctttt gtcggaggcg 300

gtgttgctgg gtcaggccct cctcgtcaac tcatcacagc cgtgggagcc cctccaactt 360

catgtcgata aagcgggtgc ggggctccgc agcttgacga cgttgcttcg ggctctgggc 420

gcacaaaagg aggctatttc gccgcctgac gcggcctccg cggcaccctt ccgaacgatc 480
 accgcggaca cgtttaggaa gctttttaga gtgtacagca atttcctccg cggaagctg 540
 aaattgtata ctggtgaagc gtgtaggaca ggggatcgc 579

<210> 82
 <211> 284
 <212> PRT
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polypeptide"

<400> 82
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20 25 30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60

Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65 70 75 80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130 135 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145 150 155 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
165 170 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
180 185 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
195 200 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
210 215 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Arg
225 230 235 240

Ser Arg Asp Ile Ser His Gly Phe Pro Pro Ala Val Ala Ala Gln Asp
245 250 255

Asp Gly Thr Leu Pro Met Ser Cys Ala Gln Glu Ser Gly Met Asp Arg
260 265 270

His Pro Ala Ala Cys Ala Ser Ala Arg Ile Asn Val
275 280

<210> 83

<211> 852

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 83

atggtgtcca aggggtgagga attgtttacc ggggtggtgc ctattctcgt cgaacttgac 60

ggggatgtga atggacacaa gttttcggtg tccggagaag gagagggtga cgccacatac 120

ggaaagctta cactcaaatt catctgtacg acggggaaac tgcccgtacc ctggcctacg 180

ctcgtaacca cgctgactta tggagtgcag tgcttttagca gataccccga ccatatgaag	240
cagcacgact tcttcaagtc ggcgatgccc gaggggtacg tgcaagagag gaccattttc	300
ttcaaagacg atggcaatta caaaacacgc gcagaagtca agtttgaggg cgatactctg	360
gtcaatcgga tcgaattgaa gggaatcgat ttcaaagaag atggaaacat ccttggccat	420
aagctcgagt acaactataa ctcgcataat gtctatatca tggctgacaa gcagaaaaac	480
ggtatcaaag tcaactttaa gatccgacac aatattgagg acggttcggg gcagcttgcg	540
gaccactatc aacagaatac gccgattggg gatgggtccgg tccttttgcc ggataaccat	600
tatctctcaa cccagtcagc cctgagcaaa gatccaaacg agaagaggga ccacatggtc	660
ttgctcgaat tcgtgacagc ggcagggatc actctgggaa tggacgagtt gtacaagaga	720
tctcgagata tcagccatgg cttcccgcgc gcggtggcgg cgcaggatga tggcacgctg	780
cccatgtctt gtgcccagga gagcgggatg gaccgtcacc ctgcagcctg tgcttctgct	840
aggatcaatg tg	852

<210> 84
 <211> 100
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 84	
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	60
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	100

<210> 85
 <211> 550
 <212> PRT
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 85	
Met Glu Asp Ala Lys Asn Ile Lys Lys Gly Pro Ala Pro Phe Tyr Pro	
1 5 10 15	

Leu Glu Asp Gly Thr Ala Gly Glu Gln Leu His Lys Ala Met Lys Arg
20 25 30

Tyr Ala Leu Val Pro Gly Thr Ile Ala Phe Thr Asp Ala His Ile Glu
35 40 45

Val Asp Ile Thr Tyr Ala Glu Tyr Phe Glu Met Ser Val Arg Leu Ala
50 55 60

Glu Ala Met Lys Arg Tyr Gly Leu Asn Thr Asn His Arg Ile Val Val
65 70 75 80

Cys Ser Glu Asn Ser Leu Gln Phe Phe Met Pro Val Leu Gly Ala Leu
85 90 95

Phe Ile Gly Val Ala Val Ala Pro Ala Asn Asp Ile Tyr Asn Glu Arg
100 105 110

Glu Leu Leu Asn Ser Met Gly Ile Ser Gln Pro Thr Val Val Phe Val
115 120 125

Ser Lys Lys Gly Leu Gln Lys Ile Leu Asn Val Gln Lys Lys Leu Pro
130 135 140

Ile Ile Gln Lys Ile Ile Ile Met Asp Ser Lys Thr Asp Tyr Gln Gly
145 150 155 160

Phe Gln Ser Met Tyr Thr Phe Val Thr Ser His Leu Pro Pro Gly Phe
165 170 175

Asn Glu Tyr Asp Phe Val Pro Glu Ser Phe Asp Arg Asp Lys Thr Ile
180 185 190

Ala Leu Ile Met Asn Ser Ser Gly Ser Thr Gly Leu Pro Lys Gly Val
195 200 205

Ala Leu Pro His Arg Thr Ala Cys Val Arg Phe Ser His Ala Arg Asp
210 215 220

Pro Ile Phe Gly Asn Gln Ile Ile Pro Asp Thr Ala Ile Leu Ser Val
225 230 235 240

Val Pro Phe His His Gly Phe Gly Met Phe Thr Thr Leu Gly Tyr Leu
245 250 255

Ile Cys Gly Phe Arg Val Val Leu Met Tyr Arg Phe Glu Glu Glu Leu
260 265 270

Phe Leu Arg Ser Leu Gln Asp Tyr Lys Ile Gln Ser Ala Leu Leu Val
275 280 285

Pro Thr Leu Phe Ser Phe Phe Ala Lys Ser Thr Leu Ile Asp Lys Tyr
290 295 300

Asp Leu Ser Asn Leu His Glu Ile Ala Ser Gly Gly Ala Pro Leu Ser
305 310 315 320

Lys Glu Val Gly Glu Ala Val Ala Lys Arg Phe His Leu Pro Gly Ile
325 330 335

Arg Gln Gly Tyr Gly Leu Thr Glu Thr Thr Ser Ala Ile Leu Ile Thr
340 345 350

Pro Glu Gly Asp Asp Lys Pro Gly Ala Val Gly Lys Val Val Pro Phe
355 360 365

Phe Glu Ala Lys Val Val Asp Leu Asp Thr Gly Lys Thr Leu Gly Val
370 375 380

Asn Gln Arg Gly Glu Leu Cys Val Arg Gly Pro Met Ile Met Ser Gly
385 390 395 400

Tyr Val Asn Asn Pro Glu Ala Thr Asn Ala Leu Ile Asp Lys Asp Gly
405 410 415

Trp Leu His Ser Gly Asp Ile Ala Tyr Trp Asp Glu Asp Glu His Phe
420 425 430

Phe Ile Val Asp Arg Leu Lys Ser Leu Ile Lys Tyr Lys Gly Tyr Gln
435 440 445

Val Ala Pro Ala Glu Leu Glu Ser Ile Leu Leu Gln His Pro Asn Ile
450 455 460

Phe Asp Ala Gly Val Ala Gly Leu Pro Asp Asp Asp Ala Gly Glu Leu
465 470 475 480

Pro Ala Ala Val Val Val Leu Glu His Gly Lys Thr Met Thr Glu Lys
485 490 495

Glu Ile Val Asp Tyr Val Ala Ser Gln Val Thr Thr Ala Lys Lys Leu
500 505 510

Arg Gly Gly Val Val Phe Val Asp Glu Val Pro Lys Gly Leu Thr Gly
515 520 525

Lys Leu Asp Ala Arg Lys Ile Arg Glu Ile Leu Ile Lys Ala Lys Lys
530 535 540

Gly Gly Lys Ile Ala Val
545 550

<210> 86
<211> 1650
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 86
atggaagatg cgaagaacat caagaaggga cctgccccgt tttacccttt ggaggacggt 60
acagcaggag aacagctcca caaggcgatg aaacgctacg ccctggtccc cggaacgatt 120
gcgtttaccg atgcacatat tgaggtagac atcacatacg cagaatactt cgaaatgtcg 180
gtgaggctgg cggaagcgat gaagagatat ggtcttaaca ctaatcaccg catcgtggtg 240
tgttcggaga actcattgca gtttttcatg ccggtccttg gagcactttt catcggggtc 300
gcagtgcgc cagcgaacga catctacaat gagcgggaac tcttgaatag catgggaatc 360
tcccagccga cggtcgtgtt tgtctccaaa aaggggctgc agaaaatcct caacgtgcag 420

aagaagctcc ccattattca aaagatcatc attatggata gcaagacaga ttaccaaggg	480
ttccagtcga tgtatacctt tgtgacatcg cttttgccgc cagggtttta cagatatgac	540
ttcgtccccg agtcatttga cagagataaa accatcgcg cgtattatgaa ttcctcgggt	600
agcaccgggt tgccaaaggg ggtggcggtg cccaccgca ctgcttgtgt gcggttctcg	660
cacgctaggg atcctatctt tggtaatcag atcattcccg acacagcaat cctgtccgtg	720
gtaccttttc atcacggttt tggcatgttc acgactctcg gctatttgat ttgcggtttc	780
agggtcgtac ttatgtatcg gttcgaggaa gaactgtttt tgagatcctt gcaagattac	840
aagatccagt cggccctcct tgtgccaacg cttttctcat tctttgcgaa atcgacactt	900
attgataagt atgacctttc caatctgcat gagattgcct cagggggagc gccgcttagc	960
aaggaagtcg gggaggcagt ggccaagcgc ttccaccttc ccggaattcg gcagggatac	1020
gggctcacgg agacaacatc cgcgatcctt atcacgccg agggtgacga taagccggga	1080
gccgtcggaa aagtgggtccc cttctttgaa gccaaagtcg tagacctcga cacgggaaaa	1140
accctcggag tgaaccagag gggcgagctc tgcgtgagag ggccgatgat catgtcaggt	1200
tacgtgaata accctgaagc gacgaatgcg ctgatcgaca aggatgggtg gttgcattcg	1260
ggagacattg cctattggga tgaggatgag cacttcttta tcgtagatcg acttaagagc	1320
ttgatcaaat acaaaggcta tcaggtagcg cctgccgagc tcgagtcaat cctgctccag	1380
cacccaaca ttttcgacgc cggagtggcc gggttgccc atgacgacgc gggtagctg	1440
ccagcggccg tggtagtcct cgaacatggg aaaacaatga ccgaaaagga gatcgtggac	1500
tacgtagcat cacaagtac gactgcgaag aaactgaggg gaggggtagt ctttgtggac	1560
gaggtcccga aaggcttgac tgggaagctt gacgctcgca aaatccggga aatcctgatt	1620
aaggcaaaga aaggcgggaa aatcgctgtc	1650

