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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 February 2017 (02.02.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/019935 A1

(51) International Patent Classification:
C12N 15/11 (2006.01) *A61K 8/11* (2006.01)
C12N 15/88 (2006.01) *A61K 8/14* (2006.01)
A61K 31/7105 (2006.01)

(21) International Application Number:
PCT/US2016/044638

(22) International Filing Date:
29 July 2016 (29.07.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/199,186 30 July 2015 (30.07.2015) US
62/352,801 21 June 2016 (21.06.2016) US

(71) Applicant: **MODERNATX, INC.** [US/US]; 320 Bent Street, Cambridge, MA 02141 (US).

(72) Inventors: **MOUSAVI, Kambiz**; 128 Parker Street, Unit 4B, Acton, MA 01720 (US). **MIHAI, Cosmin**; 110 Second Street, Cambridge, MA 02141 (US).

(74) Agent: **LOCKHART, Helen C.**; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, 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).

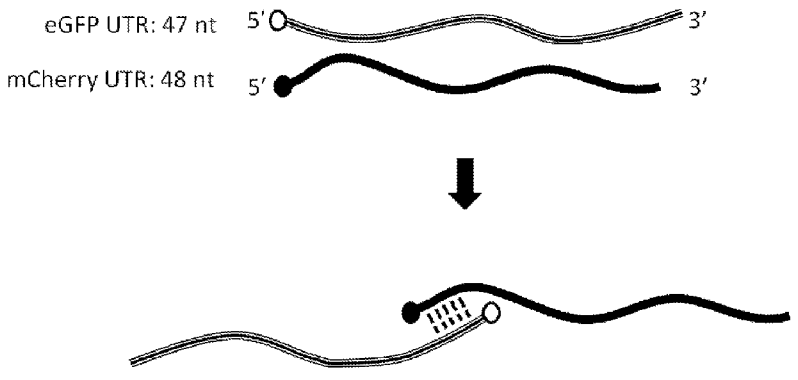
Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: MULTIMERIC MRNA

FIG. 4A



(57) Abstract: Aspects of the disclosure relate to multimeric molecules and methods of producing the same. In some embodiments, the multimeric molecules comprise at least two nucleic acid molecules (e.g., mRNA molecules) joined by non-covalent bonds between non-coding regions.

WO 2017/019935 A1

MULTIMERIC MRNA

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application serial number 61/199,186, filed July 30, 2015, and U.S. provisional application serial number 62/352,801, filed June 21, 2016, the entire contents of each of which are incorporated herein by reference.

BACKGROUND

Current mRNA therapy typically involves administration of single messenger RNAs (mRNAs). However, there are applications where multiple mRNAs must be administered for effective therapy. These applications include administration of protein complexes (*e.g.*, multimeric polypeptides such as antibodies or receptors) or multiple genes in cancer therapy. Due to the nature of the current formulation process, biopolymers (*e.g.*, multiple mRNAs) must be physically tethered for equal-molar LNP encapsulation, and release of biopolymers within subcellular compartments of target cells. Generally, biopolymers can be chemically adhered together through covalent bonds. Covalent bonds between biopolymers (*e.g.*, multiple mRNAs) can be achieved through chemical or enzymatic reactions. However, current methodology to establish covalent bonds between biopolymers (*e.g.*, multiple mRNAs) are limited as to number of biopolymers capable of being tethered, and insufficient insofar as reaching industrial scale. Moreover, covalent bonds between biopolymers (*e.g.*, multiple mRNAs) might exert unintended biological complications. Using current mRNA encapsulation processes, less than 50% of two different mRNAs are encapsulated in the same lipid nanoparticles (LNPs).

SUMMARY

A molecular design capable of equal distribution of multiple mRNAs within LNPs for uniform delivery and production of multiple polypeptides within subcellular space of targeted tissue or organ that solves these problems is described herein. The instant disclosure is based, in part, on the surprising discovery that formation of non-covalent (*e.g.*, electrostatic interactions) between the non-coding regions of mRNA molecules allows for the production of multimeric mRNA complexes that can be efficiently packaged and uniformly distributed in lipid nanoparticles (LNPs) and expressed in target cells.

Accordingly, in some aspects, the disclosure provides a lipid nanoparticle composition which includes a first nucleic acid and a second nucleic acid, wherein the first and second nucleic acids are uniformly distributed throughout the lipid nanoparticle and wherein the first and second nucleic acid molecules are not covalently linked to one another.

In some embodiments the first nucleic acid is linked to the second nucleic acid by a non-covalent bond. The non-covalent bond is located between a first non-coding region of the first nucleic acid and a second non-coding region of the second nucleic acid. In some embodiments the first nucleic acid and the second nucleic acid are RNA molecules. In other embodiments the RNA molecules are mRNA molecules. In yet other embodiments the mRNA molecules are *in vitro* transcribed mRNA molecules (IVT mRNA).

In some embodiments the first non-coding region and/or the second non-coding region is an untranslated region (UTR) and optionally a 5'UTR.

In other embodiments the first nucleic acid is linked to the second nucleic acid by 2, 3, 4, 5, or 6 non-covalent bonds. The first nucleic acid is linked to the second nucleic acid by at least 10 or at least 20 non-covalent bonds. Optionally, the non-covalent bonds are formed between complementary nucleotide bases of the first nucleic acid and the second nucleic acid. In some embodiments the complementary nucleotide bases of the first nucleic acid and the second nucleic acid have a G-C pairing ratio in a range of from about 30% to about 60%.

In other embodiments a third, fourth, fifth, sixth, seventh, eighth, ninth or tenth nucleic acid is provided. The third, fourth, fifth, sixth, seventh, eighth, ninth or tenth nucleic acid is linked to the first nucleic acid and/or the second nucleic acid by a non-covalent bond, wherein the non-covalent bond is located in a third, fourth, fifth, sixth, seventh, eighth, ninth or tenth non-coding region of the third, fourth, fifth, sixth, seventh, eighth, ninth or tenth nucleic acid.

The lipid nanoparticle in some embodiments comprises a cationic lipid, a PEG-modified lipid, a sterol and a non-cationic lipid. The cationic lipid is selected from the group consisting of 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319). The lipid nanoparticle in other embodiments has a molar ratio of about 20-60% cationic lipid: about 5-25% non-cationic lipid: about 25-55% sterol; and about 0.5-15% PEG-modified lipid. In some embodiments the lipid nanoparticle comprises a molar ratio of about 50% cationic lipid, about 1.5% PEG-

modified lipid, about 38.5% cholesterol and about 10% non-cationic lipid. The lipid nanoparticle has a mean diameter of 50-150 nm, or 80-100 nm in other embodiments.

The lipid nanoparticle may further include 4-100 additional nucleic acids, each having a different nucleic acid sequence and wherein each additional nucleic acid is linked to at least one other nucleic acid by a non-covalent bond.

In some embodiments each of the nucleic acids is an mRNA molecule encoding a different protein.

In yet other embodiments the first and second nucleic acids are non-covalently linked to one another through a splint. The splint may be an oligonucleotide having a region of complementarity with the first nucleic acid and a region of complementarity with the second nucleic acid.

A self-assembling multimeric mRNA structure is provided in other aspects of the invention. The structure comprises a first mRNA having a first linking region comprised of a part A and a part B and a second mRNA having a second linking region comprised of a part C and a part D, wherein at least part A of the first and at least part C of the second linking regions are complementary to one another.

In some embodiment the first linking region is in a non-coding region of the mRNA and/or the second linking region is in a non-coding region of the mRNA and optionally the non-coding region is a 5' untranslated region (UTR). In some embodiments the first and second linking regions are 5-100 nucleotides in length or 10-25 nucleotides in length.

The mRNA structure may also include a third, fourth, fifth, sixth, seventh, eighth, ninth or tenth mRNA, having another linking region complementary to other linking regions. In some embodiments mRNA structure further comprises 3-100 additional mRNAs, each mRNA having a linking region, wherein each linking region is complementary at least in part to at least one other linking region.

A self-assembling multimeric mRNA structure comprising 2-100 mRNAs each mRNA having a linking region and a stabilizing nucleic acid, wherein the stabilizing nucleic acid has a nucleotide sequence with regions complementary to each linking region. In some embodiments the stabilizing nucleic acid is an RNA or a DNA. In other embodiments the stabilizing nucleic acid has the following structure: $L_1X_1L_2X_2$ $L_3X_3L_4X_4L_5X_5L_6X_6$ wherein L is a nucleic acid sequence complementary to a linking region and wherein x is any nucleic acid sequence 0-50 nucleotides in length.

In other aspects the invention is a multimeric mRNA structure comprising a first mRNA and a second mRNA, wherein the first mRNA and the second mRNA are non-covalently linked to one another through a splint. In some embodiments the splint is an oligonucleotide having a first region of complementarity with the first nucleic acid and a second region of complementarity with the second nucleic acid. In other embodiments the first region of complementarity is located in the 5'UTR of the first nucleic acid and the second region of complementarity is located in the 5'UTR of the second nucleic acid. In yet other embodiments

each region of complementarity is at least 6 base pairs long.

Each of the limitations of the invention can encompass various embodiments of the invention. It is, therefore, anticipated that each of the limitations of the invention involving any one element or combinations of elements can be included in each aspect of the invention. This invention is not limited in its application to the details of construction and the arrangement of components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments and of being practiced or of being carried out in various ways.

BRIEF DESCRIPTION OF DRAWINGS

The accompanying drawings are not intended to be drawn to scale. In the drawings, each identical or nearly identical component that is illustrated in various figures is represented by a like numeral. For purposes of clarity, not every component may be labeled in every drawing. Note that in Figures 16 to 18 and elsewhere in this document, "XXXXXX" and "NNNNNN" are used interchangeably to denote a nucleic acid sequence of variable length. In the drawings:

Fig. 1 shows a schematic depiction of conventional lipid nanoparticle (LNP) formulation of nucleic acid molecules. Equimolar amounts of two nucleic acids (*e.g.*, a first mRNA labeled with Alexa 488 and a second mRNA labeled with Alexa 647 are formulated with a fluorescently-labeled lipid nanoparticle (*e.g.*, MC3, which is labeled with Rhodamine-DOPE). After encapsulation, four populations of LNP are present: empty LNPs, LNPs containing the first mRNA, LNPs containing the second mRNA, and LNPs containing both the first and second mRNAs.

Fig. 2 shows a structured illumination microscopy (SIM) photo of mRNA-loaded LNPs. Rhodamine-MC3 LNP was loaded with Alexa 488-mCherry mRNA and Alexa 647-

mCherry mRNA and imaged. Results of SIM indicate that 47% of LNPs are positive for both mRNA species.

Figs. 3A-3B show one embodiment of a splint-assisted multimeric mRNA molecule. Fig 3A shows a schematic representation of one embodiment of a splint-assisted multimeric mRNA molecule; two mRNA molecules are tethered via a short nucleic acid "splint" that hybridizes to the 5'-end of each mRNA molecule. Fig. 3B shows gel electrophoresis data confirming the formation of a splint-assisted multimeric mRNA comprising eGFP and mCherry.

Figs. 4A-4B show one embodiment of a self-assembling multimeric mRNA molecule. Fig. 4A depicts formation of non-covalent bonds between the 5' untranslated regions (UTRs) of an eGFP mRNA molecule having a 47-nucleotide long 5'UTR (untranslated region) and a mCherry mRNA molecule having a 48-nucleotide long 5'UTR. Fig. 4B shows representative bioanalyzer gel image data confirming the formation of the self-assembling multimeric mRNA molecule shown in Fig. 4A.

Fig. 5 shows self-assembled multimeric mRNA co-translation in JAWSII monocyte cells. Cells were transfected with self-assembling multimeric mRNA comprising an eGFP mRNA and a mCherry mRNA and imaged. Overlap of eGFP and mCherry signals indicates co-translation of mRNAs (arrows).

Fig. 6 shows multimeric mRNA co-translation in HeLa cells. Cells in the left panel were transfected with monomeric mRNAs encoding mCherry and eGFP (No Multiplex). Cells in the right panel were transfected with self-assembling multimeric mRNA comprising an mCherry mRNA and an eGFP mRNA (Multiplex). Cells in the right panel show much higher efficiency of co-translation than cells in the left panel.

Fig. 7 shows data related to multi-mRNA translational kinetics in HeLa cells. Lack of change in eGFP fluorescence between monomeric mRNA and multimeric mRNA indicates that formation of multimeric mRNA complex does not interfere with protein expression.

Fig. 8 shows additional data related to multi-mRNA translational kinetics in HeLa cells. Lack of change in eGFP fluorescence between monomeric mRNA and multimeric mRNA indicates that formation of multimeric mRNA complex does not interfere with protein expression.

Fig. 9 shows data related to multi-mRNA translational kinetics in HeLa cells. Lack of change in eGFP fluorescence between monomeric mRNA and multimeric mRNA across low,

medium and high doses of mRNA indicates that formation of multimeric mRNA complex does not interfere with protein expression.

Figs. 10A-10C show multiplex mRNA analysis by FACS. Fig. 10A shows a graphic depiction of a self-assembling multiplex mRNA molecule comprising eGFP and BFP. Fig. 10B shows FACS data demonstrating that delivery of multimeric mRNA molecules does not negatively affect cell viability. Fig. 10C shows FACS data related to mRNA co-translation after 24 hours. All multiplex mRNA-transfected cells are positive for both eGFP and BFP, indicating high levels of co-translation.

Fig. 11 shows additional data related to mRNA co-translation after 24 hours. Lack of change in eGFP fluorescence between monomeric mRNA and multimeric mRNA indicates that formation of the multimeric mRNA complex does not interfere with protein expression.