<210> 87

<211> 66

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 87

Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
1 5 10 15

Glu Asn Pro Gly Pro Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys
20 25 30

Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro Gly Ser Gly Gln Cys
35 40 45

Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro
50 55 60

Gly Pro
65

<210> 88

<211> 198

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 88

ggctctggcg aaggccgggg tagcctgctg acctgcggcg acgtggagga gaaccctgga 60
cctggatccg gcgccaccaa cttcagcctg ctcaagcagg ccggagacgt ggaagagaat 120
ccaggcccag gcagcgggtca gtgcaccaac tacgccctgc tgaagctggc cggcgacgta 180
gagagcaacc ccggcccc 198

<210> 89

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
peptide"

<220>

<221> VARIANT

<222> (1)..(3)
<223> /replace=" "

<220>

<221> SITE

<222> (1)..(21)

<223> /note="Variant residues given in the sequence have no
preference with respect to those in the annotations
for variant positions"

<400> 89

Gly	Ser	Gly	Glu	Gly	Arg	Gly	Ser	Leu	Leu	Thr	Cys	Gly	Asp	Val	Glu
1				5					10					15	

Glu	Asn	Pro	Gly	Pro
			20	

<210> 90

<211> 22

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
peptide"

<220>

<221> VARIANT

<222> (1)..(3)

<223> /replace=" "

<220>

<221> SITE

<222> (1)..(22)

<223> /note="Variant residues given in the sequence have no
preference with respect to those in the annotations
for variant positions"

<400> 90

Gly	Ser	Gly	Ala	Thr	Asn	Phe	Ser	Leu	Leu	Lys	Gln	Ala	Gly	Asp	Val
1				5					10					15	

Glu	Glu	Asn	Pro	Gly	Pro
				20	

<210> 91

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<220>

<221> VARIANT

<222> (1)..(3)

<223> /replace=" "

<220>

<221> SITE

<222> (1)..(23)

<223> /note="Variant residues given in the sequence have no preference with respect to those in the annotations for variant positions"

<400> 91

Gly Ser Gly Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp
1 5 10 15

Val Glu Ser Asn Pro Gly Pro
20

<210> 92

<211> 25

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic peptide"

<220>

<221> VARIANT

<222> (1)..(3)

<223> /replace=" "

<220>

<221> SITE

<222> (1)..(25)

<223> /note="Variant residues given in the sequence have no preference with respect to those in the annotations for variant positions"

<400> 92

Gly Ser Gly Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala

1 5 10 15

Gly Asp Val Glu Ser Asn Pro Gly Pro
20 25

<210> 93

<211> 1215

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polypeptide"

<400> 93

Met Glu Asp Ala Lys Asn Ile Lys Lys Gly Pro Ala Pro Phe Tyr Pro
1 5 10 15

Leu Glu Asp Gly Thr Ala Gly Glu Gln Leu His Lys Ala Met Lys Arg
20 25 30

Tyr Ala Leu Val Pro Gly Thr Ile Ala Phe Thr Asp Ala His Ile Glu
35 40 45

Val Asp Ile Thr Tyr Ala Glu Tyr Phe Glu Met Ser Val Arg Leu Ala
50 55 60

Glu Ala Met Lys Arg Tyr Gly Leu Asn Thr Asn His Arg Ile Val Val
65 70 75 80

Cys Ser Glu Asn Ser Leu Gln Phe Phe Met Pro Val Leu Gly Ala Leu
85 90 95

Phe Ile Gly Val Ala Val Ala Pro Ala Asn Asp Ile Tyr Asn Glu Arg
100 105 110

Glu Leu Leu Asn Ser Met Gly Ile Ser Gln Pro Thr Val Val Phe Val
115 120 125

Ser Lys Lys Gly Leu Gln Lys Ile Leu Asn Val Gln Lys Lys Leu Pro
130 135 140

Ile Ile Gln Lys Ile Ile Ile Met Asp Ser Lys Thr Asp Tyr Gln Gly
145 150 155 160

Phe Gln Ser Met Tyr Thr Phe Val Thr Ser His Leu Pro Pro Gly Phe
165 170 175

Asn Glu Tyr Asp Phe Val Pro Glu Ser Phe Asp Arg Asp Lys Thr Ile
180 185 190

Ala Leu Ile Met Asn Ser Ser Gly Ser Thr Gly Leu Pro Lys Gly Val
195 200 205

Ala Leu Pro His Arg Thr Ala Cys Val Arg Phe Ser His Ala Arg Asp
210 215 220

Pro Ile Phe Gly Asn Gln Ile Ile Pro Asp Thr Ala Ile Leu Ser Val
225 230 235 240

Val Pro Phe His His Gly Phe Gly Met Phe Thr Thr Leu Gly Tyr Leu
245 250 255

Ile Cys Gly Phe Arg Val Val Leu Met Tyr Arg Phe Glu Glu Glu Leu
260 265 270

Phe Leu Arg Ser Leu Gln Asp Tyr Lys Ile Gln Ser Ala Leu Leu Val
275 280 285

Pro Thr Leu Phe Ser Phe Phe Ala Lys Ser Thr Leu Ile Asp Lys Tyr
290 295 300

Asp Leu Ser Asn Leu His Glu Ile Ala Ser Gly Gly Ala Pro Leu Ser
305 310 315 320

Lys Glu Val Gly Glu Ala Val Ala Lys Arg Phe His Leu Pro Gly Ile
325 330 335

Arg Gln Gly Tyr Gly Leu Thr Glu Thr Thr Ser Ala Ile Leu Ile Thr
340 345 350

Pro Glu Gly Asp Asp Lys Pro Gly Ala Val Gly Lys Val Val Pro Phe
355 360 365

Phe Glu Ala Lys Val Val Asp Leu Asp Thr Gly Lys Thr Leu Gly Val
370 375 380

Asn Gln Arg Gly Glu Leu Cys Val Arg Gly Pro Met Ile Met Ser Gly
385 390 395 400

Tyr Val Asn Asn Pro Glu Ala Thr Asn Ala Leu Ile Asp Lys Asp Gly
405 410 415

Trp Leu His Ser Gly Asp Ile Ala Tyr Trp Asp Glu Asp Glu His Phe
420 425 430

Phe Ile Val Asp Arg Leu Lys Ser Leu Ile Lys Tyr Lys Gly Tyr Gln
435 440 445

Val Ala Pro Ala Glu Leu Glu Ser Ile Leu Leu Gln His Pro Asn Ile
450 455 460

Phe Asp Ala Gly Val Ala Gly Leu Pro Asp Asp Asp Ala Gly Glu Leu
465 470 475 480

Pro Ala Ala Val Val Val Leu Glu His Gly Lys Thr Met Thr Glu Lys
485 490 495

Glu Ile Val Asp Tyr Val Ala Ser Gln Val Thr Thr Ala Lys Lys Leu
500 505 510

Arg Gly Gly Val Val Phe Val Asp Glu Val Pro Lys Gly Leu Thr Gly
515 520 525

Lys Leu Asp Ala Arg Lys Ile Arg Glu Ile Leu Ile Lys Ala Lys Lys
530 535 540

Gly Gly Lys Ile Ala Val Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu
545 550 555 560

Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Gly Ser Gly Ala Thr
565 570 575

Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly
580 585 590

Pro Gly Ser Gly Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly
595 600 605

Asp Val Glu Ser Asn Pro Gly Pro Ala Ser Asn Phe Thr Gln Phe Val
610 615 620

Leu Val Asp Asn Gly Gly Thr Gly Asp Val Thr Val Ala Pro Ser Asn
625 630 635 640

Phe Ala Asn Gly Ile Ala Glu Trp Ile Ser Ser Asn Ser Arg Ser Gln
645 650 655

Ala Tyr Lys Val Thr Cys Ser Val Arg Gln Ser Ser Ala Gln Asn Arg
660 665 670

Lys Tyr Thr Ile Lys Val Glu Val Pro Lys Gly Ala Trp Arg Ser Tyr
675 680 685

Leu Asn Met Glu Leu Thr Ile Pro Ile Phe Ala Thr Asn Ser Asp Cys
690 695 700

Glu Leu Ile Val Lys Ala Met Gln Gly Leu Leu Lys Asp Gly Asn Pro
705 710 715 720

Ile Pro Ser Ala Ile Ala Ala Asn Ser Gly Ile Tyr Gly Gly Gly Gly
725 730 735

Ser Asp Pro Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr
740 745 750

Pro Ser Phe Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro
755 760 765

Pro Arg Gly Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu
770 775 780

Gly Pro Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile
785 790 795 800

Ile Ala Thr Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu
805 810 815

Lys Ala Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg
820 825 830

Gly Glu Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg
835 840 845

Val Arg Ser Ile Ala Asn Lys Ala Thr Pro Gln Met Ala Gln Gln Leu
850 855 860

Met Lys Gln Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys
865 870 875 880

Gly Val Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe
885 890 895

Ser Val Ala Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val
900 905 910

Pro Thr Thr Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu
915 920 925

Leu Asn Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu
930 935 940

Val Phe Glu Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu
945 950 955 960

Glu Arg Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala
965 970 975

Arg Arg Arg Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys
980 985 990

Leu Lys Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu
995 1000 1005

Leu Lys Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu
1010 1015 1020

Leu Thr Thr Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro
1025 1030 1035

Arg Met Asp Gln Tyr Phe Asn Gln Met Glu Lys Ile Ile Lys Glu
1040 1045 1050

Lys Lys Thr Ser Ser Arg Ile Arg Phe Met Leu Gln Asp Val Leu
1055 1060 1065

Asp Leu Arg Gly Ser Asn Trp Val Pro Arg Arg Gly Asp Gln Gly
1070 1075 1080

Pro Lys Thr Ile Asp Gln Ile His Lys Glu Ala Glu Met Glu Glu
1085 1090 1095

His Arg Glu His Ile Lys Val Gln Gln Leu Met Ala Lys Gly Ser
1100 1105 1110

Asp Lys Arg Arg Gly Gly Pro Pro Gly Pro Pro Ile Ser Arg Gly
1115 1120 1125

Leu Pro Leu Val Asp Asp Gly Gly Trp Asn Thr Val Pro Ile Ser
1130 1135 1140

Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu Thr Lys Ile Thr
1145 1150 1155

Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe Ala Pro Gly
1160 1165 1170

Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser Gly Ala
1175 1180 1185

Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr Ser
1190 1195 1200

Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val
1205 1210 1215

<210> 94
<211> 3645
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 94
atggaggacg ccaagaacat caagaagggg cccgctccct tctacccgct ggaggacggc 60
accgctggcg agcagctgca caaggccatg aagaggtacg cactggtgcc cggcacaatc 120
gccttcaccg acgcccacat cgaggtggac atcacctacg ccgagtactt cgagatgagc 180
gtgcggtggtg ccgaagccat gaagcgggtac ggcctgaaca ccaaccaccg catcgtgggtg 240
tgctccgaga actccctgca gttcttcatg cccgttcttg gcgccctgtt cattgggggtg 300
gccgtggcac cagccaacga catctacaac gagagggagc tgctgaacag catgggcatc 360
agccagccca ccgtgggtgtt cgtgtccaag aagggcctgc agaagatcct gaacgtgcag 420
aagaagctgc ccatcatcca gaagatcatc atcatggaca gcaagaccga ctaccagggc 480
ttccagtcca tgtacacctt cgtgaccagc cacctgcctc ccggcttcaa cgagtacgac 540
ttcgtgcccc agtccttcga cagggacaag accatcgccc tgatcatgaa ctccagcgga 600
agcaccggcc tgccaagggg tgtgggtcta cccaccgga ccgcctgtgt acggttcagc 660
cacgcccgtg acccatctt cggaaccag atcatccccg acaccgcat cctgagcgtg 720
gtgcccttcc accacggctt cggcatgttc accaccctgg gctacctgat ctgcggcttc 780
cgggtgggtg tgatgtaccg gttcaggag gagctgttcc tgcggagcct gcaggactac 840
aagatccagt ccgccctgct ggtgcccacc ctgttctcct tcttcgcaa gtccaccctg 900
atcgacaagt acgacctgag caacctgcac gagattgcct ctggcggcgc acctctgtcc 960
aaggaggtcg gggaggccgt agccaagcgg ttccacctgc ccggcatccg gcaaggctac 1020
ggcctgaccg agaccaccag cgccatcctg atcacccctg agggcgacga caagcccggc 1080
gcagtgggga aggtgggtgcc attcttcgag gccaaggtgg tggacctgga caccggcaag 1140
accctgggcg tgaaccagcg gggcgagctg tgcgtgcggg gacctatgat catgtccggc 1200
tacgtgaaca accccgaggc caccaacgcc ctgattgaca aggacgggtg gctgcacagc 1260

ggcgacatcg cctactggga cgaggacgag cacttcttca tcgtggaccg cctgaagtcc	1320
ctgatcaagt acaaggggta ccaggtggct cccgccgagc tggagtccat cctgctgcag	1380
cacccaaca tctttgacgc cggcgttgcc ggactgcctg acgacgacgc tggggaactc	1440
cccgagctg tgggtggtgct ggagcacggc aagaccatga ccgagaagga gatcgtggac	1500
tacgtcgctt cccaggtgac caccgccaag aagttacgcg gcggcgctgt gttcgtggac	1560
gaggtgcca agggcctgac cggcaagctg gacgcccgca agatccggga gatcctgac	1620
aaggccaaga agggcgggaa gatcgccgtg ggctctggcg aaggccgggg tagcctgctg	1680
acctgcggcg acgtggagga gaacctgga cctggatccg gcgccacca cttcagcctg	1740
ctcaagcagg ccggagacgt ggaagagaat ccaggcccag gcagcgggtca gtgcaccaac	1800
tacgccctgc tgaagctggc cggcgacgta gagagcaacc ccggccccgc cagcaacttc	1860
accagttcg tgctggtgga caacggcggg accggagacg tgaccgtggc cccttctaac	1920
ttcgccaacg gcatcgccga gtggatcagc agcaacagca gaagccaggc ctacaaggtg	1980
acctgcagcg tgagacagag cagcgcccag aacagaaagt acaccatcaa ggtggaggtg	2040
cctaagggcg cctggagaag ctacctgaac atggagctga ccatccctat cttcgccacc	2100
aacagcgact gcgagctgat cgtgaaggcc atgcagggcc tgctgaagga cggcaaccct	2160
atccctagcg ccatcgccgc caattcaggc atctacggtg gcggaggcag cgaccctacc	2220
agactgcagg gcatcaactg cggccctgac ttactcctt ctttcgcaaa cctgggcaga	2280
accaccctga gcaccagagg ccctcctaga ggtggtcccg gcggagaact cccaggggt	2340
cctgccggcc tgggccctag acgctcccag caaggtccta gaaaggagcc aaggaagatc	2400
atcgccaccg tgctgatgac cgaggacatc aagctgaaca aagctgagaa ggcctggaag	2460
cctagcagca agagaaccgc cgccgacaag gacagaggcg aggaggacgc cgacggatcc	2520
aagaccaggg acctgttcag aagagtgaga agcatcgcca acaaggcgac ccctcagatg	2580
gcgcagcagc tgatgaagca ggtgacgcag ctcgccatcg acaccgagga gagactgaag	2640
ggcgtgatcg acctgatctt tgagaaggct atctcagagc ctaacttcag cgtggcctac	2700
gccaacatgt gccgttgctt gatggcattg aaggtgccaa ccaccgagaa gcctactgtg	2760
accgtcaatt tccgtaaaact gctgctgaac cggtgccaga aagagttcga gaaggataag	2820
gacgacgacg aggtcttcga gaagaaacag aaagaaatgg acgaggccgc caccgcagag	2880

gaaaggggcc gattaaagga ggagctggag gaggccagag acatcgccag acggcgttct	2940
ctgggcaaca tcaagttcat aggtgagctg ttcaagctaa agatgctcac cgaggccata	3000
atgcacgact gcgtggtgaa gctactgaag aaccacgacg aagaaagcct ggagtgccctg	3060
tgagactgc tgaccacat cggaaggac ctggacttcg agaaggcaaa gcctcgaatg	3120
gaccagtact tcaaccagat ggagaagatt atcaaggaga agaagaccag cagcagaatc	3180
agattcatgc tgcaggacgt actggacctg cgcggaagca actgggtgcc aaggagaggg	3240
gaccaaggac caaagacat cgaccagatc cacaaggaag cggagatgga ggagcacaga	3300
gagcacataa aggtgcagca gcttatggcc aagggcagcg acaagcgaag aggcggccc	3360
cccggacctc ctatcagcag aggccttcct ctggtagacg acggcggctg gaacaccgtg	3420
cctatctcta agggctccag acctatcgac acttcccgtc ttaccaagat caccaagcca	3480
ggatctattg acagcaaaa ccagctgttc gcccaggag gaagacttag ctggggcaag	3540
ggaagtccg gcggatccg cgccaagcct tccgacgcc ccagcgaggc tgccagacct	3600
gccaccagca cttgaaccg cttttccgct ctgcagcaag ctgtg	3645

```

<210> 95
<211> 1783
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
      polypeptide"

```

```

<400> 95
Met Glu Asp Ala Lys Asn Ile Lys Lys Gly Pro Ala Pro Phe Tyr Pro
1             5             10             15

Leu Glu Asp Gly Thr Ala Gly Glu Gln Leu His Lys Ala Met Lys Arg
      20             25             30

Tyr Ala Leu Val Pro Gly Thr Ile Ala Phe Thr Asp Ala His Ile Glu
      35             40             45

Val Asp Ile Thr Tyr Ala Glu Tyr Phe Glu Met Ser Val Arg Leu Ala
      50             55             60

```

Glu Ala Met Lys Arg Tyr Gly Leu Asn Thr Asn His Arg Ile Val Val
65 70 75 80

Cys Ser Glu Asn Ser Leu Gln Phe Phe Met Pro Val Leu Gly Ala Leu
85 90 95

Phe Ile Gly Val Ala Val Ala Pro Ala Asn Asp Ile Tyr Asn Glu Arg
100 105 110

Glu Leu Leu Asn Ser Met Gly Ile Ser Gln Pro Thr Val Val Phe Val
115 120 125

Ser Lys Lys Gly Leu Gln Lys Ile Leu Asn Val Gln Lys Lys Leu Pro
130 135 140

Ile Ile Gln Lys Ile Ile Ile Met Asp Ser Lys Thr Asp Tyr Gln Gly
145 150 155 160

Phe Gln Ser Met Tyr Thr Phe Val Thr Ser His Leu Pro Pro Gly Phe
165 170 175

Asn Glu Tyr Asp Phe Val Pro Glu Ser Phe Asp Arg Asp Lys Thr Ile
180 185 190

Ala Leu Ile Met Asn Ser Ser Gly Ser Thr Gly Leu Pro Lys Gly Val
195 200 205

Ala Leu Pro His Arg Thr Ala Cys Val Arg Phe Ser His Ala Arg Asp
210 215 220

Pro Ile Phe Gly Asn Gln Ile Ile Pro Asp Thr Ala Ile Leu Ser Val
225 230 235 240

Val Pro Phe His His Gly Phe Gly Met Phe Thr Thr Leu Gly Tyr Leu
245 250 255

Ile Cys Gly Phe Arg Val Val Leu Met Tyr Arg Phe Glu Glu Glu Leu
260 265 270

Phe Leu Arg Ser Leu Gln Asp Tyr Lys Ile Gln Ser Ala Leu Leu Val

275

280

285

Pro Thr Leu Phe Ser Phe Phe Ala Lys Ser Thr Leu Ile Asp Lys Tyr
 290 295 300

Asp Leu Ser Asn Leu His Glu Ile Ala Ser Gly Gly Ala Pro Leu Ser
 305 310 315 320

Lys Glu Val Gly Glu Ala Val Ala Lys Arg Phe His Leu Pro Gly Ile
 325 330 335

Arg Gln Gly Tyr Gly Leu Thr Glu Thr Thr Ser Ala Ile Leu Ile Thr
 340 345 350

Pro Glu Gly Asp Asp Lys Pro Gly Ala Val Gly Lys Val Val Pro Phe
 355 360 365

Phe Glu Ala Lys Val Val Asp Leu Asp Thr Gly Lys Thr Leu Gly Val
 370 375 380

Asn Gln Arg Gly Glu Leu Cys Val Arg Gly Pro Met Ile Met Ser Gly
 385 390 395 400

Tyr Val Asn Asn Pro Glu Ala Thr Asn Ala Leu Ile Asp Lys Asp Gly
 405 410 415

Trp Leu His Ser Gly Asp Ile Ala Tyr Trp Asp Glu Asp Glu His Phe
 420 425 430

Phe Ile Val Asp Arg Leu Lys Ser Leu Ile Lys Tyr Lys Gly Tyr Gln
 435 440 445

Val Ala Pro Ala Glu Leu Glu Ser Ile Leu Leu Gln His Pro Asn Ile
 450 455 460

Phe Asp Ala Gly Val Ala Gly Leu Pro Asp Asp Asp Ala Gly Glu Leu
 465 470 475 480

Pro Ala Ala Val Val Val Leu Glu His Gly Lys Thr Met Thr Glu Lys
 485 490 495

Glu Ile Val Asp Tyr Val Ala Ser Gln Val Thr Thr Ala Lys Lys Leu
500 505 510

Arg Gly Gly Val Val Phe Val Asp Glu Val Pro Lys Gly Leu Thr Gly
515 520 525

Lys Leu Asp Ala Arg Lys Ile Arg Glu Ile Leu Ile Lys Ala Lys Lys
530 535 540

Gly Gly Lys Ile Ala Val Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu
545 550 555 560

Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Gly Ser Gly Ala Thr
565 570 575

Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly
580 585 590

Pro Gly Ser Gly Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly
595 600 605

Asp Val Glu Ser Asn Pro Gly Pro Ala Ser Asn Phe Thr Gln Phe Val
610 615 620

Leu Val Asp Asn Gly Gly Thr Gly Asp Val Thr Val Ala Pro Ser Asn
625 630 635 640

Phe Ala Asn Gly Ile Ala Glu Trp Ile Ser Ser Asn Ser Arg Ser Gln
645 650 655

Ala Tyr Lys Val Thr Cys Ser Val Arg Gln Ser Ser Ala Gln Asn Arg
660 665 670

Lys Tyr Thr Ile Lys Val Glu Val Pro Lys Gly Ala Trp Arg Ser Tyr
675 680 685

Leu Asn Met Glu Leu Thr Ile Pro Ile Phe Ala Thr Asn Ser Asp Cys
690 695 700

Glu Leu Ile Val Lys Ala Met Gln Gly Leu Leu Lys Asp Gly Asn Pro

705		710		715		720									
Ile	Pro	Ser	Ala	Ile	Ala	Ala	Asn	Ser	Gly	Ile	Tyr	Gly	Gly	Gly	Gly
			725						730					735	
Ser	Ser	Phe	Thr	Leu	Thr	Asn	Lys	Asn	Val	Ile	Phe	Val	Ala	Gly	Leu
			740						745					750	
Gly	Gly	Ile	Gly	Leu	Asp	Thr	Ser	Lys	Glu	Leu	Leu	Lys	Arg	Asp	Pro
		755					760						765		
Val	Val	Leu	Gln	Arg	Arg	Asp	Trp	Glu	Asn	Pro	Gly	Val	Thr	Gln	Leu
		770					775					780			
Asn	Arg	Leu	Ala	Ala	His	Pro	Pro	Phe	Ala	Ser	Trp	Arg	Asn	Ser	Glu
785					790						795				800
Glu	Ala	Arg	Thr	Asp	Arg	Pro	Ser	Gln	Gln	Leu	Arg	Ser	Leu	Asn	Gly
				805						810				815	
Glu	Trp	Arg	Phe	Ala	Trp	Phe	Pro	Ala	Pro	Glu	Ala	Val	Pro	Glu	Ser
			820					825					830		
Trp	Leu	Glu	Cys	Asp	Leu	Pro	Glu	Ala	Asp	Thr	Val	Val	Val	Pro	Ser
		835					840					845			
Asn	Trp	Gln	Met	His	Gly	Tyr	Asp	Ala	Pro	Ile	Tyr	Thr	Asn	Val	Thr
		850				855						860			
Tyr	Pro	Ile	Thr	Val	Asn	Pro	Pro	Phe	Val	Pro	Thr	Glu	Asn	Pro	Thr
865					870					875					880
Gly	Cys	Tyr	Ser	Leu	Thr	Phe	Asn	Val	Asp	Glu	Ser	Trp	Leu	Gln	Glu
				885					890					895	
Gly	Gln	Thr	Arg	Ile	Ile	Phe	Asp	Gly	Val	Asn	Ser	Ala	Phe	His	Leu
			900					905					910		
Trp	Cys	Asn	Gly	Arg	Trp	Val	Gly	Tyr	Gly	Gln	Asp	Ser	Arg	Leu	Pro
		915					920					925			

Ser Glu Phe Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu
930 935 940

Ala Val Met Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln
945 950 955 960

Asp Met Trp Arg Met Ser Gly Ile Phe Arg Asp Val Ser Leu Leu His
965 970 975

Lys Pro Thr Thr Gln Ile Ser Asp Phe His Val Ala Thr Arg Phe Asn
980 985 990

Asp Asp Phe Ser Arg Ala Val Leu Glu Ala Glu Val Gln Met Cys Gly
995 1000 1005

Glu Leu Arg Asp Tyr Leu Arg Val Thr Val Ser Leu Trp Gln Gly
1010 1015 1020

Glu Thr Gln Val Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu Ile
1025 1030 1035

Ile Asp Glu Arg Gly Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu
1040 1045 1050

Asn Val Glu Asn Pro Lys Leu Trp Ser Ala Glu Ile Pro Asn Leu
1055 1060 1065

Tyr Arg Ala Val Val Glu Leu His Thr Ala Asp Gly Thr Leu Ile
1070 1075 1080

Glu Ala Glu Ala Cys Asp Val Gly Phe Arg Glu Val Arg Ile Glu
1085 1090 1095

Asn Gly Leu Leu Leu Leu Asn Gly Lys Pro Leu Leu Ile Arg Gly
1100 1105 1110

Val Asn Arg His Glu His His Pro Leu His Gly Gln Val Met Asp
1115 1120 1125

Glu Gln Thr Met Val Gln Asp Ile Leu Leu Met Lys Gln Asn Asn

1130						1135						1140			
Phe	Asn	Ala	Val	Arg	Cys	Ser	His	Tyr	Pro	Asn	His	Pro	Leu	Trp	
	1145					1150					1155				
Tyr	Thr	Leu	Cys	Asp	Arg	Tyr	Gly	Leu	Tyr	Val	Val	Asp	Glu	Ala	
	1160					1165					1170				
Asn	Ile	Glu	Thr	His	Gly	Met	Val	Pro	Met	Asn	Arg	Leu	Thr	Asp	
	1175					1180					1185				
Asp	Pro	Arg	Trp	Leu	Pro	Ala	Met	Ser	Glu	Arg	Val	Thr	Arg	Met	
	1190					1195					1200				
Val	Gln	Arg	Asp	Arg	Asn	His	Pro	Ser	Val	Ile	Ile	Trp	Ser	Leu	
	1205					1210					1215				
Gly	Asn	Glu	Ser	Gly	His	Gly	Ala	Asn	His	Asp	Ala	Leu	Tyr	Arg	
	1220					1225					1230				
Trp	Ile	Lys	Ser	Val	Asp	Pro	Ser	Arg	Pro	Val	Gln	Tyr	Glu	Gly	
	1235					1240					1245				
Gly	Gly	Ala	Asp	Thr	Thr	Ala	Thr	Asp	Ile	Ile	Cys	Pro	Met	Tyr	
	1250					1255					1260				
Ala	Arg	Val	Asp	Glu	Asp	Gln	Pro	Phe	Pro	Ala	Val	Pro	Lys	Trp	
	1265					1270					1275				
Ser	Ile	Lys	Lys	Trp	Leu	Ser	Leu	Pro	Gly	Glu	Thr	Arg	Pro	Leu	
	1280					1285					1290				
Ile	Leu	Cys	Glu	Tyr	Ala	His	Ala	Met	Gly	Asn	Ser	Leu	Gly	Gly	
	1295					1300					1305				
Phe	Ala	Lys	Tyr	Trp	Gln	Ala	Phe	Arg	Gln	Tyr	Pro	Arg	Leu	Gln	
	1310					1315					1320				
Gly	Gly	Phe	Val	Trp	Asp	Trp	Val	Asp	Gln	Ser	Leu	Ile	Lys	Tyr	
	1325					1330					1335				

Asp Glu Asn Gly Asn Pro Trp Ser Ala Tyr Gly Gly Asp Phe Gly
1340 1345 1350

Asp Thr Pro Asn Asp Arg Gln Phe Cys Met Asn Gly Leu Val Phe
1355 1360 1365

Ala Asp Arg Thr Pro His Pro Ala Leu Thr Glu Ala Lys His Gln
1370 1375 1380

Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly Gln Thr Ile Glu Val
1385 1390 1395

Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn Glu Leu Leu His
1400 1405 1410

Trp Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser Gly Glu Val
1415 1420 1425

Pro Leu Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu Leu Pro
1430 1435 1440

Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr Val
1445 1450 1455

Arg Val Val Gln Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His
1460 1465 1470

Ile Ser Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn Leu Ser Val
1475 1480 1485

Thr Leu Pro Ala Ala Ser His Ala Ile Pro His Leu Thr Thr Ser
1490 1495 1500

Glu Met Asp Phe Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln Phe
1505 1510 1515

Asn Arg Gln Ser Gly Phe Leu Ser Gln Met Trp Ile Gly Asp Lys
1520 1525 1530

Lys Gln Leu Leu Thr Pro Leu Arg Asp Gln Phe Thr Arg Ala Pro

1535		1540		1545
Leu Asp Asn Asp Ile Gly Val Ser Glu Ala Thr Arg Ile Asp Pro	1550	1555		1560
Asn Ala Trp Val Glu Arg Trp Lys Ala Ala Gly His Tyr Gln Ala	1565	1570		1575
Glu Ala Ala Leu Leu Gln Cys Thr Ala Asp Thr Leu Ala Asp Ala	1580	1585		1590
Val Leu Ile Thr Thr Ala His Ala Trp Gln His Gln Gly Lys Thr	1595	1600		1605
Leu Phe Ile Ser Arg Lys Thr Tyr Arg Ile Asp Gly Ser Gly Gln	1610	1615		1620
Met Ala Ile Thr Val Asp Val Glu Val Ala Ser Asp Thr Pro His	1625	1630		1635
Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala Gln Val Ala Glu	1640	1645		1650
Arg Val Asn Trp Leu Gly Leu Gly Pro Gln Glu Asn Tyr Pro Asp	1655	1660		1665
Arg Leu Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu Pro Leu Ser	1670	1675		1680
Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly Leu Arg	1685	1690		1695
Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His Gln Trp Arg Gly	1700	1705		1710
Asp Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln Gln Gln Leu Met	1715	1720		1725
Glu Thr Ser His Arg His Leu Leu His Ala Glu Glu Gly Thr Trp	1730	1735		1740