Figs. 12A-12B show a self-assembling multimeric mRNA comprising three mRNA sequences. Fig. 12A shows a graphic representation of a trimeric molecule comprising mTagBFP mRNA, eGFP mRNA and mCherry mRNA. Non-covalent bonds are formed between the 5'UTRs of each mRNA molecule. Fig. 12B shows bioanalyzer gel image confirming formation of the self-assembling trimeric mRNA molecule shown in Fig. 12A.

Fig. 13 shows fluorescence microscopy data demonstrating co-translation of all three mRNAs (mTagBFP, eGFP, and mCherry) from the trimeric molecule (far right column).

Figs. 14A-14B show a self-assembling multimeric mRNA comprising four mRNA sequences. Fig. 14A shows a graphic representation of a tetrameric molecule comprising mTagBFP mRNA, eGFP mRNA, mCherry mRNA, and nanoLuc mRNA. Non-covalent bonds are formed between the 5'UTRs of each mRNA molecule. Fig. 14B shows bioanalyzer gel confirming formation of the self-assembling tetrameric mRNA molecule shown in Fig. 14A.

Fig. 15 shows a SIM photo of a self-assembling dimeric mRNA molecule formulated in a LNP. An mRNA comprising Alexa 488 was tethered to an mRNA comprising Alexa 647 to form a self-assembling dimeric mRNA molecule. The dimeric molecule was formulated into a LNP and imaged. The overlay image (right) demonstrates co-localization of the two mRNAs in the same LNP at a higher efficiency than LNP loaded using the conventional method depicted in Fig. 1.

Fig. 16 shows nucleic acid sequences of mRNA #1 (SEQ ID NO: 1) and mRNA #2 (SEQ ID NO: 2), which form a dimer by non-covalent bonding between their respective

5'UTRs (**bold**). The coding sequence of each mRNA sequence is denoted by "NNNNNN" and each 3'UTRs is underlined.

Fig. 17 shows nucleic acid sequences of mRNA #1 (SEQ ID NO: 1), mRNA #2 (SEQ ID NO: 2), and mRNA #3 (SEQ ID NO: 3) which form a trimer by non-covalent bonding between their respective 5'UTRs (**bold**). The coding sequence of each mRNA sequence is denoted by "NNNNNN" and each 3'UTRs is underlined.

Fig. 18 shows nucleic acid sequences of mRNA #1 (SEQ ID NO: 1), mRNA #2 (SEQ ID NO: 2), mRNA #3 (SEQ ID NO: 3), and mRNA #4 (SEQ ID NO: 4) which form a tetramer by non-covalent bonding between their respective 5'UTRs (**bold**). The coding sequence of each mRNA sequence is denoted by "NNNNNN" and each 3'UTR is underlined.

Fig. 19 shows a bioanalyzer gel confirming formation of the self-assembling dimer multi-mRNA complex comprising eGFP and mCherry.

Fig. 20 shows microscopy data for GFP IHC on mouse liver embedded in paraffin. The arrows indicate the GFP signal localized to Kupffer cells. Positive staining is also shown in the cytoplasm of hepatocytes in the right panel. Samples were taken 6 hours after intravenous (IV) administration of PBS (left panel) or GFP (right panel). The GFP was diluted 1:1500 and a single dose of 2 mg/kg was administered.

Fig. 21 shows microscopy data for mCherry IF on mouse liver embedded in paraffin. The left panel shows background staining in the control due to endogenous mouse immunoglobulin (Ig) staining from mouse primary antibodies used on mouse tissues. The right panel shows high mCherry protein expression in hepatocytes, 6 hours after IV administration of 2 mg/kg mCherry diluted 1:800.

Figs. 22A-22B show microscopy data for GFP (Fig. 22A) and mCherry/GFP (Fig. 22B).

Fig. 23 shows a graphical representation of IHC mCherry protein expression 6 hours after IV administration of a single 2 mg/kg dose of the indicated compound in mouse liver. Dunnett's multiple comparisons tests were used to determine the statistics of $p < 0.01$ vs. vehicle (**) and $p < 0.0001$ vs. vehicle (****).

Fig. 24 shows a graphical representation of IHC GFP protein expression 6 hours after IV administration of a single 2 mg/kg dose of the indicated compound in mouse liver. Dunnett's multiple comparisons tests were used to determine the statistics of $p < 0.05$ vs. vehicle (*) and $p < 0.01$ vs. vehicle (**).

Fig. 25 shows a graphical representation of IHC dual mCherry and GFP protein expression 6 hours after IV administration of a single 2 mg/kg dose of the indicated compound in mouse liver. Dunnett's multiple comparisons tests were used to determine the statistics of $p < 0.05$ vs. vehicle (*) and $p < 0.0001$ vs. vehicle (***).

Figs. 26A-26E show microscopy data from the mCherry and GFP dual IHC assay from three different mice. Fig. 26A shows staining of the PBS control. Fig. 26B shows staining after a 2mg/kg dose of GFP mRNA. Fig. 26C shows staining after a 2mg/kg dose of mCherry mRNA. Fig. 26D shows staining after a 2 mg/kg dose of non-dimerized eGFP and mCherry. Fig. 26E shows staining after a 2 mg/kg dose of dimerized eGFP and mCherry.

Figs. 27A-27E show microscopy data from the GFP singleplex IHC assay from three different mice. Fig. 27A shows staining of the PBS control. Fig. 27B shows staining after a 2mg/kg dose of GFP mRNA. Fig. 27C shows staining after a 2mg/kg dose of mCherry mRNA. Note the non-specific staining in Fig. 27C. Fig. 27D shows staining after a 2 mg/kg dose of non-dimerized eGFP and mCherry. Fig. 27E shows staining after a 2 mg/kg dose of dimerized eGFP and mCherry.

Figs. 28A-28C show microscopy data from the mCherry singleplex IHC assay from three different mice. Fig. 28A shows staining of the PBS control. Fig. 28B shows staining after a 2mg/kg dose of GFP mRNA. Fig. 28C shows staining after a 2mg/kg dose of mCherry mRNA.

Fig. 29 shows microscopy data for GFP IF on mouse liver embedded in paraffin. The green channel (auto fluorescence) and Tritc channel (GFP) are overlaid. The arrows on the right panel indicate GFP signal localized to Kupffer cells. Samples were taken 6 hours after IV administration of PBS (left panel) or GFP (right panel). The GFP was diluted 1:100 and a single dose of 2 mg/kg was administered.

Fig. 30 shows microscopy data for the Fitc channel of mCherry IF on mouse liver embedded in paraffin. The left panel shows background staining in the control due to endogenous mouse Ig staining from mouse primary antibodies used on mouse tissues. The right panel shows high mCherry protein expression in hepatocytes. Samples were taken 6 hours after IV administration of PBS (left panel) or mCherry (right panel). The mCherry was diluted 1:50 and a single dose of 2 mg/kg was administered.

DETAILED DESCRIPTION

Some challenges exist for mRNA therapy wherein multiple mRNAs must to be administered for effective therapy, for example administration of protein complexes (*e.g.*, multimeric polypeptides such as antibodies or receptors) or multiple genes in cancer therapy. Current encapsulation processes use monomeric mRNAs, which result in random encapsulation of different ratios of mRNAs in lipid nanoparticles (LNPs). This presents several challenges from both manufacturing and clinical perspectives. For example, current formulation methodology is limited as to number of biopolymers (*e.g.*, multiple mRNAs) capable of being tethered. Encapsulation efficiency for multiple biopolymers is also low and therefore insufficient for industrial scale-up. Furthermore, currently used strategies that rely on covalent bonds between biopolymers (*e.g.,* multiple mRNAs) might exert unintended biological complications. Accordingly, the discoveries described herein provide novel compositions for the delivery of multiplex biopolymers, such as multiple mRNAs and overcome prior art issues.

The instant invention is based, in part, on the surprising discovery that formation of multimeric complexes based on non-covalent (*e.g.*, hydrogen bonds) linkages between mRNA molecules allows for uniform distribution of the mRNA in a therapeutic composition. When multiple nucleic acids such as RNA are formulated, for instance, in a lipid based formulation, a relatively uniform distribution of the total nucleic acid through the formulation may be achieved. However, the distribution of a particular nucleic acid with respect to the other nucleic acids in the mixture is not uniform. For instance when the nucleic acid mixture is composed of two distinct mRNA sequences, some of the lipid particles or other formulatory agents will house a single mRNA sequence, while others will house the other mRNA sequence and a few will house both of the mRNA sequences. In a therapeutic context this uneven distribution of mRNA is undesirable because the dosage of the mRNA being delivered to a patient will vary from administration to administration. Quite surprisingly, the methods of the invention have enabled the production of formulations having multiple nucleic acids wherein the nucleic acid has a uniform distribution throughout the formulation. The methods are achieved through the use of a non-covalent interaction. It was surprising that a non-covalent interaction between the individual nucleic acids would be capable of producing such a uniform distribution of the nucleic acids in a formulation.

It was also discovered according to aspects of the invention that the multimeric nucleic acid complexes generated according to the invention did not interfere with activity

such as mRNA expression activity. It was quite surprising that mRNA formed into multimeric complexes did not experience a loss of expression activity as a result of the structures.

Described herein are compositions (including pharmaceutical compositions) and methods for the delivery of multimeric nucleic acid molecules. In some embodiments the multimeric structures are uniformly distributed throughout a composition such as a lipid nanoparticle. Uniformly distributed, as used herein in the context of multiple nucleic acids (each having a unique nucleotide sequence), refers to the distribution of each of the nucleic acids relative to one another in the formulation. Distribution of the nucleic acids in a formulation may be assessed using methods known in the art. For instance, several exemplary methods are shown in the Examples below. A nucleic acid is uniformly distributed relative to another nucleic acid if the nucleic acid is associated in proximity within a particular area of the formulation to the other nucleic acid at an approximately 1:1 ratio. In some embodiments the nucleic acid is uniformly distributed relative to another nucleic acid if the nucleic acid is positioned within a particular area of the formulation to the other nucleic acid at an approximately 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, or 1:2 ratio.

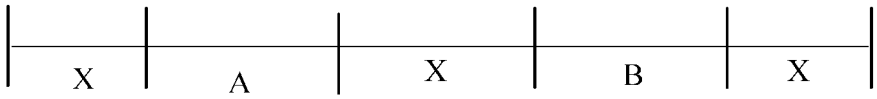
The multimeric structures of the invention are comprised of nucleic acid molecules, specifically polynucleotides which, in some embodiments, encode one or more peptides or polypeptides of interest. The term "nucleic acid," in its broadest sense, includes any compound and/or substance that comprise a polymer of nucleotides. These polymers are often referred to as polynucleotides.

A multimeric structure as used herein is series of at least nucleic acids linked together to form a multimeric structure. In some embodiments a multimeric structure is composed of 3 or more, 4 or more, 5 or more 6 or more 7 or more, 8 or more, 9 or more nucleic acids. In other embodiments the multimeric structure is composed of 1000 or less, 900 or less, 500 or less, 100 or less, 75 or less, 50 or less, 40 or less, 30 or less, 20 or less or 100 or less nucleic acids. In yet other embodiments a multimeric structure has 3-100, 5-100, 10-100, 15-100, 20-100, 25-100, 30-100, 35-100, 40-100, 45-100, 50-100, 55-100, 60-100, 65-100, 70-100, 75-100, 80-100, 90-100, 5-50, 10-50, 15-50, 20-50, 25-50, 30-50, 35-50, 40-50, 45-50, 100-150, 100-200, 100-300, 100-400, 100-500, 50-500, 50-800, 50-1,000, or 100-1,000 nucleic acids. In preferred embodiments a multimeric structure is composed of 3-5 nucleic acids.

In some embodiments the upper limit on the number of nucleic acids in a multimeric structure depends on the length of dimerizable region. A greater than 20-nucleotide space

between mRNAs can provide specificity and enough force to keep the multi-mRNA complex intact for downstream processing and is thus preferred in some embodiments. In some embodiments 4-5 nucleic acids in a multimeric structure may be desirable. For instance, cell conversion/differentiation (*e.g.*, Induced Pluripotent Stem Cells-iPS) may be achieved with four protein factors. A similar number of proteins may be effective for inhibition of tumor growth.

The multimeric structures may be self-assembling multimeric mRNA structures composed of a first mRNA having a first linking region comprised of a part A and a part B and a second mRNA having a second linking region comprised of a part C and a part D, wherein at least part A of the first and at least part C of the second linking regions are complementary to one another. Preferably the nucleic acids are linked to one another through a non-covalent bond in the linking regions. The following is an exemplary linking region, wherein X is any nucleic acid sequence of 0-100 nucleotides and A and B are complementary parts, which are complementary to one or more other nucleic acids.



A linking region, as used herein, refers to a nucleic acid sequence having one or more regions or parts that are complementary to one or more regions of other linking regions. A pair of linking regions, each having one complementary region, may be at least 70% complementary to one another. In some embodiments a pair of linking regions are at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% complementary to one another. A linking region may be composed of sub-parts, optionally referred to as parts A, B, C, D, …, which have shorter regions of complementarity between one another, such that the subparts may be complementary with other sub-parts. For instance, Figure 4 depicts a simple multimeric structure of two mRNAs, each having a linking region with a single region of complementarity. The two linking regions are able to form non-covalent interactions with one another through base pairing. Figures 12 and 14 each depict a more complex multimeric structure wherein a linking region of each nucleic acid has at least two parts, each part having complementarity with a part on another nucleic acid linking region. Linking regions having multiple parts with different complementarity enables the production of larger multimeric complexes of 3, 4, 5 or more nucleic acids.

The linking regions in some embodiments are 5-100 nucleotides in length. In other embodiments the linking regions are 10-25 nucleotides in length.