Leu Asn Ile Asp Gly Phe His Met Gly Ile Gly Gly Asp Asp Ser
 1745 1750 1755

Trp Ser Pro Ser Val Ser Ala Glu Leu Gln Leu Ser Ala Gly Arg
 1760 1765 1770

Tyr His Tyr Gln Leu Val Trp Cys Gln Lys
 1775 1780

<210> 96
 <211> 5349
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

<400> 96
 atggaggacg ccaagaacat caagaagggg cccgctccct tctacccgct ggaggacggc 60
 accgctggcg agcagctgca caaggccatg aagaggtacg cactggtgcc cggcacaatc 120
 gccttcaccg acgcccacat cgaggtggac atcacctacg ccgagtactt cgagatgagc 180
 gtgcggctgg ccgaagccat gaagcggtag ggcctgaaca ccaaccaccg catcgtggtg 240
 tgctccgaga actccctgca gttcttcatg cccgttcttg gcgccctgtt cattgggggtg 300
 gccgtggcac cagccaacga catctacaac gagagggagc tgctgaacag catgggcatc 360
 agccagccca ccgtggtgtt cgtgtccaag aagggcctgc agaagatcct gaacgtgcag 420
 aagaagctgc ccatcatcca gaagatcatc atcatggaca gcaagaccga ctaccagggc 480
 ttccagtcca tgtacacctt cgtgaccagc cacctgcctc ccggcttcaa cgagtacgac 540
 ttcgtgcccc agtccttcga cagggacaag accatcgccc tgatcatgaa ctccagcgga 600
 agcaccggcc tgcccaaggg tgtggctcta cccaccgga ccgcctgtgt acggttcagc 660
 cacgcccgtg accccatctt cgggaaccag atcatccccg acaccgcat cctgagcgtg 720
 gtgcccttcc accacggctt cggcatgttc accaccctgg gctacctgat ctgcggcttc 780
 cgggtggtgc tgatgtaccg gttcaggag gagctgttcc tgcggagcct gcaggactac 840
 aagatccagt ccgccctgct ggtgcccacc ctgttctcct tcttcgcaa gtccaccctg 900

atcgacaagt acgacctgag caacctgcac gagattgcct ctggcggcgc acctctgtcc	960
aaggaggtcg gggaggccgt agccaagcgg ttccacctgc ccggcatccg gcaaggctac	1020
ggcctgaccg agaccaccag cgccatcctg atcacccctg agggcgacga caagcccggc	1080
gcagtgggga aggtggtgcc attcttcgag gccaaaggtgg tggacctgga caccggcaag	1140
accctgggcg tgaaccagcg gggcgagctg tgcgtgcggg gacctatgat catgtccggc	1200
tacgtgaaca accccgaggc caccaacgcc ctgattgaca aggacgggtg gctgcacagc	1260
ggcgacatcg cctactggga cgaggacgag cacttcttca tcgtggaccg cctgaagtcc	1320
ctgatcaagt acaaggggta ccaggtggct cccgccgagc tggagtccat cctgctgcag	1380
cacccaaca tctttgacgc cggcgttgcc ggactgcctg acgacgacgc tggggaactc	1440
cccgagctg tgggtggtgct ggagcacggc aagaccatga ccgagaagga gatcgtggac	1500
tacgtgcct cccaggtgac caccgccaag aagttacgcg gcggcgctcgt gttcgtggac	1560
gaggtgcca agggcctgac cggcaagctg gacgccgca agatccggga gatcctgac	1620
aaggccaaga agggcgggaa gatcgccgtg ggctctggcg aaggccgggg tagcctgctg	1680
acctgcggcg acgtggagga gaacctgga cctggatccg gcgccacca cttcagcctg	1740
ctcaagcagg ccggagacgt ggaagagaat ccaggcccag gcagcggta gtgcaccaac	1800
tacgccctgc tgaagctggc cggcgacgta gagagcaacc ccggccccgc cagcaacttc	1860
accagttcg tgctggtgga caacggcggt accggagacg tgaccgtggc cccttctaac	1920
ttcgccaacg gcatcgccga gtggatcagc agcaacagca gaagccaggc ctacaaggtg	1980
acctgcagcg tgagacagag cagcgcccag aacagaaagt acaccatcaa ggtggaggtg	2040
cctaagggcg cctggagaag ctacctgaac atggagctga ccatccctat cttcgccacc	2100
aacagcgact gcgagctgat cgtgaaggcc atgcagggcc tgctgaagga cggcaaccct	2160
atccctagcg ccatcgccgc caattcaggc atctacggag gcggtggaag ctcgttcacg	2220
ctaacgaaca agaacgtcat cttcgtagcg ggacttggcg gtatcggcct agacacgtcg	2280
aaggaactac taaagcgtga cccggtagtc ctccaacgtc gcgattggga gaaccgggc	2340
gtaacgcaac taaaccgtct tgcggcgcac ccgccgtttg cgtcgtggcg taactcggag	2400
gaggcgcgaa cggatcgtcc gtcgcaaca ctacgttcgc tcaacgggga gtggcgcttc	2460
gcgtggttcc cggcgccgga ggcggtaccg gagtcgtggc tcgagtgcga tctaccggag	2520

gcggacacgg tcgtcgtacc gtcgaactgg caaatgcacg gttacgacgc gccgatatac	2580
acgaacgtca cgtacccgat aacggtaaac ccgccgttcg tcccgcacgga gaacccgacg	2640
gggtgctact cgctaacgtt caacgttgac gagtcgtggt tgcaagaggg tcaaacgcgt	2700
atcatattcg acggtgtaaa ctcggcggtt cacctgtggt gcaacgggcg ctgggtaggg	2760
tacggccaag actcgcgtct accgtcggag ttcgacctat cggcgttcct acgagcgggt	2820
gagaaccggc tagcgggtcat ggtcctacgt tggtcggacg gttcgtacct cgaggaccaa	2880
gacatgtggc gaatgtcggg tatcttccgc gacgtatcg tcctacacaa gccgacgacg	2940
caaatctcgg acttccacgt cgcgacgcgt ttcaacgacg atttctcgcg ggcagtccta	3000
gaggcggagg tccaaatgtg cggggagcta cgtgactacc tccgtgtcac ggtatcgctc	3060
tggcaaggtg agacgcaagt agcgtcgggt acggcgccgt tcggcgggtga gatcatcgac	3120
gagcgtggtg ggtacgcgga ccgtgtaacg ctacgtctaa acgtcgagaa cccgaagctc	3180
tggtcggcgg agatcccgaa cctataccgt gcggtcgtcg agctacatac ggcggacggg	3240
acgctaatag aggcggaagc gtgcgacgtc gggtttcgag aggttcgtat agagaacggg	3300
ctgctacttc taaacgggaa gccgttgctc atacgtggtg tcaaccgtca cgagcaccac	3360
ccgctacacg gtcaagtaat ggacgagcaa acgatggtac aagacatcct actaatgaag	3420
cagaacaact tcaacgcggt acgctgttcg cattaccga accatccgtt gtggtacacg	3480
ctttgcgacc gatacggctc atacgtcgta gacgaggcga acatagagac gcacgggatg	3540
gtaccgatga atcgcctaac ggacgacccg cgttggttac cggcgatgtc ggagcgagtc	3600
acgcgtatgg tccaacggga ccgtaaccac ccgtcggtaa taatctggtc gctaggcaac	3660
gaatcggggc acggggcgaa ccacgacgcg ctataccgtt ggatcaagtc ggtagacccg	3720
tcgcgtccgg tacaatacga aggtggcggt gcggacacga cggcgacgga catcatctgc	3780
ccgatgtacg cgcgcgtcga cgaagaccaa ccgttcccgg cggtagcgaa gtggtcgatc	3840
aagaagtggc tctcgttgcc ggggtgaaacg cgtccgttga tactttgcga gtacgcgcac	3900
gcgatgggca actcgttggg tgggttcgcg aagtactggc aggcgttcg tcaatacccg	3960
cgtctacagg gtgggttcgt ctgggactgg gtagaccaat cgctaataca gtacgacgag	4020
aacggcaacc cgtggtcggc gtacggtggg gacttcgggg acacgccgaa cgaccgcaa	4080
ttctgtatga acggcctagt cttcgcggac cgaacgccgc acccggcgtt gacggaggcg	4140

aagcatcaac aacaattctt ccaattccgt ctatcggggc aaacgacga ggtaacgtcg	4200
gagtacttgt tccggcactc ggacaacgag ctactacact ggatggtagc actagacggc	4260
aagccgctag cgtcgggaga agtccccttg gacgtcgcgc cgcaaggtaa gcaactaatc	4320
gagctaccgg agctaccga accggagtcg gcgggtcaac tgtggttgac ggtccgtgtc	4380
gttcaaccga acgcgacggc gtggtcggag gcgggtcaca tctcggcgtg gcagcagtgg	4440
cgtctagcgg agaacctctc ggtcacgcta ccggcggcgt cgcacgcgat accgcatcta	4500
acgacgtcgg agatggactt ctgcatcgag ttggggaaca agaggtggca gttcaaccgt	4560
caatcgggat tcctatcgca aatgtggata ggtgacaaga agcaactact aacgccgcta	4620
cgtgatcagt tcacgcgtgc tccgctagac aacgacatag gtgtttcgga ggcgacgcgt	4680
atagacccga acgcgtgggt ggagcgggtg aaggcggcgg ggcactacca agcggaggcg	4740
gcgctactac agtgcacggc ggacacgcta gcggacgcgg tattgatcac gacggcgcac	4800
gcgtggcaac accaggggaa gacgctattc atctcgcgta agacgtaccg tatcgacggt	4860
tcgggccaaa tggcgatcac ggtcgacgta gaggtagcgt cggacacgcc gcatccggcg	4920
cgcatcggtc taaactgcca actagcgcaa gtagcggagc gtgtaaactg gctagggcta	4980
gggccgcaag agaactatcc ggaccgccta acggcggcgt gcttcgaccg ttgggacct	5040
ccgctttcgg acatgtatac cccgtacgtc ttcccgtcgg agaacggggt gaggtgcggg	5100
acgcgcgagc taaactacgg gccgcaccag tggcgagggg acttccaatt caacatatcg	5160
cgttactcgc aacaacaact aatggagacg tcgcaccgtc acctactaca cgcggaggag	5220
gggacgtggc taaacatcga cgggttcac atgggcatag gtggggacga ctcgtggtcg	5280
ccgtcgggtc cggcggagct ccaactctcg gcgggtcgtt accattacca actagtttgg	5340
tgccagaag	5349