As used herein, the term "region of complementarity" refers to a region on a first nucleic acid strand that is substantially complementary to a second region on a second nucleic acid strand. Generally, two nucleic acids sharing a region of complementarity are capable, under suitable conditions, of hybridizing (*e.g.*, via nucleic acid base pairing) to form a duplex structure. A region of complementarity can vary in size. In some embodiments, a region of complementarity ranges in length from about 2 base pairs to about 100 base pairs. In some embodiments, a region of complementarity ranges in length from about 5 base pairs to about 75 base pairs. In some embodiments, a region of complementarity ranges in length from about 10 base pairs to about 50 base pairs. In some embodiments, a region of complementarity ranges in length from about 20 base pairs to about 30 base pairs.

The number of nucleic acid bases shared between two nucleic acids across a region of complementarity can vary. In some embodiments, two nucleic acids share 100% complementary base pairs (*e.g.*, no mismatches) across a region of complementarity. In some embodiments, two nucleic acids share at least 99.9%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% complementary base pairs across a region of complementarity. In some embodiments, a region of complementarity shared between two nucleic acids includes at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 base pair mismatches. In some embodiments, a region of complementarity shared between two nucleic acids includes more than 10 base pair mismatches.

As used herein, the term "non-covalent bond" refers to a chemical interaction (*e.g.*, joining) between molecules that does not involve the sharing of electrons. Generally, non-covalent bonds are formed via electromagnetic interactions between charged molecules. Examples of non-covalent bonds include, but are not limited to, ionic bonds, hydrogen bonds, halogen bonds, Van der Waals forces (*e.g.*, dipole-dipole interactions, London dispersion forces, *etc.*), π-effects (π-π interactions, cation-π interactions, anion-π interactions), and hydrophobic effect.

In some embodiments, at least one non-covalent bond formed between the nucleic acid molecules (*e.g.*, mRNA molecules) of a multimeric molecule is a result of Watson-Crick base-pairing. The term "Watson-Crick base-pairing", or "base-pairing" refers to the formation of hydrogen bonds between specific pairs of nucleotide bases ("complementary

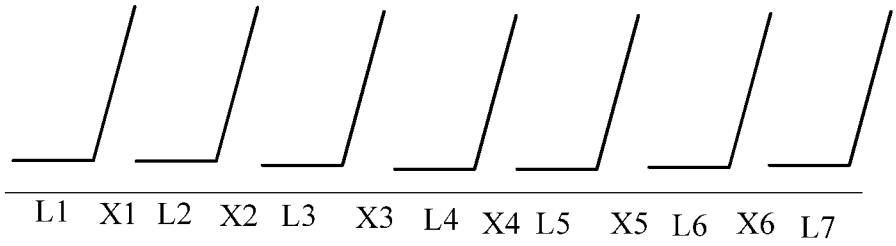
base pairs"). For example, two hydrogen bonds form between adenine (A) and uracil (U), and three hydrogen bonds form between guanine (G) and cytosine (C). One method of assessing the strength of bonding between two polynucleotides is by quantifying the percentage of bonds formed between the guanine and cytosine bases of the two polynucleotides ("GC content"). In some embodiments, the GC content of bonding between two nucleic acids of a multimeric molecule (*e.g.*, a multimeric mRNA molecule) is at least 10%, at least 20%, at least 30%, at least 40%, or at least 50%. In some embodiments, the GC content of bonding between two nucleic acids of a multimeric molecule (*e.g.*, a multimeric mRNA molecule) is between 10% and 70%, about 20% to about 60%, or about 30% to about 60%. The formation of a nucleic acid duplex via bonding of complementary base pairs can also be referred to as "hybridization".

In some embodiments, two nucleic acid molecules (*e.g.*, mRNA molecules) hybridize to form a multimeric molecule. Hybridization can result from the formation of at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 non-covalent bonds between two polynucleotides (*e.g.*, mRNA molecules). In some embodiments, between about 2 non-covalent bonds and about 10 non-covalent bonds are formed between two nucleic acid molecules. In some embodiments, between about 5 and about 15 non-covalent bonds are formed between two nucleic acid molecules. In some embodiments, between about 10 and about 20 non-covalent bonds are formed between two nucleic acid molecules. In some embodiments, between about 15 and about 30 non-covalent bonds are formed between two nucleic acid molecules. In some embodiments, between about 20 and about 50 non-covalent bonds are formed between two nucleic acid molecules. In some embodiments, the number of non-covalent bonds formed between two nucleic acid molecules (*e.g.*, mRNA molecules) is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 non-covalent bonds.

In some embodiments the self-assembling multimeric mRNA structure is comprised of at least 2-100 mRNAs each mRNA having a linking region and a stabilizing nucleic acid, wherein the stabilizing nucleic acid has a nucleotide sequence with regions complementary to each linking region. A stabilizing nucleic acid as used herein is any nucleic acid that has multiple linking regions and is capable of forming non-covalent interactions with at least 2, but more preferably, 3, 4, 5, 6,7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49

or 50 other nucleic acids. For instance the stabilizing nucleic acid may have the following structure:

$L_1X_1L_2X_2$ $L_3X_3L_4X_4L_5X_5L_6X_6$ wherein L is a nucleic acid sequence complementary to a linking region and wherein x is any nucleic acid sequence 0-50 nucleotides in length. Such a structure may look like the following:



Exemplary nucleic acids or polynucleotides of the invention include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), 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), ethylene nucleic acids (ENA), cyclohexenyl nucleic acids (CeNA) or hybrids or combinations thereof.

In some aspects, the disclosure provides a multimeric molecule comprising at least two nucleic acid molecules, wherein a first nucleic acid molecule is joined to a second nucleic acid molecule by at least one non-covalent bond, and wherein the at least one covalent bond is located between a first non-coding region of the first nucleic acid molecule and a second non-coding region of the second nucleic acid molecule.

In addition to having at least two distinct nucleic acids with unique sequences, the multimeric molecules may comprise multiple copies of the same gene or protein (*e.g.*, 2, 3, 4, 5, or more mRNA encoding the same protein), as long as it includes at least two distinct nucleic acids. This type of multimeric molecule may be useful for increasing expression level of a particular protein in a cell. Multimeric molecules can also comprise nucleic acids (*e.g.*, mRNA) encoding different gene or protein (*e.g.*, 4 mRNA molecules, wherein each mRNA molecule encodes a different subunit protein of tetrameric receptor). Multimeric molecules comprising nucleic acids encoding different genes or proteins may also be useful for delivering combination biological therapies, for example in the context of cancer chemotherapy.

In some embodiments, a multimeric mRNA molecule comprises a first mRNA and a second mRNA, wherein the first mRNA and the second mRNA are non-covalently linked to one another through a splint. As used herein, the term "splint" refers to an oligonucleotide having a first region of complementarity with the first nucleic acid and a second region of complementarity with the second nucleic acid. A splint can be a DNA oligonucleotide or an RNA oligonucleotide. In some embodiments, a splint comprises one or more modified oligonucleotides. In some embodiments, a splint is non-covalently linked to a 5'UTR of an mRNA. In some embodiments, a splint is non-covalently linked to a 3'UTR of an mRNA.

In some embodiments, non-covalent bonds between nucleic acid molecules (*e.g.*, mRNA molecules) are formed in a non-coding region of each molecule. As used herein, the term "non-coding region" refers to a location of a polynucleotide (*e.g.*, an mRNA) that is not translated into a protein. Examples of non-coding regions include regulatory regions (*e.g.*, DNA binding domains, promoter sequences, enhancer sequences), and untranslated regions (*e.g.*, 5'UTR, 3'UTR). In some embodiments, the non-coding region is an untranslated region (UTR).

By definition, wild type untranslated regions (UTRs) of a gene are transcribed but not translated. In mRNA, the 5′UTR starts at the transcription start site and continues to the start codon but does not include the start codon; whereas, the 3′UTR starts immediately following the stop codon and continues until the transcriptional termination signal.

Natural 5′UTRs bear features which play roles in translation initiation. They harbor signatures like Kozak sequences which 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, where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G'. 5′UTR also have been known to form secondary structures which 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 the polynucleotides of the invention. 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, could be used to enhance expression of a nucleic acid molecule, such as a 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 (MyoD, Myosin, Myoglobin, Myogenin, Herculin), for endothelial cells (Tie-1, CD36), for

myeloid cells (C/EBP, AML1, G-CSF, GM-CSF, CD11b, MSR, Fr-1, i-NOS), for leukocytes (CD45, CD18), for adipose tissue (CD36, GLUT4, ACRP30, adiponectin) and for lung epithelial cells (SP-A/B/C/D).

Other non-UTR sequences may also be used as regions or subregions within the polynucleotides. For example, introns or portions of introns sequences may be incorporated into regions of the polynucleotides of the invention. Incorporation of intronic sequences may increase protein production as well as polynucleotide levels.

Combinations of features may be included in flanking regions and may be contained within other features. For example, the ORF may be flanked by a 5' UTR which may contain a strong Kozak translational initiation signal and/or a 3' UTR which may include an oligo(dT) sequence for templated addition of a poly-A tail. 5'UTR may comprise a first polynucleotide fragment and a second polynucleotide fragment from the same and/or different genes.

It should be understood that any UTR from any gene may be incorporated into the regions of the polynucleotide. Furthermore, multiple wild-type UTRs of any known gene may be utilized. It is also within the scope of the present invention to provide artificial UTRs which are not variants of wild type regions. These UTRs or portions thereof may be placed in the same orientation as in the transcript from which they were selected or may be altered in orientation or location. Hence a 5′ or 3′ UTR may be inverted, shortened, lengthened, made with one or more other 5′ UTRs or 3′ UTRs. As used herein, the term "altered" as it relates to a UTR sequence, means that the UTR has been changed in some way in relation to a reference sequence. For example, a 3′ or 5′ UTR may be altered relative to a wild type or native UTR by the change in orientation or location as taught above or may be altered by the inclusion of additional nucleotides, deletion of nucleotides, swapping or transposition of nucleotides. Any of these changes producing an "altered" UTR (whether 3′ or 5′) comprise a variant UTR.

In one embodiment, a double, triple or quadruple UTR such as a 5′ or 3′ UTR may be used. As used herein, a "double" UTR is one in which two copies of the same UTR are encoded either in series or substantially in series.

It is also within the scope of the present invention to have patterned UTRs. As used herein "patterned UTRs" are those UTRs which reflect a repeating or alternating pattern, such as ABABAB or AABBAABBAABB or ABCABCABC or variants thereof repeated once, twice, or more than 3 times. In these patterns, each letter, A, B, or C represent a different UTR at the nucleotide level.

In one embodiment, flanking regions are selected from a family of transcripts whose proteins share a common function, structure, feature of property. For example, polypeptides of interest may belong to a family of proteins which are expressed in a particular cell, tissue or at some time during development. The UTRs from any of these genes may be swapped for any other UTR of the same or different family of proteins to create a new polynucleotide. As used herein, a "family of proteins" is used in the broadest sense to refer to a group of two or more polypeptides of interest which share at least one function, structure, feature, localization, origin, or expression pattern. The untranslated region may also include translation enhancer elements (TEE).

In some embodiments, an UTR of a polynucleotide (*e.g.*, a first nucleic acid) of the present invention is engineered or modified to have regions of complementarity with an UTR of another polynucleotide (a second nucleic acid). For example, UTR nucleotide sequences of two polynucleotides sought to be joined (*e.g.*, in a multimeric molecule) can be modified to include a region of complementarity such that the two UTRs hybridize to form a multimeric molecule.

In some embodiments, multimeric nucleic acid molecules comprise RNA molecules. In some embodiments, the RNA molecules are mRNA molecules. As used herein, the term "messenger RNA" (mRNA) refers to any polynucleotide which encodes at least one peptide or polypeptide of interest and which is capable of being translated to produce the encoded peptide polypeptide of interest in vitro, in vivo, in situ or ex vivo. An mRNA has been transcribed from a DNA sequence by an RNA polymerase enzyme, and interacts with a ribosome synthesize genetic information encoded by DNA. Generally, mRNA are classified into two sub-classes: pre-mRNA and mature mRNA. Precursor mRNA (pre-mRNA) is mRNA that has been transcribed by RNA polymerase but has not undergone any post-transcriptional processing (*e.g.*, 5'capping, splicing, editing, and polyadenylation). Mature mRNA has been modified via post-transcriptional processing (*e.g.*, spliced to remove introns and polyadenylated) and is capable of interacting with ribosomes to perform protein synthesis. mRNA can be isolated from tissues or cells by a variety of methods. For example, a total RNA extraction can be performed on cells or a cell lysate and the resulting extracted total RNA can be purified (*e.g.*, on a column comprising oligo-dT beads) to obtain extracted mRNA.

Alternatively, mRNA can be synthesized in a cell-free environment, for example by *in vitro* transcription (IVT). An "*in vitro* transcription template" as used herein, refers to

deoxyribonucleic acid (DNA) suitable for use in an IVT reaction for the production of messenger RNA (mRNA). In some embodiments, an IVT template encodes a 5′ untranslated region, contains an open reading frame, and encodes a 3′ untranslated region and a polyA tail. The particular nucleotide sequence composition and length of an IVT template will depend on the mRNA of interest encoded by the template.

A "5′ untranslated region (UTR)" refers to a region of an mRNA that is directly upstream (*i.e.*, 5′) from the start codon (*i.e.*, the first codon of an mRNA transcript translated by a ribosome) that does not encode a protein or peptide.

A "3′ untranslated region (UTR)" refers to a region of an mRNA that is directly downstream (*i.e.*, 3′) from the stop codon (*i.e.*, the codon of an mRNA transcript that signals a termination of translation) that does not encode a protein or peptide.

An "open reading frame" is a continuous stretch of DNA beginning with a start codon (*e.g.*, methionine (ATG)), and ending with a stop codon (*e.g.*, TAA, TAG or TGA) and encodes a protein or peptide.

A "polyA tail" is a region of mRNA that is downstream, *e.g.*, directly downstream (*i.e.*, 3'), from the 3' UTR that contains multiple, consecutive adenosine monophosphates. A polyA tail may contain 10 to 300 adenosine monophosphates. For example, a polyA tail may contain 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290 or 300 adenosine monophosphates. In some embodiments, a polyA tail contains 50 to 250 adenosine monophosphates. In a relevant biological setting (*e.g.*, in cells, *in vivo*, *etc.*) the poly(A) tail functions to protect mRNA from enzymatic degradation, *e.g.*, in the cytoplasm, and aids in transcription termination, export of the mRNA from the nucleus, and translation.

Thus, the polynucleotide may in some embodiments comprise (a) a first region of linked nucleosides encoding a polypeptide of interest; (b) a first terminal region located 5' relative to said first region comprising a 5' untranslated region (UTR); (c) a second terminal region located 3' relative to said first region; and (d) a tailing region. The terms poly nucleotide and nucleic acid are used interchangeably herein. For example, SEQ ID NO: 1-4 in FIG. 18 each comprise a first region of linked nucleosides encoding a polypeptide of interest represented by "NNNNNN". The skilled artisan recognizes that the six-mer "NNNNNN" is representative of polypeptide encoding sequence(s) of varying length. In some embodiments, the first region of linked nucleosides (*e.g.*, polypeptide encoding sequence) ranges from about 30 to about 3,000 nucleotides in length. In some embodiments,

the first region of linked nucleosides (*e.g.*, polypeptide encoding sequence) ranges from about 200 to about 3,000 nucleotides in length.

In some embodiments, the polynucleotide includes from about 30 to about 300 nucleotides (*e.g.*, from about 30 to about 50, from about 40 to about 60, from about 50 to about 100, from about 75 to about 150, from about 125 to about 200, from about 175 to about 250, from about 225 to about 300). In some embodiments, the polynucleotide includes from about 200 to about 3,000 nucleotides (*e.g.*, from 200 to 500, from 200 to 1,000, from 200 to 1,500, from 200 to 3,000, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to 3,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 3,000, from 1,500 to 3,000, and from 2,000 to 3,000).

IVT mRNA may function as mRNA but are distinguished from wild-type mRNA in their functional and/or structural design features which serve to overcome existing problems of effective polypeptide production using nucleic-acid based therapeutics. For example, IVT mRNA may be structurally modified or chemically modified. As used herein, a "structural" modification is one in which two or more linked nucleosides are inserted, deleted, duplicated, inverted or randomized in a polynucleotide without significant chemical modification to the nucleotides themselves. Because chemical bonds will necessarily be broken and reformed to effect a structural modification, structural modifications are of a chemical nature and hence are chemical modifications. However, structural modifications will result in a different sequence of nucleotides. For example, the polynucleotide "ATCG" may be chemically modified to "AT-5meC-G". The same polynucleotide may be structurally modified from "ATCG" to "ATCCCG". Here, the dinucleotide "CC" has been inserted, resulting in a structural modification to the polynucleotide.

cDNA encoding the polynucleotides described herein may be transcribed using an in vitro transcription (IVT) system. The system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and a polymerase. The NTPs may be manufactured in house, may be selected from a supplier, or may be synthesized as described herein. The NTPs may be selected from, but are not limited to, those described herein including natural and unnatural (modified) NTPs. The polymerase may be selected from, but is not limited to, T7 RNA polymerase, T3 RNA polymerase and mutant polymerases such as, but not limited to, polymerases able to incorporate polynucleotides (*e.g.*, modified nucleic acids).

In some aspects, the disclosure provides a method of producing a multimeric mRNA complex. In some embodiments, a multimeric mRNA complex is formed by a heating and stepwise cooling protocol. For example, a mixture of 5 μM of each mRNA desired to be incorporated into the multimeric complex can be placed in a buffer containing 50 mM 2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris) pH 7.5, 150 mM sodium chloride (NaCl), and 1 mM ethylene-diamine-tetra-acetic acid (EDTA). The mixture can be heated to 65°C for 5 minutes, 60°C for 5 minutes, 40°C for 2 minutes, and then cooled to 4°C for 10 minutes, resulting in the formation of a multimeric complex.

Thus, in an exemplary aspect, polynucleotides of the invention may include at least one chemical modification. The polynucleotides can include various substitutions and/or insertions from native or naturally occurring polynucleotides. As used herein in a polynucleotide, the terms "chemical modification" or, as appropriate, "chemically modified" refer to modification with respect to adenosine (A), guanosine (G), uridine (U), thymidine (T) or cytidine (C) ribo- or deoxyribnucleosides in one or more of their position, pattern, percent or population. Generally, herein, these terms are not intended to refer to the ribonucleotide modifications in naturally occurring 5′-terminal mRNA cap moieties.

The modifications may be various distinct modifications. In some embodiments, the regions may contain one, two, or more (optionally different) nucleoside or nucleotide modifications. In some embodiments, a modified polynucleotide, introduced to a cell may exhibit reduced degradation in the cell, as compared to an unmodified polynucleotide.

Modifications of the polynucleotides of the multimeric structures include, but are not limited to those listed in detail below. The polynucleotide may comprise modifications which are naturally occurring, non-naturally occurring or the polynucleotide can comprise both naturally and non-naturally occurring modifications.

The polynucleotides of the multimeric structures of the invention can include any useful modification, such as to the sugar, the nucleobase, or the internucleoside linkage (*e.g.*, to a linking phosphate / to a phosphodiester linkage / to the phosphodiester backbone). One or more atoms of a pyrimidine nucleobase may be replaced or substituted with optionally substituted amino, optionally substituted thiol, optionally substituted alkyl (*e.g.*, methyl or ethyl), or halo (*e.g.*, chloro or fluoro). In certain embodiments, modifications (*e.g.*, one or more modifications) are present in each of the sugar and the internucleoside linkage. Modifications according to the present invention may be modifications of ribonucleic acids (RNAs) to deoxyribonucleic acids (DNAs), threose nucleic acids (TNAs), glycol nucleic

acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs) or hybrids thereof). Additional modifications are described herein.

Non-natural modified nucleotides may be introduced to polynucleotides during synthesis or post-synthesis of the chains to achieve desired functions or properties. The modifications may be on internucleotide lineage, the purine or pyrimidine bases, or sugar. The modification may be introduced at the terminal of a chain or anywhere else in the chain; with chemical synthesis or with a polymerase enzyme. Any of the regions of the polynucleotides may be chemically modified.

The present disclosure provides for multimeric structures comprised of unmodified or modified nucleosides and nucleotides and combinations thereof. As described herein "nucleoside" is defined as 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"). As described herein, "nucleotide" is defined as a nucleoside including a phosphate group. The modified nucleotides may by synthesized by any useful method, as described herein (*e.g.*, chemically, enzymatically, or recombinantly to include one or more modified or non-natural nucleosides). The polynucleotides may comprise a region or regions of linked nucleosides. Such regions may have variable backbone linkages. The linkages may be standard phosphodiester linkages, in which case the polynucleotides would comprise regions of nucleotides. Any combination of base/sugar or linker may be incorporated into the polynucleotides of the invention.

Modifications of the polynucleotides of the multimeric structures which are useful in the present invention include, but are not limited to the following: 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine; 2-methylthio-N6-methyladenosine; 2-methylthio-N6-threonyl carbamoyladenosine; N6-glycinylcarbamoyladenosine; N6-isopentenyladenosine; N6-methyladenosine; N6-threonylcarbamoyladenosine; 1,2′-O-dimethyladenosine; 1-methyladenosine; 2′-O-methyladenosine; 2′-O-ribosyladenosine (phosphate); 2-methyladenosine; 2-methylthio-N6 isopentenyladenosine; 2-methylthio-N6-hydroxynorvalyl carbamoyladenosine; 2'-O-methyladenosine; 2'-O-ribosyladenosine (phosphate); Isopentenyladenosine; N6-(cis-hydroxyisopentenyl)adenosine; N6,2′-O-dimethyladenosine; N6,2'-O-dimethyladenosine; N6,N6,2′-O-trimethyladenosine; N6,N6-dimethyladenosine; N6-acetyladenosine; N6-hydroxynorvalylcarbamoyladenosine; N6-methyl-N6-threonylcarbamoyladenosine; 2-methyladenosine; 2-methylthio-N6-isopentenyladenosine; 7-deaza-adenosine; N1-methyl-adenosine; N6, N6 (dimethyl)adenine; N6-cis-hydroxy-

isopentenyl-adenosine; α-thio-adenosine; 2 (amino)adenine; 2 (aminopropyl)adenine; 2 (methylthio) N6 (isopentenyl)adenine; 2-(alkyl)adenine; 2-(aminoalkyl)adenine; 2-(aminopropyl)adenine; 2-(halo)adenine; 2-(halo)adenine; 2-(propyl)adenine; 2'-Amino-2'-deoxy-ATP; 2'-Azido-2'-deoxy-ATP; 2'-Deoxy-2'-a-aminoadenosine TP; 2'-Deoxy-2'-a-azidoadenosine TP; 6 (alkyl)adenine; 6 (methyl)adenine; 6-(alkyl)adenine ; 6-(methyl)adenine; 7 (deaza)adenine; 8 (alkenyl)adenine; 8 (alkynyl)adenine; 8 (amino)adenine; 8 (thioalkyl)adenine; 8-(alkenyl)adenine; 8-(alkyl)adenine; 8-(alkynyl)adenine; 8-(amino)adenine; 8-(halo)adenine; 8-(hydroxyl)adenine; 8-(thioalkyl)adenine; 8-(thiol)adenine; 8-azido-adenosine; aza adenine; deaza adenine; N6 (methyl)adenine; N6-(isopentyl)adenine; 7-deaza-8-aza-adenosine; 7-methyladenine; 1-Deazaadenosine TP; 2'Fluoro-N6-Bz-deoxyadenosine TP; 2'-OMe-2-Amino-ATP; 2'O-methyl-N6-Bz-deoxyadenosine TP; 2'-a-Ethynyladenosine TP; 2-aminoadenine; 2-Aminoadenosine TP; 2-Amino-ATP; 2'-a-Trifluoromethyladenosine TP; 2-Azidoadenosine TP; 2'-b-Ethynyladenosine TP; 2-Bromoadenosine TP; 2'-b-Trifluoromethyladenosine TP; 2-Chloroadenosine TP; 2'-Deoxy-2',2'-difluoroadenosine TP; 2'-Deoxy-2'-a-mercaptoadenosine TP; 2'-Deoxy-2'-a-thiomethoxyadenosine TP; 2'-Deoxy-2'-b-aminoadenosine TP; 2'-Deoxy-2'-b-azidoadenosine TP; 2'-Deoxy-2'-b-bromoadenosine TP; 2'-Deoxy-2'-b-chloroadenosine TP; 2'-Deoxy-2'-b-fluoroadenosine TP; 2'-Deoxy-2'-b-iodoadenosine TP; 2'-Deoxy-2'-b-mercaptoadenosine TP; 2'-Deoxy-2'-b-thiomethoxyadenosine TP; 2-Fluoroadenosine TP; 2-Iodoadenosine TP; 2-Mercaptoadenosine TP; 2-methoxy-adenine; 2-methylthio-adenine; 2-Trifluoromethyladenosine TP; 3-Deaza-3-bromoadenosine TP; 3-Deaza-3-chloroadenosine TP; 3-Deaza-3-fluoroadenosine TP; 3-Deaza-3-iodoadenosine TP; 3-Deazaadenosine TP; 4'-Azidoadenosine TP; 4'-Carbocyclic adenosine TP; 4'-Ethynyladenosine TP; 5'-Homo-adenosine TP; 8-Aza-ATP; 8-bromo-adenosine TP; 8-Trifluoromethyladenosine TP; 9-Deazaadenosine TP; 2-aminopurine; 7-deaza-2,6-diaminopurine; 7-deaza-8-aza-2,6-diaminopurine; 7-deaza-8-aza-2-aminopurine; 2,6-diaminopurine; 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine; 2-thiocytidine; 3-methylcytidine; 5-formylcytidine; 5-hydroxymethylcytidine; 5-methylcytidine; N4-acetylcytidine; 2′-O-methylcytidine; 2'-O-methylcytidine; 5,2′-O-dimethylcytidine; 5-formyl-2′-O-methylcytidine; Lysidine; N4,2′-O-dimethylcytidine; N4-acetyl-2′-O-methylcytidine; N4-methylcytidine; N4,N4-Dimethyl-2'-OMe-Cytidine TP; 4-methylcytidine; 5-aza-cytidine; Pseudo-iso-cytidine; pyrrolo-cytidine; α-thio-cytidine; 2-(thio)cytosine; 2'-Amino-2'-deoxy-CTP; 2'-Azido-2'-deoxy-CTP; 2'-Deoxy-2'-a-aminocytidine TP; 2'-Deoxy-2'-a-azidocytidine TP; 3 (deaza) 5 (aza)cytosine; 3