<210> 97

<211> 1047

<212> PRT

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 97

Met Ser Phe Thr Leu Thr Asn Lys Asn Val Ile Phe Val Ala Gly Leu
1 5 10 15

Gly Gly Ile Gly Leu Asp Thr Ser Lys Glu Leu Leu Lys Arg Asp Pro
20 25 30

Val Val Leu Gln Arg Arg Asp Trp Glu Asn Pro Gly Val Thr Gln Leu
35 40 45

Asn Arg Leu Ala Ala His Pro Pro Phe Ala Ser Trp Arg Asn Ser Glu
50 55 60

Glu Ala Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly
65 70 75 80

Glu Trp Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Glu Ser
85 90 95

Trp Leu Glu Cys Asp Leu Pro Glu Ala Asp Thr Val Val Val Pro Ser
100 105 110

Asn Trp Gln Met His Gly Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr
115 120 125

Tyr Pro Ile Thr Val Asn Pro Pro Phe Val Pro Thr Glu Asn Pro Thr
130 135 140

Gly Cys Tyr Ser Leu Thr Phe Asn Val Asp Glu Ser Trp Leu Gln Glu
145 150 155 160

Gly Gln Thr Arg Ile Ile Phe Asp Gly Val Asn Ser Ala Phe His Leu
165 170 175

Trp Cys Asn Gly Arg Trp Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro
180 185 190

Ser Glu Phe Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu
195 200 205

Ala Val Met Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln
210 215 220

Asp Met Trp Arg Met Ser Gly Ile Phe Arg Asp Val Ser Leu Leu His
225 230 235 240

Lys Pro Thr Thr Gln Ile Ser Asp Phe His Val Ala Thr Arg Phe Asn
245 250 255

Asp Asp Phe Ser Arg Ala Val Leu Glu Ala Glu Val Gln Met Cys Gly
260 265 270

Glu Leu Arg Asp Tyr Leu Arg Val Thr Val Ser Leu Trp Gln Gly Glu
275 280 285

Thr Gln Val Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp
290 295 300

Glu Arg Gly Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu Asn Val Glu
305 310 315 320

Asn Pro Lys Leu Trp Ser Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val
325 330 335

Val Glu Leu His Thr Ala Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys
340 345 350

Asp Val Gly Phe Arg Glu Val Arg Ile Glu Asn Gly Leu Leu Leu Leu
355 360 365

Asn Gly Lys Pro Leu Leu Ile Arg Gly Val Asn Arg His Glu His His
370 375 380

Pro Leu His Gly Gln Val Met Asp Glu Gln Thr Met Val Gln Asp Ile
385 390 395 400

Leu Leu Met Lys Gln Asn Asn Phe Asn Ala Val Arg Cys Ser His Tyr
405 410 415

Pro Asn His Pro Leu Trp Tyr Thr Leu Cys Asp Arg Tyr Gly Leu Tyr
420 425 430

Val Val Asp Glu Ala Asn Ile Glu Thr His Gly Met Val Pro Met Asn
435 440 445

Arg Leu Thr Asp Asp Pro Arg Trp Leu Pro Ala Met Ser Glu Arg Val
450 455 460

Thr Arg Met Val Gln Arg Asp Arg Asn His Pro Ser Val Ile Ile Trp
465 470 475 480

Ser Leu Gly Asn Glu Ser Gly His Gly Ala Asn His Asp Ala Leu Tyr
485 490 495

Arg Trp Ile Lys Ser Val Asp Pro Ser Arg Pro Val Gln Tyr Glu Gly
500 505 510

Gly Gly Ala Asp Thr Thr Ala Thr Asp Ile Ile Cys Pro Met Tyr Ala
515 520 525

Arg Val Asp Glu Asp Gln Pro Phe Pro Ala Val Pro Lys Trp Ser Ile
530 535 540

Lys Lys Trp Leu Ser Leu Pro Gly Glu Thr Arg Pro Leu Ile Leu Cys
545 550 555 560

Glu Tyr Ala His Ala Met Gly Asn Ser Leu Gly Gly Phe Ala Lys Tyr
565 570 575

Trp Gln Ala Phe Arg Gln Tyr Pro Arg Leu Gln Gly Gly Phe Val Trp
580 585 590

Asp Trp Val Asp Gln Ser Leu Ile Lys Tyr Asp Glu Asn Gly Asn Pro
595 600 605

Trp Ser Ala Tyr Gly Gly Asp Phe Gly Asp Thr Pro Asn Asp Arg Gln
610 615 620

Phe Cys Met Asn Gly Leu Val Phe Ala Asp Arg Thr Pro His Pro Ala
625 630 635 640

Leu Thr Glu Ala Lys His Gln Gln Gln Phe Phe Gln Phe Arg Leu Ser
645 650 655

Gly Gln Thr Ile Glu Val Thr Ser Glu Tyr Leu Phe Arg His Ser Asp
660 665 670

Asn Glu Leu Leu His Trp Met Val Ala Leu Asp Gly Lys Pro Leu Ala
675 680 685

Ser Gly Glu Val Pro Leu Asp Val Ala Pro Gln Gly Lys Gln Leu Ile
690 695 700

Glu Leu Pro Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu
705 710 715 720

Thr Val Arg Val Val Gln Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly
725 730 735

His Ile Ser Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn Leu Ser Val
740 745 750

Thr Leu Pro Ala Ala Ser His Ala Ile Pro His Leu Thr Thr Ser Glu
755 760 765

Met Asp Phe Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg
770 775 780

Gln Ser Gly Phe Leu Ser Gln Met Trp Ile Gly Asp Lys Lys Gln Leu
785 790 795 800

Leu Thr Pro Leu Arg Asp Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp
805 810 815

Ile Gly Val Ser Glu Ala Thr Arg Ile Asp Pro Asn Ala Trp Val Glu
820 825 830

Arg Trp Lys Ala Ala Gly His Tyr Gln Ala Glu Ala Ala Leu Leu Gln
835 840 845

Cys Thr Ala Asp Thr Leu Ala Asp Ala Val Leu Ile Thr Thr Ala His
850 855 860

Ala Trp Gln His Gln Gly Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr
865 870 875 880

Arg Ile Asp Gly Ser Gly Gln Met Ala Ile Thr Val Asp Val Glu Val
885 890 895

Ala Ser Asp Thr Pro His Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu
900 905 910

Ala Gln Val Ala Glu Arg Val Asn Trp Leu Gly Leu Gly Pro Gln Glu
915 920 925

Asn Tyr Pro Asp Arg Leu Thr Ala Ala Cys Phe Asp Arg Trp Asp Leu
930 935 940

Pro Leu Ser Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser Glu Asn Gly
945 950 955 960

Leu Arg Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His Gln Trp Arg
965 970 975

Gly Asp Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln Gln Gln Leu Met
980 985 990

Glu Thr Ser His Arg His Leu Leu His Ala Glu Glu Gly Thr Trp Leu
995 1000 1005

Asn Ile Asp Gly Phe His Met Gly Ile Gly Gly Asp Asp Ser Trp
1010 1015 1020

Ser Pro Ser Val Ser Ala Glu Leu Gln Leu Ser Ala Gly Arg Tyr
1025 1030 1035

His Tyr Gln Leu Val Trp Cys Gln Lys
1040 1045

<210> 98

<211> 3141

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 98

atgagcttca ccctgaccaa caagaacgtg atcttcgtgg ccggcctggg cggcatcggc	60
ctggacacca gcaaggagct gctgaagcgg gaccccgtgg tgctgcagcg ccgggactgg	120
gagaaccccg gcgtgaccca gctgaacagg ctggccgccc atcctccgtt cgcctcctgg	180
cggaactccg aggaggcccg gaccgaccgg ccctcccagc agctgcggtc tctgaacggc	240
gagtggaggt tcgcctgggt ccccgcccca gaggccgtgc ctgagagctg gctggagtgc	300
gacctgcctg aagccgacac cgtgggtggg cctagcaact ggagatgca cggctacgac	360
gcacccatct acaccaacgt gacctaccct atcaccgtga atccgccttt cgtgcccacc	420
gagaacccca ccggctgcta cagcctgacc ttcaacgtgg acgagagttg gctgcaggag	480
ggccagaccc gcatcatctt cgacggcgtg aacagcgcct tccacctgtg gtgcaacggc	540
cggtgggtgg gctacggcca ggactcccgg ctgccctccg agtttgacct gtccgccttc	600
ctgcgcgccg gcgagaaccg gctggccgtg atgggtgctga ggtggagcga cggcagctac	660
ctggaggacc aggacatgtg gcgcatgtcc ggaatcttcc gggacgtgag cctgctgcac	720
aagcccacca cccagatcag cgacttcac gtggccaccc ggttcaacga cgacttcagc	780
cgggccgtgc tggaggccga ggtgcagatg tgcggcgagc tgcgggacta cctgcgcgtg	840
accgtgagcc tgtggcaggg cgagaccag gtggcttccg gcaccgcccc attcggcggc	900
gagatcatcg acgagcgggg cggctacgcc gaccgggtga ccctgcggct gaacgtggag	960
aacccaagc tgtggagcgc cgagatcccc aacctgtacc gggccgtggt ggagctgcac	1020
accgccgacg gcaccctgat cgaggccgag gcctgcgacg tgggcttccg ggaggtgagg	1080
atcgagaacg ggctgctgct gctgaacggg aagccactgc tgatcagggg cgtgaaccgc	1140
cacgagcacc atcccctgca cggccaggtg atggacgagc agaccatggt gcaggacatc	1200
ctgctgatga agcagaacaa cttcaacgcc gtgaggtgct ccactaccc caaccatccc	1260
ctgtggtaca ccctgtgcga ccggtacggc ctgtacgtgg tggacgaggc caacatcgag	1320
accacggga tgggtccaat gaaccggctg actgacgacc ctgctggct gcccgccatg	1380
agcgagcggg tgaccaggat ggtgcagcgc gaccggaacc acccagcgt gatcatctgg	1440
agcctgggga acgagtccgg ccacggcgcc aaccacgacg ccctgtaccg ctggatcaag	1500