(methyl)cytosine; 3-(alkyl)cytosine; 3-(deaza) 5 (aza)cytosine; 3-(methyl)cytidine; 4,2'-O-dimethylcytidine; 5 (halo)cytosine; 5 (methyl)cytosine; 5 (propynyl)cytosine; 5 (trifluoromethyl)cytosine; 5-(alkyl)cytosine; 5-(alkynyl)cytosine; 5-(halo)cytosine; 5-(propynyl)cytosine; 5-(trifluoromethyl)cytosine; 5-bromo-cytidine; 5-iodo-cytidine; 5-propynyl cytosine; 6-(azo)cytosine; 6-aza-cytidine; aza cytosine; deaza cytosine; N4 (acetyl)cytosine; 1-methyl-1-deaza-pseudoisocytidine; 1-methyl-pseudoisocytidine; 2-methoxy-5-methyl-cytidine; 2-methoxy-cytidine; 2-thio-5-methyl-cytidine; 4-methoxy-1-methyl-pseudoisocytidine; 4-methoxy-pseudoisocytidine; 4-thio-1-methyl-1-deaza-pseudoisocytidine; 4-thio-1-methyl-pseudoisocytidine; 4-thio-pseudoisocytidine; 5-aza-zebularine; 5-methyl-zebularine; pyrrolo-pseudoisocytidine; Zebularine; (E)-5-(2-Bromo-vinyl)cytidine TP; 2,2'-anhydro-cytidine TP hydrochloride; 2'Fluor-N4-Bz-cytidine TP; 2'Fluoro-N4-Acetyl-cytidine TP; 2'-O-Methyl-N4-Acetyl-cytidine TP; 2'O-methyl-N4-Bz-cytidine TP; 2'-a-Ethynylcytidine TP; 2'-a-Trifluoromethylcytidine TP; 2'-b-Ethynylcytidine TP; 2'-b-Trifluoromethylcytidine TP; 2'-Deoxy-2',2'-difluorocytidine TP; 2'-Deoxy-2'-a-mercaptocytidine TP; 2'-Deoxy-2'-a-thiomethoxycytidine TP; 2'-Deoxy-2'-b-aminocytidine TP; 2'-Deoxy-2'-b-azidocytidine TP; 2'-Deoxy-2'-b-bromocytidine TP; 2'-Deoxy-2'-b-chlorocytidine TP; 2'-Deoxy-2'-b-fluorocytidine TP; 2'-Deoxy-2'-b-iodocytidine TP; 2'-Deoxy-2'-b-mercaptocytidine TP; 2'-Deoxy-2'-b-thiomethoxycytidine TP; 2'-O-Methyl-5-(1-propynyl)cytidine TP; 3'-Ethynylcytidine TP; 4'-Azidocytidine TP; 4'-Carbocyclic cytidine TP; 4'-Ethynylcytidine TP; 5-(1-Propynyl)ara-cytidine TP; 5-(2-Chloro-phenyl)-2-thiocytidine TP; 5-(4-Amino-phenyl)-2-thiocytidine TP; 5-Aminoallyl-CTP; 5-Cyanocytidine TP; 5-Ethynylara-cytidine TP; 5-Ethynylcytidine TP; 5'-Homo-cytidine TP; 5-Methoxycytidine TP; 5-Trifluoromethyl-Cytidine TP; N4-Amino-cytidine TP; N4-Benzoyl-cytidine TP; Pseudoisocytidine; 7-methylguanosine; N2,2'-O-dimethylguanosine; N2-methylguanosine; Wyosine; 1,2'-O-dimethylguanosine; 1-methylguanosine; 2'-O-methylguanosine; 2'-O-ribosylguanosine (phosphate); 2'-O-methylguanosine; 2'-O-ribosylguanosine (phosphate); 7-aminomethyl-7-deazaguanosine; 7-cyano-7-deazaguanosine; Archaeosine; Methylwyosine; N2,7-dimethylguanosine; N2,N2,2'-O-trimethylguanosine; N2,N2,7-trimethylguanosine; N2,N2-dimethylguanosine; N2,7,2'-O-trimethylguanosine; 6-thio-guanosine; 7-deaza-guanosine; 8-oxo-guanosine; N1-methyl-guanosine; α-thio-guanosine; 2 (propyl)guanine; 2-(alkyl)guanine; 2'-Amino-2'-deoxy-GTP; 2'-Azido-2'-deoxy-GTP; 2'-Deoxy-2'-a-aminoguanosine TP; 2'-Deoxy-2'-a-azidoguanosine TP; 6 (methyl)guanine; 6-(alkyl)guanine; 6-(methyl)guanine; 6-methyl-guanosine; 7

(alkyl)guanine; 7 (deaza)guanine; 7 (methyl)guanine; 7-(alkyl)guanine; 7-(deaza)guanine; 7-(methyl)guanine; 8 (alkyl)guanine; 8 (alkynyl)guanine; 8 (halo)guanine; 8 (thioalkyl)guanine; 8-(alkenyl)guanine; 8-(alkyl)guanine; 8-(alkynyl)guanine; 8-(amino)guanine; 8-(halo)guanine; 8-(hydroxyl)guanine; 8-(thioalkyl)guanine; 8-(thiol)guanine; aza guanine; deaza guanine; N (methyl)guanine; N-(methyl)guanine; 1-methyl-6-thio-guanosine; 6-methoxy-guanosine; 6-thio-7-deaza-8-aza-guanosine; 6-thio-7-deaza-guanosine; 6-thio-7-methyl-guanosine; 7-deaza-8-aza-guanosine; 7-methyl-8-oxo-guanosine; N2,N2-dimethyl-6-thio-guanosine; N2-methyl-6-thio-guanosine; 1-Me-GTP; 2'Fluoro-N2-isobutyl-guanosine TP; 2'O-methyl-N2-isobutyl-guanosine TP; 2'-a-Ethynylguanosine TP; 2'-a-Trifluoromethylguanosine TP; 2'-b-Ethynylguanosine TP; 2'-b-Trifluoromethylguanosine TP; 2'-Deoxy-2',2'-difluoroguanosine TP; 2'-Deoxy-2'-a-mercaptoguanosine TP; 2'-Deoxy-2'-a-thiomethoxyguanosine TP; 2'-Deoxy-2'-b-aminoguanosine TP; 2'-Deoxy-2'-b-azidoguanosine TP; 2'-Deoxy-2'-b-bromoguanosine TP; 2'-Deoxy-2'-b-chloroguanosine TP; 2'-Deoxy-2'-b-fluoroguanosine TP; 2'-Deoxy-2'-b-iodoguanosine TP; 2'-Deoxy-2'-b-mercaptoguanosine TP; 2'-Deoxy-2'-b-thiomethoxyguanosine TP; 4'-Azidoguanosine TP; 4'-Carbocyclic guanosine TP; 4'-Ethynylguanosine TP; 5'-Homo-guanosine TP; 8-bromo-guanosine TP; 9-Deazaguanosine TP; N2-isobutyl-guanosine TP; 1-methylinosine; Inosine; 1,2'-O-dimethylinosine; 2'-O-methylinosine; 7-methylinosine; 2'-O-methylinosine; Epoxyqueuosine; galactosyl-queuosine; Mannosylqueuosine; Queuosine; allyamino-thymidine; aza thymidine; deaza thymidine; deoxy-thymidine; 2'-O-methyluridine; 2-thiouridine; 3-methyluridine; 5-carboxymethyluridine; 5-hydroxyuridine; 5-methyluridine; 5-taurinomethyl-2-thiouridine; 5-taurinomethyluridine; Dihydrouridine; Pseudouridine; (3-(3-amino-3-carboxypropyl)uridine; 1-methyl-3-(3-amino-5-carboxypropyl)pseudouridine; 1-methylpseduouridine; 1-methyl-pseudouridine; 2'-O-methyluridine; 2'-O-methylpseudouridine; 2'-O-methyluridine; 2-thio-2'-O-methyluridine; 3-(3-amino-3-carboxypropyl)uridine; 3,2'-O-dimethyluridine; 3-Methyl-pseudo-Uridine TP; 4-thiouridine; 5-(carboxyhydroxymethyl)uridine; 5-(carboxyhydroxymethyl)uridine methyl ester; 5,2'-O-dimethyluridine; 5,6-dihydro-uridine; 5-aminomethyl-2-thiouridine; 5-carbamoylmethyl-2'-O-methyluridine; 5-carbamoylmethyluridine; 5-carboxyhydroxymethyluridine; 5-carboxyhydroxymethyluridine methyl ester; 5-carboxymethylaminomethyl-2'-O-methyluridine; 5-carboxymethylaminomethyl-2-thiouridine; 5-carboxymethylaminomethyl-2-thiouridine; 5-carboxymethylaminomethyluridine; 5-carboxymethylaminomethyluridine; 5-Carbamoylmethyluridine TP; 5-methoxycarbonylmethyl-2'-O-methyluridine; 5-

methoxycarbonylmethyl-2-thiouridine; 5-methoxycarbonylmethyluridine; 5-methoxyuridine; 5-methyl-2-thiouridine; 5-methylaminomethyl-2-selenouridine; 5-methylaminomethyl-2-thiouridine; 5-methylaminomethyluridine; 5-Methyldihydrouridine; 5-Oxyacetic acid-Uridine TP; 5-Oxyacetic acid-methyl ester-Uridine TP; N1-methyl-pseudo-uridine; uridine 5-oxyacetic acid; uridine 5-oxyacetic acid methyl ester; 3-(3-Amino-3-carboxypropyl)-Uridine TP; 5-(iso-Pentenylaminomethyl)- 2-thiouridine TP; 5-(iso-Pentenylaminomethyl)-2'-O-methyluridine TP; 5-(iso-Pentenylaminomethyl)uridine TP; 5-propynyl uracil; α-thio-uridine; 1 (aminoalkylamino-carbonylethylenyl)-2(thio)-pseudouracil; 1 (aminoalkylaminocarbonylethylenyl)-2,4-(dithio)pseudouracil; 1 (aminoalkylaminocarbonylethylenyl)-4 (thio)pseudouracil; 1 (aminoalkylaminocarbonylethylenyl)-pseudouracil; 1 (aminocarbonylethylenyl)-2(thio)-pseudouracil; 1 (aminocarbonylethylenyl)-2,4-(dithio)pseudouracil; 1 (aminocarbonylethylenyl)-4 (thio)pseudouracil; 1 (aminocarbonylethylenyl)-pseudouracil; 1 substituted 2(thio)-pseudouracil; 1 substituted 2,4-(dithio)pseudouracil; 1 substituted 4 (thio)pseudouracil; 1 substituted pseudouracil; 1-(aminoalkylamino-carbonylethylenyl)-2-(thio)-pseudouracil; 1-Methyl-3-(3-amino-3-carboxypropyl) pseudouridine TP; 1-Methyl-3-(3-amino-3-carboxypropyl)pseudo-UTP ; 1-Methyl-pseudo-UTP ; 2 (thio)pseudouracil; 2' deoxy uridine; 2' fluorouridine; 2-(thio)uracil; 2,4-(dithio)psuedouracil; 2' methyl, 2'amino, 2'azido, 2'fluro-guanosine; 2'-Amino-2'-deoxy-UTP; 2'-Azido-2'-deoxy-UTP; 2'-Azido-deoxyuridine TP; 2'-O-methylpseudouridine; 2′ deoxy uridine; 2′ fluorouridine; 2'-Deoxy-2'-a-aminouridine TP; 2'-Deoxy-2'-a-azidouridine TP; 2-methylpseudouridine; 3 (3 amino-3 carboxypropyl)uracil; 4 (thio)pseudouracil; 4-(thio)pseudouracil; 4-(thio)uracil; 4-thiouracil; 5 (1,3-diazole-1-alkyl)uracil; 5 (2-aminopropyl)uracil; 5 (aminoalkyl)uracil; 5 (dimethylaminoalkyl)uracil; 5 (guanidiniumalkyl)uracil; 5 (methoxycarbonylmethyl)-2-(thio)uracil; 5 (methoxycarbonyl-methyl)uracil; 5 (methyl) 2 (thio)uracil; 5 (methyl) 2,4 (dithio)uracil; 5 (methyl) 4 (thio)uracil; 5 (methylaminomethyl)-2 (thio)uracil; 5 (methylaminomethyl)-2,4 (dithio)uracil; 5 (methylaminomethyl)-4 (thio)uracil; 5 (propynyl)uracil; 5 (trifluoromethyl)uracil; 5-(2-aminopropyl)uracil; 5-(alkyl)-2-(thio)pseudouracil; 5-(alkyl)-2,4 (dithio)pseudouracil; 5-(alkyl)-4 (thio)pseudouracil; 5-(alkyl)pseudouracil; 5-(alkyl)uracil; 5-(alkynyl)uracil; 5-(allylamino)uracil; 5-(cyanoalkyl)uracil; 5-(dialkylaminoalkyl)uracil; 5-(dimethylaminoalkyl)uracil; 5-(guanidiniumalkyl)uracil; 5-(halo)uracil; 5-(l,3-diazole-l-alkyl)uracil; 5-(methoxy)uracil; 5-(methoxycarbonylmethyl)-2-(thio)uracil; 5-(methoxycarbonyl-methyl)uracil; 5-(methyl)