agcgtggacc cctcccgccc cgtgcagtac gagggcggcg gcgccgacac caccgccacc	1560
gacatcatct gccccatgta cgccccgggtg gacgaggacc agcccttccc tgccgtgccc	1620
aagtggagca tcaagaagtg gctgagcttg cccggcgaga ccagaccgct gatcctgtgc	1680
gagtacgccc acgccatggg caactccctg ggcggttctg ccaagtactg gcaggccttc	1740
cggcagtacc cgcggtgca gggcggttct gtgtgggact ggggtggacca gtccctgata	1800
aagtacgacg agaacggcaa cccctggagc gcctacggcg gggacttcgg ggacacaccc	1860
aacgaccgcc agttctgcat gaacggcctg gtgttcgccg accgcacccc gcaccctgcc	1920
ctgaccgagg ccaagcacca gcagcagttc ttccagttcc ggctgagcgg gcagaccatc	1980
gaggtgacca gcgagtacct gttccggcac agcgacaacg agctgctgca ctggatgggtg	2040
gccctggacg gcaagcccct ggccagcggc gaggtgcctc tggacgtggc tcccagggc	2100
aagcagctga tcgagctgcc cgagctgccc cagcccgaga gcgccggcca gctgtggctg	2160
accgtgaggg tggcgcagcc caacgccacc gcctggagcg aggccggcca catcagcgcc	2220
tggcaacagt ggcggttggc cgagaacctg agcgtgacct tgcccggcg cagccacgcc	2280
atccctcacc tgaccacctc cgagatggac ttctgcatcg agctgggcaa caagcggtgg	2340
cagttcaacc ggcagtccgg cttcctgagc cagatgtgga tcggcgacaa gaagcagctg	2400
ctgacccccc tgcggggacca gttcaccagg gccccactgg acaacgacat cggcgtgagc	2460
gaggccacca gaatcgacct caacgcctgg gtggagcggg ggaaggccgc cggccactac	2520
caggccgagg ccgcccctgt gcagtgcacc gccgacacct tggccgacgc cgtgctgata	2580
accaccgccc acgcctggca gcaccagggg aagaccctgt tcatcagccg gaagacctac	2640
cggatcgacg ggtccggcca gatggccatc accgtggacg tggaggtggc ctccgacacc	2700
cctcaccgccc cccggatcgg cctgaactgt cagctggccc aggtggccga gcgggtgaac	2760
tggctggggc tgggcccaca ggagaactac cccgaccggc tgaccgccgc ctgcttcgac	2820
cgggtgggacc tgcccctgag cgacatgtac accccatacg tgttcccctc tgagaacggc	2880
ctgcggtgcg gcacccggga gctgaactac gggccccacc agtggcgcgg cgacttcag	2940
ttcaacatca gccggtacag ccagcagcag ctgatggaga ccagccaccg gcacctgctg	3000
cacgccgagg agggcacctg gctgaacatc gacggcttcc acatgggcat cggcggcgac	3060
gacagctggg ccccatccgt gagcgccgag ctgcagctga gcgccggccg gtaccactac	3120

cagctggtgt ggtgccagaa g 3141

<210> 99
<211> 133
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 99
gcuggagccu cgguggccau gcuucuugcc ccuugggccu cccccagcc ccuccucucc 60
uuccugcacc cguacccccc ccauaaagua ggaaacacua caguggucuu ugaauaaagu 120
cugagugggc ggc 133

<210> 100
<211> 141
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 100
ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagccc 60
cuccucuccu uccugcacc guaccccccg cauauuacu cacgguacga guggucuuug 120
aaauaaagucu gagugggcgg c 141

<210> 101
<211> 119
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 101
ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagccc 60
cuccucuccu uccugcacc guacccccgu ggucuuugaa uaaagucuga gugggcggc 119

<210> 102
<211> 142
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 102
ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagccc 60
cuccuccccu uccugcaccc guacccccuc cauaaaguag gaaacacuac aguggucuuu 120
gaauaaaguc ugagugggcg gc 142

<210> 103
<211> 164
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 103
ugauaaauagu ccauaaagua ggaaacacua cagcuggagc cucgguggcc augcuucuug 60
cccuuggggc cuccccccag cccuccucc ccuuccugca cccguacccc ccgcauuauu 120
acucacggua cgaguggucu uugaauaaag ucugaguggg cggc 164

<210> 104
<211> 188
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 104
ugauaaauagu ccauaaagua ggaaacacua cagcuggagc cucgguggcc augcuucuug 60
cccuuggggc cuccaauaaag uaggaaacac uacaucaccc cagccccucc ucccuuccu 120
gcacccguac ccccucaua aaguaggaaa cacuacagug gucuugaau aaagucugag 180

ugggcggc 188

<210> 105
<211> 140
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 105
ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagccc 60
cuccucuccu uccugcacc guacccccag uagugcuuuc uacuuuauagg uggucuuga 120
auaaagucug agugggcggc 140

<210> 106
<211> 182
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 106
ugauaaauaga guagugcuuu cuacuuuau gcuuggagccu cgguggccau gcuucuugcc 60
ccuugggcca guagugcuuu cuacuuuau uccccccagc ccuccuccu cuuccugcac 120
ccguaccccc aguagugcuu ucuacuuuau gguggucuuu gaauaaaguc ugagugggcg 180
gc 182

<210> 107
<211> 184
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 107
ugauaaauaga guagugcuuu cuacuuuau gcuuggagccu cgguggccau gcuucuugcc 60

ccuugggccu ccauaaagua ggaaacacua caucuuuuuu gcccuccuc ccuuccugc	120
acccguaccc ccaguagugc uuucuacuuu augguggucu uugaauaaag ucugaguggg	180
cggc	184

<210> 108
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 108	
ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc ccccagccc	60
cuccucuccu uccugcacc guacccccac ccuaucau auuagcauu aguggucuuu	120
gaauaaaguc ugagugggcg gc	142

<210> 109
 <211> 188
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 109	
ugauaaauaga cccuaucau aauuagcau aagcuggagc cucgguggcc augcuucuug	60
cccuugggc caccuauca acauuagca uuaaucccc cagccccucc ucccuuccu	120
gcaccguac cccaccccu aucacaaua gcauuagug gucuugaau aaagucugag	180
ugggcggc	188

<210> 110
 <211> 188
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 110
ugauaaauaga ccccuauacac aauuagcauu aagcuggagc cucgguggcc augcuucuug 60
cccuuugggc cuccaauaaag uaggaaacac uacauccccc cagccccucc ucccuuccu 120
gcacccguac ccccaccccu aucacaauua gcauuagug gucuuugaau aaagucugag 180
ugggcggc 188

<210> 111
<211> 142
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 111
ugauaaauagu ccauaaagua ggaaacacua cagcuggagc cucgguggcc augcuucuug 60
cccuuugggc cuccccccag cccuccucc ccuuccugca cccguacccc cguggucuuu 120
gaauaaaguc ugagugggcg gc 142

<210> 112
<211> 142
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 112
ugauaaauagg cuggagccuc gguggcucca uaaaguagga aacacuacac augcuucuug 60
cccuuugggc cuccccccag cccuccucc ccuuccugca cccguacccc cguggucuuu 120
gaauaaaguc ugagugggcg gc 142

<210> 113
<211> 142
<212> RNA
<213> Artificial Sequence

<220>

<221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 113
 ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cauaaaguag 60
 gaaacacuac auccccccag cccuccucc ccuuccugca cccguacccc cguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 114
 <211> 181
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 114
 ugauaauga guagugcuuu cuacuuuauug gcuggagccu cgguggccau gcuucuugcc 60
 ccuugggccca guagugcuuu cuacuuuauug uccccccagc cccucucucc uuccugcacc 120
 cguacccccca guagugcuuu cuacuuuauug guggucuuug aauaaagucu gagugggcgg 180
 c 181

<210> 115
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 115
 ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuuccauaaa guaggaaaca 60
 cuacaugggc cuccccccag cccuccucc ccuuccugca cccguacccc cguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 116
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 116
 ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagucc 60
 auaaaguagg aaacacuaca cccuccucc ccuuccugca cccguacccc cguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 117
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 117
 ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagccc 60
 cuccuccccu ucuccauaaa guaggaaaca cuacacugca cccguacccc cguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 118
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 118
 ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc cccccagccc 60
 cuccuccccu uccugcacc guacccccgu ggucuuugaa uaaaguucca uaaaguagga 120
 aacacuacac ugagugggcg gc 142

<210> 119
 <211> 164
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 119
 ugauaaugu ccauaaagua ggaaacacua cagcuggagc cucgguggcc uagcuucuug 60
 ccccuugggc cuccccccag cccuccucc ccuuccugca cccguacccc ccgcauuauu 120
 acucacggua cgaguggucu uugaauaaag ucugaguggg cggc 164

<210> 120
 <211> 119
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 120
 ugauaaugg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 60
 cuccucccu uccugcacc guacccccgu ggucuugaa uaaagucuga gugggcggc 119

<210> 121
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 121
 ugauaaugg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 60
 cuccucccu uccugcacc guacccccuc cauaaaguag gaaacacuac aguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 122
 <211> 141
 <212> RNA
 <213> Artificial Sequence

<220>

<221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 122
 ugauaaauagg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 60
 cuccuccccu uccugcaccc guaccccccg cauuaauacu cacgguacga guggucuuug 120
 aaauaaagucu gagugggcgg c 141

<210> 123
 <211> 188
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 123
 ugauaaauagu ccauaaagua ggaaacacua cagcuggagc cucgguggcc uagcuucuug 60
 ccccuugggc cuccauaaag uaggaaacac uacaucuccc cagccccucc ucccuuccu 120
 gcacccguac ccccucaua aaguaggaaa cacuacagug gucuuugaau aaagucugag 180
 ugggcggc 188

<210> 124
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 124
 ugauaaauagu ccauaaagua ggaaacacua cagcuggagc cucgguggcc uagcuucuug 60
 ccccuugggc cccccccag cccuccucc ccuuccugca cccguacccc cguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 125
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 125
 ugauaaauagg cuggagccuc gguggcucca uaaaguagga aacacuacac uagcuucuug 60
 ccccuugggc cuccccccag cccuccucc ccuuccugca cccguacccc cguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 126
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 126
 ugauaaauagg cuggagccuc gguggccuag cuucuugccc cuugggccuc cauaaaguag 60
 gaaacacuac auccccccag cccuccucc ccuuccugca cccguacccc cguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 127
 <211> 142
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 127
 ugauaaauagg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 60
 cuccuccccu uccugcacc guacccccac ccuauacaca auuagcauua aguggucuuu 120
 gaauaaaguc ugagugggcg gc 142