2(thio)uracil; 5-(methyl) 2,4 (dithio)uracil; 5-(methyl) 4 (thio)uracil; 5-(methyl)-2-(thio)pseudouracil; 5-(methyl)-2,4 (dithio)pseudouracil; 5-(methyl)-4 (thio)pseudouracil; 5-(methyl)pseudouracil; 5-(methylaminomethyl)-2 (thio)uracil; 5-(methylaminomethyl)-2,4(dithio)uracil; 5-(methylaminomethyl)-4-(thio)uracil; 5-(propynyl)uracil; 5-(trifluoromethyl)uracil; 5-aminoallyl-uridine; 5-bromo-uridine; 5-iodo-uridine; 5-uracil; 6 (azo)uracil; 6-(azo)uracil; 6-aza-uridine; allyamino-uracil; aza uracil; deaza uracil; N3 (methyl)uracil; P seudo-UTP-1-2-ethanoic acid ; Pseudouracil; 4-Thio-pseudo-UTP ; 1-carboxymethyl-pseudouridine; 1-methyl-1-deaza-pseudouridine; 1-propynyl-uridine; 1-taurinomethyl-1-methyl-uridine; 1-taurinomethyl-4-thio-uridine; 1-taurinomethyl-pseudouridine; 2-methoxy-4-thio-pseudouridine; 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; (±)1-(2-Hydroxypropyl)pseudouridine TP; (2R)-1-(2-Hydroxypropyl)pseudouridine TP; (2S)-1-(2-Hydroxypropyl)pseudouridine TP; (E)-5-(2-Bromo-vinyl)ara-uridine TP; (E)-5-(2-Bromo-vinyl)uridine TP; (Z)-5-(2-Bromo-vinyl)ara-uridine TP; (Z)-5-(2-Bromo-vinyl)uridine TP; 1-(2,2,2-Trifluoroethyl)-pseudo-UTP ; 1-(2,2,3,3,3-Pentafluoropropyl)pseudouridine TP; 1-(2,2-Diethoxyethyl)pseudouridine TP; 1-(2,4,6-Trimethylbenzyl)pseudouridine TP; 1-(2,4,6-Trimethyl-benzyl)pseudo-UTP ; 1-(2,4,6-Trimethyl-phenyl)pseudo-UTP ; 1-(2-Amino-2-carboxyethyl)pseudo-UTP; 1-(2-Amino-ethyl)pseudo-UTP; 1-(2-Hydroxyethyl)pseudouridine TP; 1-(2-Methoxyethyl)pseudouridine TP; 1-(3,4-Bis-trifluoromethoxybenzyl)pseudouridine TP; 1-(3,4-Dimethoxybenzyl)pseudouridine TP; 1-(3-Amino-3-carboxypropyl)pseudo-UTP ; 1-(3-Amino-propyl)pseudo-UTP; 1-(3-Cyclopropyl-prop-2-ynyl)pseudouridine TP; 1-(4-Amino-4-carboxybutyl)pseudo-UTP; 1-(4-Amino-benzyl)pseudo-UTP; 1-(4-Amino-butyl)pseudo-UTP; 1-(4-Amino-phenyl)pseudo-UTP ; 1-(4-Azidobenzyl)pseudouridine TP; 1-(4-Bromobenzyl)pseudouridine TP; 1-(4-Chlorobenzyl)pseudouridine TP; 1-(4-Fluorobenzyl)pseudouridine TP; 1-(4-Iodobenzyl)pseudouridine TP; 1-(4-Methanesulfonylbenzyl)pseudouridine TP; 1-(4-Methoxybenzyl)pseudouridine TP; 1-(4-Methoxy-benzyl)pseudo-UTP; 1-(4-Methoxy-phenyl)pseudo-UTP ; 1-(4-Methylbenzyl)pseudouridine TP; 1-(4-Methyl-benzyl)pseudo-UTP ; 1-(4-Nitrobenzyl)pseudouridine TP; 1-(4-Nitro-benzyl)pseudo-UTP; 1(4-Nitro-phenyl)pseudo-UTP ; 1-(4-Thiomethoxybenzyl)pseudouridine TP; 1-(4-

Trifluoromethoxybenzyl)pseudouridine TP; 1-(4-Trifluoromethylbenzyl)pseudouridine TP; 1-(5-Amino-pentyl)pseudo-UTP; 1-(6-Amino-hexyl)pseudo-UTP; 1,6-Dimethyl-pseudo-UTP; 1-[3-(2-{2-[2-(2-Aminoethoxy)-ethoxy]-ethoxy}-ethoxy)-propionyl]pseudouridine TP; 1-{3-[2-(2-Aminoethoxy)-ethoxy]-propionyl } pseudouridine TP; 1-Acetylpseudouridine TP; 1-Alkyl-6-(1-propynyl)-pseudo-UTP; 1-Alkyl-6-(2-propynyl)-pseudo-UTP; 1-Alkyl-6-allyl-pseudo-UTP; 1-Alkyl-6-ethynyl-pseudo-UTP; 1-Alkyl-6-homoallyl-pseudo-UTP; 1-Alkyl-6-vinyl-pseudo-UTP; 1-Allylpseudouridine TP; 1-Aminomethyl-pseudo-UTP ; 1-Benzoylpseudouridine TP; 1-Benzyloxymethylpseudouridine TP; 1-Benzyl-pseudo-UTP ; 1-Biotinyl-PEG2-pseudouridine TP; 1-Biotinylpseudouridine TP; 1-Butyl-pseudo-UTP ; 1-Cyanomethylpseudouridine TP; 1-Cyclobutylmethyl-pseudo-UTP ; 1-Cyclobutyl-pseudo-UTP ; 1-Cycloheptylmethyl-pseudo-UTP ; 1-Cycloheptyl-pseudo-UTP ; 1-Cyclohexylmethyl-pseudo-UTP ; 1-Cyclohexyl-pseudo-UTP ; 1-Cyclooctylmethyl-pseudo-UTP; 1-Cyclooctyl-pseudo-UTP ; 1-Cyclopentylmethyl-pseudo-UTP ; 1-Cyclopentyl-pseudo-UTP ; 1-Cyclopropylmethyl-pseudo-UTP ; 1-Cyclopropyl-pseudo-UTP ; 1-Ethyl-pseudo-UTP ; 1-Hexyl-pseudo-UTP ; 1-Homoallylpseudouridine TP; 1-Hydroxymethylpseudouridine TP; 1-iso-propyl-pseudo-UTP ; 1-Me-2-thio-pseudo-UTP; 1-Me-4-thio-pseudo-UTP; 1-Me-alpha-thio-pseudo-UTP ; 1-Methanesulfonylmethylpseudouridine TP; 1-Methoxymethylpseudouridine TP; 1-Methyl-6-(2,2,2-Trifluoroethyl)pseudo-UTP; 1-Methyl-6-(4-morpholino)-pseudo-UTP; 1-Methyl-6-(4-thiomorpholino)-pseudo-UTP; 1-Methyl-6-(substituted phenyl)pseudo-UTP; 1-Methyl-6-amino-pseudo-UTP; 1-Methyl-6-azido-pseudo-UTP; 1-Methyl-6-bromo-pseudo-UTP; 1-Methyl-6-butyl-pseudo-UTP; 1-Methyl-6-chloro-pseudo-UTP; 1-Methyl-6-cyano-pseudo-UTP; 1-Methyl-6-dimethylamino-pseudo-UTP; 1-Methyl-6-ethoxy-pseudo-UTP; 1-Methyl-6-ethylcarboxylate-pseudo-UTP; 1-Methyl-6-ethyl-pseudo-UTP; 1-Methyl-6-fluoro-pseudo-UTP; 1-Methyl-6-formyl-pseudo-UTP; 1-Methyl-6-hydroxyamino-pseudo-UTP; 1-Methyl-6-hydroxy-pseudo-UTP; 1-Methyl-6-iodo-pseudo-UTP; 1-Methyl-6-iso-propyl-pseudo-UTP; 1-Methyl-6-methoxy-pseudo-UTP; 1-Methyl-6-methylamino-pseudo-UTP; 1-Methyl-6-phenyl-pseudo-UTP; 1-Methyl-6-propyl-pseudo-UTP; 1-Methyl-6-tert-butyl-pseudo-UTP; 1-Methyl-6-trifluoromethoxy-pseudo-UTP; 1-Methyl-6-trifluoromethyl-pseudo-UTP; 1-Morpholinomethylpseudouridine TP; 1-Pentyl-pseudo-UTP ; 1-Phenyl-pseudo-UTP ; 1-Pivaloylpseudouridine TP; 1-Propargylpseudouridine TP; 1-Propyl-pseudo-UTP ; 1-propynyl-pseudouridine; 1-p-tolyl-pseudo-UTP ; 1-tert-Butyl-pseudo-UTP ; 1-Thiomethoxymethylpseudouridine TP; 1-Thiomorpholinomethylpseudouridine TP; 1-

Trifluoroacetylpseudouridine TP; 1-Trifluoromethyl-pseudo-UTP ; 1-Vinylpseudouridine TP; 2,2'-anhydro-uridine TP; 2'-bromo-deoxyuridine TP; 2'-F-5-Methyl-2'-deoxy-UTP; 2'-OMe-5-Me-UTP; 2'-OMe-pseudo-UTP; 2'-a-Ethynyluridine TP; 2'-a-Trifluoromethyluridine TP; 2'-b-Ethynyluridine TP; 2'-b-Trifluoromethyluridine TP; 2'-Deoxy-2',2'-difluorouridine TP; 2'-Deoxy-2'-a-mercaptouridine TP; 2'-Deoxy-2'-a-thiomethoxyuridine TP; 2'-Deoxy-2'-b-aminouridine TP; 2'-Deoxy-2'-b-azidouridine TP; 2'-Deoxy-2'-b-bromouridine TP; 2'-Deoxy-2'-b-chlorouridine TP; 2'-Deoxy-2'-b-fluorouridine TP; 2'-Deoxy-2'-b-iodouridine TP; 2'-Deoxy-2'-b-mercaptouridine TP; 2'-Deoxy-2'-b-thiomethoxyuridine TP; 2-methoxy-4-thio-uridine; 2-methoxyuridine; 2'-O-Methyl-5-(1-propynyl)uridine TP; 3-Alkyl-pseudo-UTP; 4'-Azidouridine TP; 4'-Carbocyclic uridine TP; 4'-Ethynyluridine TP; 5-(1-Propynyl)ara-uridine TP; 5-(2-Furanyl)uridine TP; 5-Cyanouridine TP; 5-Dimethylaminouridine TP; 5'-Homo-uridine TP; 5-iodo-2'-fluoro-deoxyuridine TP; 5-Phenylethynyluridine TP; 5-Trideuteromethyl-6-deuterouridine TP; 5-Trifluoromethyl-Uridine TP; 5-Vinylarauridine TP; 6-(2,2,2-Trifluoroethyl)-pseudo-UTP; 6-(4-Morpholino)-pseudo-UTP; 6-(4-Thiomorpholino)-pseudo-UTP; 6-(Substituted-Phenyl)-pseudo-UTP; 6-Amino-pseudo-UTP; 6-Azido-pseudo-UTP; 6-Bromo-pseudo-UTP ; 6-Butyl-pseudo-UTP; 6-Chloro-pseudo-UTP ; 6-Cyano-pseudo-UTP; 6-Dimethylamino-pseudo-UTP; 6-Ethoxy-pseudo-UTP; 6-Ethylcarboxylate-pseudo-UTP; 6-Ethyl-pseudo-UTP; 6-Fluoro-pseudo-UTP ; 6-Formyl-pseudo-UTP; 6-Hydroxyamino-pseudo-UTP; 6-Hydroxy-pseudo-UTP; 6-Iodo-pseudo-UTP ; 6-iso-Propyl-pseudo-UTP; 6-Methoxy-pseudo-UTP ; 6-Methylamino-pseudo-UTP; 6-Methyl-pseudo-UTP ; 6-Phenyl-pseudo-UTP; 6-Phenyl-pseudo-UTP ; 6-Propyl-pseudo-UTP; 6-tert-Butyl-pseudo-UTP; 6-Trifluoromethoxy-pseudo-UTP; 6-Trifluoromethyl-pseudo-UTP ; Alpha-thio-pseudo-UTP ; Pseudouridine 1-(4-methylbenzenesulfonic acid) TP; Pseudouridine 1-(4-methylbenzoic acid) TP; Pseudouridine TP 1-[3-(2-ethoxy)]propionic acid; Pseudouridine TP 1-[3-{2-(2-[2-(2-ethoxy)-ethoxy]-ethoxy)-ethoxy}]propionic acid; Pseudouridine TP 1-[3-{2-(2-[2-{2(2-ethoxy)-ethoxy}-ethoxy]-ethoxy)-ethoxy}]propionic acid; Pseudouridine TP 1-[3-{2-(2-[2-ethoxy]-ethoxy)-ethoxy}]propionic acid; Pseudouridine TP 1-[3-{2-(2-ethoxy)-ethoxy}] propionic acid; Pseudouridine TP 1-methylphosphonic acid; Pseudouridine TP 1-methylphosphonic acid diethyl ester; Pseudo-UTP-N1-3-propionic acid; Pseudo-UTP-N1-4-butanoic acid; Pseudo-UTP-N1-5-pentanoic acid; Pseudo-UTP-N1-6-hexanoic acid; Pseudo-UTP-N1-7-heptanoic acid; Pseudo-UTP-N1-methyl-p-benzoic acid; Pseudo-UTP-N1-p-benzoic acid ; Wybutosine; Hydroxywybutosine; Isowyosine; Peroxywybutosine; undermodified hydroxywybutosine; 4-demethylwyosine;

2,6-(diamino)purine;1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl: 1,3-(diaza)-2-(oxo)-phenthiazin-l-yl;1,3-(diaza)-2-(oxo)-phenoxazin-1-yl;1,3,5-(triaza)-2,6-(dioxa)-naphthalene;2(amino)purine;2,4,5-(trimethyl)phenyl;2‘ methyl, 2’amino, 2’azido, 2’fluro-cytidine;2’methyl, 2’amino, 2’azido, 2’fluro-adenine;2’methyl, 2’amino, 2’azido, 2’fluro-uridine;2'-amino-2'-deoxyribose; 2-amino-6-Chloro-purine; 2-aza-inosinyl; 2'-azido-2'-deoxyribose; 2'fluoro-2'-deoxyribose; 2'-fluoro-modified bases; 2'-O-methyl-ribose; 2-oxo-7-aminopyridopyrimidin-3-yl; 2-oxo-pyridopyrimidine-3-yl; 2-pyridinone; 3 nitropyrrole; 3-(methyl)-7-(propynyl)isocarbostyrilyl; 3-(methyl)isocarbostyrilyl; 4-(fluoro)-6-(methyl)benzimidazole; 4-(methyl)benzimidazole; 4-(methyl)indolyl; 4,6-(dimethyl)indolyl; 5 nitroindole; 5 substituted pyrimidines; 5-(methyl)isocarbostyrilyl; 5-nitroindole; 6-(aza)pyrimidine; 6-(azo)thymine; 6-(methyl)-7-(aza)indolyl; 6-chloro-purine; 6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; 7-(aminoalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenthiazin-l-yl; 7-(aminoalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 7-(aminoalkylhydroxy)-1,3-(diaza)-2-(oxo)-phenoxazin-1-yl; 7-(aminoalkylhydroxy)-l,3-(diaza)-2-(oxo)-phenthiazin-l-yl ; 7-(aminoalkylhydroxy)-l,3-(diaza)-2-(oxo)-phenoxazin-l-yl ; 7-(aza)indolyl; 7-(guanidiniumalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenoxazinl-yl; 7-(guanidiniumalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenthiazin-l-yl; 7-(guanidiniumalkylhydroxy)-1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 7-(guanidiniumalkylhydroxy)-1,3-(diaza)-2-(oxo)-phenoxazin-1-yl; 7-(guanidiniumalkyl-hydroxy)-l,3-(diaza)-2-(oxo)-phenthiazin-l-yl ; 7-(guanidiniumalkylhydroxy)-l,3-(diaza)-2-(oxo)-phenoxazin-l-yl; 7-(propynyl)isocarbostyrilyl; 7-(propynyl)isocarbostyrilyl, propynyl-7-(aza)indolyl; 7-deaza-inosinyl; 7-substituted 1-(aza)-2-(thio)-3-(aza)-phenoxazin-1-yl; 7-substituted 1,3-(diaza)-2-(oxo)-phenoxazin-1-yl; 9-(methyl)-imidizopyridinyl; Aminoindolyl; Anthracenyl; bis-ortho-(aminoalkylhydroxy)-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; bis-ortho-substituted-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; Difluorotolyl; Hypoxanthine; Imidizopyridinyl; Inosinyl; Isocarbostyrilyl; Isoguanisine; N2-substituted purines; N6-methyl-2-amino-purine; N6-substituted purines; N-alkylated derivative; Napthalenyl; Nitrobenzimidazolyl; Nitroimidazolyl; Nitroindazolyl; Nitropyrazolyl; Nubularine; O6-substituted purines; O-alkylated derivative; ortho-(aminoalkylhydroxy)-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; ortho-substituted-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; Oxoformycin TP; para-(aminoalkylhydroxy)-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; para-substituted-6-phenyl-pyrrolo-pyrimidin-2-on-3-yl; Pentacenyl; Phenanthracenyl; Phenyl; propynyl-7-(aza)indolyl; Pyrenyl; pyridopyrimidin-3-yl; pyridopyrimidin-3-yl, 2-oxo-7-amino-

pyridopyrimidin-3-yl; pyrrolo-pyrimidin-2-on-3-yl; Pyrrolopyrimidinyl; Pyrrolopyrizinyl; Stilbenzyl; substituted 1,2,4-triazoles; Tetracenyl; Tubercidine; Xanthine; Xanthosine-5'-TP; 2-thio-zebularine; 5-aza-2-thio-zebularine; 7-deaza-2-amino-purine; pyridin-4-one ribonucleoside; 2-Amino-riboside-TP; Formycin A TP; Formycin B TP; Pyrrolosine TP; 2'-OH-ara-adenosine TP; 2'-OH-ara-cytidine TP; 2'-OH-ara-uridine TP; 2'-OH-ara-guanosine TP; 5-(2-carbomethoxyvinyl)uridine TP; and N6-(19-Amino-pentaoxanonadecyl)adenosine TP.

In some embodiments, an mRNA of the invention includes a combination of one or more of the aforementioned modified nucleobases (*e.g.*, a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

In some embodiments, the modified nucleobase is pseudouridine (ψ), N1-methylpseudouridine ($m^1\psi$), 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-methoxyuridine, or 2'-O-methyl uridine. In some embodiments, an mRNA of the invention includes a combination of one or more of the aforementioned modified nucleobases (*e.g.*, a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

In some embodiments, the modified nucleobase is 1-methyl-pseudouridine ($m^1\psi$), 5-methoxy-uridine (mo^5U), 5-methyl-cytidine (m^5C), pseudouridine (ψ), α-thio-guanosine, or α-thio-adenosine. In some embodiments, an mRNA of the invention includes a combination of one or more of the aforementioned modified nucleobases (*e.g.*, a combination of 2, 3 or 4 of the aforementioned modified nucleobases.)

In some embodiments, the mRNA comprises pseudouridine (ψ) and 5-methyl-cytidine (m^5C). In some embodiments, the mRNA comprises 1-methyl-pseudouridine ($m^1\psi$). In some embodiments, the mRNA comprises 1-methyl-pseudouridine ($m^1\psi$) and 5-methyl-cytidine (m^5C). In some embodiments, the mRNA comprises 2-thiouridine (s^2U). In some embodiments, the mRNA comprises 2-thiouridine and 5-methyl-cytidine (m^5C). In some embodiments, the mRNA comprises 5-methoxy-uridine (mo^5U). In some embodiments, the mRNA comprises 5-methoxy-uridine (mo^5U) and 5-methyl-cytidine (m^5C). In some embodiments, the mRNA comprises 2'-O-methyl uridine. In some embodiments, the mRNA comprises 2'-O-methyl uridine and 5-methyl-cytidine (m^5C). In some embodiments, the

mRNA comprises N6-methyl-adenosine (m^6A). In some embodiments, the mRNA comprises N6-methyl-adenosine (m^6A) and 5-methyl-cytidine (m^5C).

In certain embodiments, an mRNA of the invention is uniformly modified (i.e., fully modified, modified through-out the entire sequence) for a particular modification. For example, an mRNA can be uniformly modified with 5-methyl-cytidine (m^5C), meaning that all cytosine residues in the mRNA sequence are replaced with 5-methyl-cytidine (m^5C). Similarly, mRNAs of the invention 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.

In some embodiments, the modified nucleobase is a modified cytosine. Exemplary nucleobases and nucleosides having a modified cytosine include N4-acetyl-cytidine (ac4C), 5-methyl-cytidine (m5C), 5-halo-cytidine (*e.g.*, 5-iodo-cytidine), 5-hydroxymethyl-cytidine (hm5C), 1-methyl-pseudoisocytidine, 2-thio-cytidine (s2C), 2-thio-5-methyl-cytidine.

In some embodiments, the modified nucleobase is a modified uridine. Exemplary nucleobases and nucleosides having a modified uridine include 5-cyano uridine or 4'-thio uridine.

In some embodiments, the modified nucleobase is a modified adenine. Exemplary nucleobases and nucleosides having a modified adenine include 7-deaza-adenine, 1-methyl-adenosine (m1A), 2-methyl-adenine (m2A), N6-methyl-adenosine (m6A), and 2,6-Diaminopurine.

In some embodiments, the modified nucleobase is a modified guanine. Exemplary nucleobases and nucleosides having a modified guanine include inosine (I), 1-methyl-inosine (m1I), wyosine (imG), methylwyosine (mimG), 7-deaza-guanosine, 7-cyano-7-deaza-guanosine (preQ0), 7-aminomethyl-7-deaza-guanosine (preQ1), 7-methyl-guanosine (m7G), 1-methyl-guanosine (m1G), 8-oxo-guanosine, 7-methyl-8-oxo-guanosine.

In one embodiment, the polynucleotides of the present invention, such as IVT polynucleotides, may have a uniform chemical modification of all or any of the same nucleoside type or a population of modifications produced by mere downward titration of the same starting modification in all or any of the same nucleoside type, or a measured percent of a chemical modification of any of the same nucleoside type but with random incorporation, such as where all uridines are replaced by a uridine analog, *e.g.*, pseudouridine. In another embodiment, the polynucleotides may have a uniform chemical modification of two, three, or four of the same nucleoside type throughout the entire polynucleotide (such as all uridines

and all cytosines, *etc.* are modified in the same way). When the polynucleotides of the present invention are chemically and/or structurally modified the polynucleotides may be referred to as "modified polynucleotides."

Generally, the length of the IVT polynucleotide (*e.g.*, IVT mRNA) encoding a polypeptide of interest is greater than about 30 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, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000 or up to and including 100,000 nucleotides).

In some embodiments, the IVT polynucleotide (*e.g.*, IVT mRNA) includes from about 30 to about 100,000 nucleotides (*e.g.,* from 30 to 50, from 30 to 100, from 30 to 250, from 30 to 500, from 30 to 1,000, from 30 to 1,500, from 30 to 3,000, from 30 to 5,000, from 30 to 7,000, from 30 to 10,000, from 30 to 25,000, from 30 to 50,000, from 30 to 70,000, from 100 to 250, from 100 to 500, from 100 to 1,000, from 100 to 1,500, from 100 to 3,000, from 100 to 5,000, from 100 to 7,000, from 100 to 10,000, from 100 to 25,000, from 100 to 50,000, from 100 to 70,000, from 100 to 100,000, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to 3,000, from 500 to 5,000, from 500 to 7,000, from 500 to 10,000, from 500 to 25,000, from 500 to 50,000, from 500 to 70,000, from 500 to 100,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 3,000, from 1,000 to 5,000, from 1,000 to 7,000, from 1,000 to 10,000, from 1,000 to 25,000, from 1,000 to 50,000, from 1,000 to 70,000, from 1,000 to 100,000, from 1,500 to 3,000, from 1,500 to 5,000, from 1,500 to 7,000, from 1,500 to 10,000, from 1,500 to 25,000, from 1,500 to 50,000, from 1,500 to 70,000, from 1,500 to 100,000, from 2,000 to 3,000, from 2,000 to 5,000, from 2,000 to 7,000, from 2,000 to 10,000, from 2,000 to 25,000, from 2,000 to 50,000, from 2,000 to 70,000, and from 2,000 to 100,000).

In some embodiments, a nucleic acid of a multimeric molecule as described herein is a chimeric polynucleotide. Chimeric polynucleotides or RNA constructs maintain a modular organization similar to IVT polynucleotides, but the chimeric polynucleotides comprise one or more structural and/or chemical modifications or alterations which impart useful properties to the polynucleotide. As such, the chimeric polynucleotides which are modified mRNA molecules of the present invention are termed "chimeric modified mRNA" or "chimeric mRNA." Chimeric polynucleotides have portions or regions which differ in size and/or

chemical modification pattern, chemical modification position, chemical modification percent or chemical modification population and combinations of the foregoing.

In some embodiments, the multimeric nucleic acids are therapeutic mRNAs. As used herein, the term "therapeutic mRNA" refers to an mRNA that encodes a therapeutic protein. Therapeutic proteins mediate a variety of effects in a host cell or a subject in order to treat a disease or ameliorate the signs and symptoms of a disease. For example, a therapeutic protein can replace a protein that is deficient or abnormal, augment the function of an endogenous protein, provide a novel function to a cell (*e.g.*, inhibit or activate an endogenous cellular activity, or act as a delivery agent for another therapeutic compound (*e.g.*, an antibody-drug conjugate). Therapeutic mRNA may be useful for the treatment of the following diseases and conditions: bacterial infections, viral infections, parasitic infections, cell proliferation disorders, genetic disorders, and autoimmune disorders.

Thus, the multimeric structures of the invention can be used as therapeutic or prophylactic agents. They are provided for use in medicine. For example, the mRNA of the multimeric structures described herein can be administered to a subject, wherein the polynucleotides are translated in vivo to produce a therapeutic peptide. Provided are compositions, methods, kits, and reagents for diagnosis, treatment or prevention of a disease or condition in humans and other mammals. The active therapeutic agents of the invention include the multimeric structures, cells containing multimeric structures or polypeptides translated from the polynucleotides contained in the multimeric structures.

The multimeric structures may be induced for translation in a cell, tissue or organism. Such translation can be in vivo, ex vivo, in culture, or in vitro. The cell, tissue or organism is contacted with an effective amount of a composition containing a multimeric structure which contains the multiple mRNA polynucleotides each of which has at least one translatable region encoding a peptide.

An "effective amount" of the multimeric structures are provided based, at least in part, on the target tissue, target cell type, means of administration, physical characteristics of the polynucleotide (*e.g.*, size, and extent of modified nucleosides) and other components of the multimeric structures, and other determinants. In general, an effective amount of the multimeric structure provides an induced or boosted peptide production in the cell, preferably more efficient than a composition containing a corresponding unmodified polynucleotide encoding the same peptide or about the same or more efficient than separate mRNAs that are not part of a multimeric structure. Increased peptide production may be demonstrated by

increased cell transfection (i.e., the percentage of cells transfected with the multimeric structures), increased protein translation from the polynucleotide, decreased nucleic acid degradation (as demonstrated, *e.g.*, by increased duration of protein translation from a modified polynucleotide), or altered peptide production in the host cell.

The mRNA of the present invention may be designed to encode polypeptides of interest selected from any of several target categories including, but not limited to, biologics, antibodies, vaccines, therapeutic proteins or peptides, cell penetrating peptides, secreted proteins, plasma membrane proteins, cytoplasmic or cytoskeletal proteins, intracellular membrane bound proteins, nuclear proteins, proteins associated with human disease, targeting moieties or those proteins encoded by the human genome for which no therapeutic indication has been identified but which nonetheless have utility in areas of research and discovery. "Therapeutic protein" refers to a protein that, when administered to a cell has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect.

The mRNA disclosed herein, may encode one or more biologics. As used herein, a "biologic" is a polypeptide-based molecule produced by the methods provided herein and which may be used to treat, cure, mitigate, prevent, or diagnose a serious or life-threatening disease or medical condition. Biologics, according to the present invention include, but are not limited to, allergenic extracts (*e.g.*, for allergy shots and tests), blood components, gene therapy products, human tissue or cellular products used in transplantation, vaccines, monoclonal antibodies, cytokines, growth factors, enzymes, thrombolytics, and immunomodulators, among others.

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

The mRNA disclosed herein, may encode one or more antibodies or fragments thereof. The term "antibody" includes monoclonal antibodies (including full length antibodies which have an immunoglobulin Fc region), antibody compositions with polyepitopic specificity, multispecific antibodies (*e.g.*, bispecific antibodies, diabodies, and single-chain molecules), as well as antibody fragments. The term "immunoglobulin" (Ig) is used interchangeably with "antibody" herein. As used herein, the term "monoclonal antibody"

refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations and/or post-translation modifications (*e.g.*, isomerizations, amidations) that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site.