<210> 128
 <211> 188
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

 <400> 128
 ugauaaauaga ccccuauacac aaauagcauu aagcuggagc cucgguggcc uagcuucuug 60

 ccccuugggc caccuuauac acaauuagca uuaauccccc cagccccucc ucccuuccu 120

 gcacccguac cccaccccu aucacaauua gcauuagug gucuuugaau aaagucugag 180

 ugggcggc 188

<210> 129
 <211> 188
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

 <400> 129
 ugauaaauaga ccccuauacac aaauagcauu aagcuggagc cucgguggcc uagcuucuug 60

 ccccuugggc cucauaaaag uaggaaacac uacauccccc cagccccucc ucccuuccu 120

 gcacccguac cccaccccu aucacaauua gcauuagug gucuuugaau aaagucugag 180

 ugggcggc 188

<210> 130
 <211> 141
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polynucleotide"

 <400> 130
 ugauaaauagg cuggagccuc gguggccaug cuucuugccc cuugggccuc ccccgagccc 60

 cuccucccu uccugcacc guacccccca aacaccauug ucacacucca guggucuug 120

 aaauaaagucu gagugggcgg c 141

<210> 131
<211> 141
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 131
ugauaaauagg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 60
cuccucuccu uccugcacc guacccccca aacaccauug ucacacucca guggucuuug 120
aaauaaagucu gagugggcgg c 141

<210> 132
<211> 141
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 132
ugauaaauagg cuggagccuc gguggccuag cuucuugccc cuugggccuc cccccagccc 60
cuccucuccu uccugcacc guaccccccg cauuauuacu cacgguacga guggucuuug 120
aaauaaagucu gagugggcgg c 141

<210> 133
<211> 47
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 133
gggaaauaag agagaaaaga agaguaagaa gaaauauaag agccacc 47

<210> 134
<211> 70
<212> RNA
<213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 134
 gggaaaauaag aguccauaaa guaggaaaca cuacaagaaa agaagaguua gaagaaauau 60
 aagagccacc 70

<210> 135
 <211> 70
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 135
 gggaaaauaag agagaaaaga agaguaaucc auaaaguagg aaacacuaca gaagaaauau 60
 aagagccacc 70

<210> 136
 <211> 70
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 136
 gggaaaauaag agagaaaaga agaguaagaa gaaauuaau ccauaaagua ggaaacacua 60
 cagagccacc 70

<210> 137
 <211> 47
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 137
aggaaauaag agagaaaaga agaguaagaa gaaauauaag agccacc 47

<210> 138
<211> 57
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 138
gggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccacc 57

<210> 139
<211> 57
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 139
aggaaauaag agagaaaaga agaguaagaa gaaauauaag accccggcgc cgccacc 57

<210> 140
<211> 47
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 140
gggagaucag agagaaaaga agaguaagaa gaaauauaag agccacc 47

<210> 141
<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic

oligonucleotide"

<400> 141
gggagacaag cuuggcauuc cgguacuguu gguaaagcca cc 42

<210> 142
<211> 47
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 142
gggaauuaac agagaaaaga agaguaagaa gaaauuaag agccacc 47

<210> 143
<211> 47
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 143
gggaaauuag acagaaaaga agaguaagaa gaaauuaag agccacc 47

<210> 144
<211> 47
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 144
gggaaauaag agaguaaaga acaguaagaa gaaauuaag agccacc 47

<210> 145
<211> 47
<212> RNA
<213> Artificial Sequence

<220>

<221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 145
 gggaaaaaag agagaaaaga agacuaagaa gaaauuaag agccacc 47

<210> 146
 <211> 47
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 146
 gggaaauaag agagaaaaga agaguaagaa gauauuaag agccacc 47

<210> 147
 <211> 47
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 147
 gggaaauaag agacaaaaca agaguaagaa gaaauuaag agccacc 47

<210> 148
 <211> 47
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 148
 gggaaauuag agaguaaaga acaguaagua gaauuaaaag agccacc 47

<210> 149
 <211> 47
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 149
 gggaaaauaag agagaauaga agaguaagaa gaaauuaag agccacc 47

<210> 150
 <211> 47
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 150
 gggaaaauaag agagaaaaga agaguaagaa gaaaauuaag agccacc 47

<210> 151
 <211> 47
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 151
 gggaaaauaag agagaaaaga agaguaagaa gaaaauuaag agccacc 47

<210> 152
 <211> 92
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 152
 ucaagcuuuu ggaccucgu acagaagcua auacgacuca cuauaggga auaagagaga 60
 aaagaagagu aagaagaaau auaagagcca cc 92

<210> 153
<211> 120
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 153
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 120

<210> 154
<211> 925
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 154
tgataatagg cgcatgatc acacgcgcac gcgaagcttg tgttcgcacg gcaagcagca 60
gctgccgtgg aggcttcac ggcgaaatgt ccacgaatac gggacaaata attgatatag 120
ggccggcccg taagagtact acgggatgat tctcttgact taagtgccgg accactaggc 180
cggcaacgga cagacaaaga ggctctggga ctaccaagcc cagtcacacg gtacttcaag 240
gcaccgggaa cagtagtccc ggagtgttga cttatagaag tagtcccatt attaaaggga 300
cttcgggact tcccttctcg ctcccgacca gtaggcggga aatgcgccct cccgtactgt 360
cctgcacat tagggcaggt tctttgaggc tctctcgctc gtgtatgaat acaacacgtg 420
tgcgtcctgt cccctcgtga ctcaggacta ccgagtcaca gcaaacttaa aggatcgcac 480
ggacgaacac gccgtgacat agctaatacag acaatggcac cagtaccaac ggtgcacgtt 540
gacactcgaa cgtgattgac aggattaccg tcaactcttc cactagcccg ttactggcaa 600
ggattacctg ccactgtttt gtcttaaccg gttgacgtat caacagaacg tctggtagca 660
gccacgctta ctgggccatt aataaaggcc caaggaactt aaagccagtg cccggcacta 720
tcaaggccgg gcatcattga ttcgacgccg gacggaggag cccccgtcg ggcacgcgg 780
ccctgcgggt gctaaacagt agttagcaga catttagcgc cggtttgaag cccacagta 840

cggggccttag cagctccaga tacacggcgg gaactactaa gtcccgtaaa ctctccacta 900
 tcgccggggcc aatcaataga tggcc 925

<210> 155
 <211> 478
 <212> PRT
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 polypeptide"

<400> 155
 Asp Pro Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro
 1 5 10 15

Ser Phe Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro
 20 25 30

Arg Gly Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Ala Gly Leu Gly
 35 40 45

Pro Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile
 50 55 60

Ala Thr Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys
 65 70 75 80

Ala Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly
 85 90 95

Glu Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val
 100 105 110

Arg Ser Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Gln Gln Leu Met
 115 120 125

Lys Gln Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly
 130 135 140

Val Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser

145		150		155		160
Val	Ala	Tyr	Ala	Asn	Met	Pro
			165			170
Thr	Thr	Glu	Lys	Pro	Thr	Val
		180			185	190
Asn	Arg	Cys	Gln	Lys	Glu	Phe
		195			200	205
Phe	Glu	Lys	Lys	Gln	Lys	Glu
	210				215	220
Arg	Gly	Arg	Leu	Lys	Glu	Glu
225				230		235
Arg	Arg	Ser	Leu	Gly	Asn	Ile
			245			250
Lys	Met	Leu	Thr	Glu	Ala	Ile
		260				265
Lys	Asn	His	Asp	Glu	Glu	Ser
		275				280
Thr	Ile	Gly	Lys	Asp	Leu	Asp
	290				295	
Gln	Tyr	Phe	Asn	Gln	Met	Glu
305					310	
Ser	Arg	Ile	Arg	Phe	Met	Leu
			325			330
Asn	Trp	Val	Pro	Arg	Arg	Gly
		340				345
Ile	His	Lys	Glu	Ala	Glu	Met
		355				360

Gln Gln Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro
370 375 380

Gly Pro Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp
385 390 395 400

Asn Thr Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg
405 410 415

Leu Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu
420 425 430

Phe Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly
435 440 445

Ser Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala
450 455 460

Thr Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val
465 470 475

<210> 156
<211> 1434
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 156
gaccctacca gactgcaggg catcaactgc ggccctgact tcactccttc tttcgcaaac 60
ctgggcagaa ccaccctgag caccagaggc ctcctagag gtggtcccgg cggagaactc 120
cccaggggtc ctgccggcct gggccctaga cgctcccagc aaggtcctag aaaggagcca 180
aggaagatca tcgccaccgt gctgatgacc gaggacatca agctgaacaa agctgagaag 240
gcctggaagc ctagcagcaa gagaaccgcc gccgacaagg acagaggcga ggaggacgcc 300
gacggatcca agaccagga cctgttcaga agagtgagaa gcatcctcaa caagctgacc 360
cctcagatgt tccagcagct gatgaagcag gtgacgcagc tcgccatcga caccgaggag 420

agactgaagg gcgtgatcga cctgatcttt gagaaggcta tctcagagcc taacttcagc	480
gtggcctacg ccaacatgtg ccgttgcctg atggcattga aggtgccaac caccgagaag	540
cctactgtga ccgtcaattt ccgtaaactg ctgctgaacc ggtgccagaa agagttcgag	600
aaggataagg acgacgacga ggtcttcgag aagaaacaga aagaaatgga cgaggccgcc	660
accgcagagg aaaggggccc attaaaggag gagctggagg aggccagaga catcgccaga	720
cggcgttctc tgggcaacat caaggcgata ggtgagctgt tcaagctaaa gatgctcacc	780
gaggccataa tgcacgactg cgtggtgaag ctactgaaga accacgacga agaaagcctg	840
gagtgcctgt gcagactgct gaccaccatc ggcaaggacc tggacttcga gaaggcaaag	900
cctcgaatgg accagtactt caaccagatg gagaagatta tcaaggagaa gaagaccagc	960
agcagaatca gattcatgct gcaggacgta ctggacctgc gcggaagcaa ctgggtgcc	1020
aggagagggg accaaggacc aaagaccatc gaccagatcc acaaggaagc ggagatggag	1080
gagcacagag agcacataaa ggtgcagcag cttatggcca agggcagcga caagcgaaga	1140
ggcggcccg ccggacctcc tatcagcaga ggccttcctc tggtagacga cggcggtgg	1200
aacaccgtgc ctatctctaa gggctccaga cctatcgaca cttcccgtct taccaagatc	1260
accaagccag gatctattga cagcaacaac cagctgttcg cccagaggag aagacttagc	1320
tggggcaagg gaagttccgg cggatccggc gccaaagcctt ccgacgccgc cagcgaggct	1380
gccagacctg ccaccagcac cttgaaccgc ttttcgctc tgcagcaagc tgtg	1434

<210> 157

<211> 9

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 157

ccrccaugg