The monoclonal antibodies herein specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is(are) identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity. Chimeric antibodies of interest herein include, but are not limited to, "primatized" antibodies comprising variable domain antigen-binding sequences derived from a non-human primate (*e.g.*, Old World Monkey, Ape etc.) and human constant region sequences.

An "antibody fragment" comprises a portion of an intact antibody, preferably the antigen binding and/or the variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', $F(ab')_2$ and Fv fragments; diabodies; linear antibodies; nanobodies; single-chain antibody molecules and multispecific antibodies formed from antibody fragments.

Any of the five classes of immunoglobulins, IgA, IgD, IgE, IgG and IgM, may be encoded by the mRNA of the invention, including the heavy chains designated alpha, delta, epsilon, gamma and mu, respectively. Also included are polynucleotide sequences encoding the subclasses, gamma and mu. Hence any of the subclasses of antibodies may be encoded in part or in whole and include the following subclasses: IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2. According to the present invention, one or more antibodies or fragments currently being marketed or in development may be encoded by the mRNA of the present invention.

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

In one embodiment, mRNA disclosed herein may encode monoclonal antibodies and/or variants thereof. Variants of antibodies may also include, but are not limited to,

substitutional variants, conservative amino acid substitution, insertional variants, deletional variants and/or covalent derivatives. In one embodiment, the mRNA disclosed herein may encode an immunoglobulin Fc region. In another embodiment, the mRNA may encode a variant immunoglobulin Fc region.

The multimeric mRNA disclosed herein, may encode one or more vaccine antigens. As used herein, a "vaccine antigen" is a biological preparation that improves immunity to a particular disease or infectious agent. According to the present invention, one or more vaccine antigens currently being marketed or in development may be encoded by the multimeric mRNA of the present invention.

Vaccine antigens encoded in the mRNA of the invention may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, cancer, allergy and infectious disease.

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

In some embodiments, the anti-microbial polypeptide may have at least about 75%, 80%, 85%, 90%, 95%, 100% sequence homology to the corresponding viral protein sequence. In some embodiments, the anti-microbial polypeptide may have at least about 75%, 80%, 85%, 90%, 95%, or 100% sequence homology to the corresponding viral protein sequence.

In other embodiments, the anti-microbial polypeptide may comprise or consist of a synthetic peptide corresponding to a region, *e.g.*, a consecutive sequence of at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, or 60 amino acids of the binding domain of a capsid binding protein. In some embodiments, the anti-microbial polypeptide may have at least

about 75%, 80%, 85%, 90%, 95%, or 100% sequence homology to the corresponding sequence of the capsid binding protein.

The anti-microbial polypeptides described herein may block protease dimerization and inhibit cleavage of viral proproteins (*e.g.*, HIV Gag-pol processing) into functional proteins thereby preventing release of one or more enveloped viruses (*e.g.*, HIV, HCV). In some embodiments, the anti-microbial polypeptide may have at least about 75%, 80%, 85%, 90%, 95%, 100% sequence homology to the corresponding viral protein sequence.

In other embodiments, the anti-microbial polypeptide can comprise or consist of a synthetic peptide corresponding to a region, *e.g.*, a consecutive sequence of at least about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, or 60 amino acids of the binding domain of a protease binding protein. In some embodiments, the anti-microbial polypeptide may have at least about 75%, 80%, 85%, 90%, 95%, 100% sequence homology to the corresponding sequence of the protease binding protein.

A non-limiting list of infectious diseases that the mRNA vaccine antigens or anti-microbial peptides may treat is presented below: human immunodeficiency virus (HIV), HIV resulting in mycobacterial infection, AIDS related Cacheixa, AIDS related Cytomegalovirus infection, HIV-associated nephropathy, Lipodystrophy, AID related cryptococcal meningitis, AIDS related neutropaenia, Pneumocysitis jiroveci (Pneumocystis carinii) infections, AID related toxoplasmosis, hepatitis A, B, C, D or E, herpes, herpes zoster (chicken pox), German measles (rubella virus), yellow fever, dengue fever etc. (flavi viruses), flu (influenza viruses), haemorrhagic infectious diseases (Marburg or Ebola viruses), bacterial infectious diseases such as Legionnaires' disease (Legionella), gastric ulcer (Helicobacter), cholera (Vibrio), E. coli infections, staphylococcal infections, salmonella infections or streptococcal infections, tetanus (Clostridium tetani), protozoan infectious diseases (malaria, sleeping sickness, leishmaniasis, toxoplasmosis, i.e. infections caused by plasmodium, trypanosomes, leishmania and toxoplasma), diphtheria, leprosy, measles, pertussis, rabies, tetanus, tuberculosis, typhoid, varicella, diarrheal infections such as Amoebiasis, Clostridium difficile-associated diarrhea (CDAD), Cryptosporidiosis, Giardiasis, Cyclosporiasis and Rotaviral gastroenteritis, encephalitis such as Japanese encephalitis, Wester equine encephalitis and Tick-borne encephalitis (TBE), fungal skin diseases such as candidiasis, onychomycosis, Tinea captis/scal ringworm, Tinea corporis/body ringworm, Tinea cruris/jock itch, sporotrichosis and Tinea pedis/Athlete's foot, Meningitis such as Haemophilus influenza type b (Hib), Meningitis, viral, meningococcal infections and

pneumococcal infection, neglected tropical diseases such as Argentine haemorrhagic fever, Leishmaniasis, Nematode/roundworm infections, Ross river virus infection and West Nile virus (WNV) disease, Non-HIV STDs such as Trichomoniasis, Human papillomavirus (HPV) infections, sexually transmitted chlamydial diseases, Chancroid and Syphilis, Non-septic bacterial infections such as cellulitis, lyme disease, MRSA infection, pseudomonas, staphylococcal infections, Boutonneuse fever, Leptospirosis, Rheumatic fever, Botulism, Rickettsial disease and Mastoiditis, parasitic infections such as Cysticercosis, Echinococcosis, Trematode/Fluke infections, Trichinellosis, Babesiosis, Hypodermyiasis, Diphyllobothriasis and Trypanosomiasis, respiratory infections such as adenovirus infection, aspergillosis infections, avian (H5N1) influenza, influenza, RSV infections, severe acute respiratory syndrome (SARS), sinusitis, Legionellosis, Coccidioidomycosis and swine (H1N1) influenza, sepsis such as bacteraemia, sepsis/septic shock, sepsis in premature infants, urinary tract infection such as vaginal infections (bacterial), vaginal infections (fungal) and gonococcal infection, viral skin diseases such as B19 parvovirus infections, warts, genital herpes, orofacial herpes, shingles, inner ear infections, fetal cytomegalovirus syndrome, foodborn illnesses such as brucellosis (Brucella species), Clostridium perfringens (Epsilon toxin), E. Coli O157:H7 (Escherichia coli), Salmonellosis (Salmonella species), Shingellosis (Shingella), Vibriosis and Listeriosis, bioterrorism and potential epidemic diseases such as Ebola haemorrhagic fever, Lassa fever, Marburg haemorrhagic fever, plague, Anthrax Nipah virus disease, Hanta virus, Smallpox, Glanders (Burkholderia mallei), Melioidosis (Burkholderia pseudomallei), Psittacosis (Chlamydia psittaci), Q fever (Coxiella burnetii), Tularemia (Fancisella tularensis), rubella, mumps and polio.

The mRNA disclosed herein, may encode one or more validated or "in testing" therapeutic proteins or peptides. According to the present invention, one or more therapeutic proteins or peptides currently being marketed or in development may be encoded by the mRNA of the present invention. Therapeutic proteins and peptides encoded in the mRNA of the invention may be utilized to treat conditions or diseases in many therapeutic areas such as, but not limited to, blood, cardiovascular, CNS, poisoning (including antivenoms), dermatology, endocrinology, genetic, genitourinary, gastrointestinal, musculoskeletal, oncology, and immunology, respiratory, sensory and anti-infective.

The mRNA disclosed herein, may encode one or more cell-penetrating polypeptides. As used herein, “cell-penetrating polypeptide” or CPP refers to a polypeptide which may facilitate the cellular uptake of molecules. A cell-penetrating polypeptide of the present

invention may contain one or more detectable labels. The polypeptides may be partially labeled or completely labeled throughout. The mRNA may encode the detectable label completely, partially or not at all. The cell-penetrating peptide may also include a signal sequence. As used herein, a "signal sequence" refers to a sequence of amino acid residues bound at the amino terminus of a nascent protein during protein translation. The signal sequence may be used to signal the secretion of the cell-penetrating polypeptide.

In one embodiment, the mRNA may also encode a fusion protein. The fusion protein may be created by operably linking a charged protein to a therapeutic protein. As used herein, "operably linked" refers to the therapeutic protein and the charged protein being connected in such a way to permit the expression of the complex when introduced into the cell. As used herein, "charged protein" refers to a protein that carries a positive, negative or overall neutral electrical charge. Preferably, the therapeutic protein may be covalently linked to the charged protein in the formation of the fusion protein. The ratio of surface charge to total or surface amino acids may be approximately 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9.

The cell-penetrating polypeptide encoded by the mRNA may form a complex after being translated. The complex may comprise a charged protein linked, *e.g.*, covalently linked, to the cell-penetrating polypeptide.

In one embodiment, the cell-penetrating polypeptide may comprise a first domain and a second domain. The first domain may comprise a supercharged polypeptide. The second domain may comprise a protein-binding partner. As used herein, "protein-binding partner" includes, but is not limited to, antibodies and functional fragments thereof, scaffold proteins, or peptides. The cell-penetrating polypeptide may further comprise an intracellular binding partner for the protein-binding partner. The cell-penetrating polypeptide may be capable of being secreted from a cell where the mRNA may be introduced. The cell-penetrating polypeptide may also be capable of penetrating the first cell.

In one embodiment, the mRNA may encode a cell-penetrating polypeptide which may comprise a protein-binding partner. The protein binding partner may include, but is not limited to, an antibody, a supercharged antibody or a functional fragment. The mRNA may be introduced into the cell where a cell-penetrating polypeptide comprising the protein-binding partner is introduced.

Human and other eukaryotic cells are subdivided by membranes into many functionally distinct compartments. Each membrane-bound compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting

signals" which are amino acid motifs located within the protein, to target proteins to particular cellular organelles. One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER).

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane. While not wishing to be bound by theory, the molecules of the present invention may be used to exploit the cellular trafficking described above. As such, in some embodiments of the invention, mRNA are provided to express a secreted protein. In one embodiment, these may be used in the manufacture of large quantities of valuable human gene products.

In some embodiments of the invention, mRNA are provided to express a protein of the plasma membrane.

In some embodiments of the invention, mRNA are provided to express a cytoplasmic or cytoskeletal protein.

In some embodiments of the invention, mRNA are provided to express an intracellular membrane bound protein.

In some embodiments of the invention, mRNA are provided to express a nuclear protein.

In some embodiments of the invention, mRNA are provided to express a protein associated with human disease.

The mRNA may have a nucleotide sequence of a native or naturally occurring mRNA or encoding a native or naturally occurring peptide. Alternatively the mRNA may have a nucleotide sequence having a percent identity to the nucleotide sequence of a native or naturally occurring mRNA or mRNA may have a nucleotide sequence encoding a peptide having a percent identity to the nucleotide sequence of a native or naturally occurring peptide. The term "identity" as known in the art, refers to a relationship between the sequences of two or more peptides, as determined by comparing the sequences. In the art, identity also means the degree of sequence relatedness between peptides, as determined by the number of matches between strings of two or more amino acid residues. Identity measures the percent of identical matches between the smaller of two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer program

(i.e., "algorithms"). Identity of related peptides can be readily calculated by known methods. Such methods include, but are not limited to, those described in Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M. Stockton Press, New York, 1991; and Carillo et al., SIAM J. Applied Math. 48, 1073 (1988).

Thus, in some embodiments, the peptides encoded by the mRNAs of the multimeric structure are polypeptide variants that may have the same or a similar activity as a reference polypeptide. Alternatively, the variant may have an altered activity (*e.g.*, increased or decreased) relative to a reference polypeptide. Generally, variants of a particular polynucleotide or polypeptide of the invention will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% but less than 100% sequence identity to that particular reference polynucleotide or polypeptide as determined by sequence alignment programs and parameters described herein and known to those skilled in the art. Such tools for alignment include those of the BLAST suite (Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.) Other tools are described herein, specifically in the definition of “Identity.” Default parameters in the BLAST algorithm include, for example, an expect threshold of 10, Word size of 28, Match/Mismatch Scores 1, -2, Gap costs Linear. Any filter can be applied as well as a selection for species specific repeats, *e.g.*, Homo sapiens.

According to the present invention, the multimeric structures include mRNA to encode one or more polypeptides of interest or fragments thereof. A polypeptide of interest may include, but is not limited to, whole polypeptides, a plurality of polypeptides or fragments of polypeptides. As used herein, the term “polypeptides of interest" refer to any polypeptide which is selected to be encoded in the primary construct of the present invention. As used herein, “polypeptide" means a polymer of amino acid residues (natural or unnatural) linked together most often by peptide bonds. The term, as used herein, refers to proteins, polypeptides, and peptides of any size, structure, or function. In some instances the polypeptide encoded is smaller than about 50 amino acids and the polypeptide is then termed

a peptide. If the polypeptide is a peptide, it will be at least about 2, 3, 4, or at least 5 amino acid residues long. Thus, polypeptides include gene products, naturally occurring polypeptides, synthetic polypeptides, homologs, orthologs, paralogs, fragments and other equivalents, variants, and analogs of the foregoing. A polypeptide may be a single molecule or may be a multi-molecular complex such as a dimer, trimer or tetramer. They may also comprise single chain or multichain polypeptides such as antibodies or insulin and may be associated or linked. Most commonly disulfide linkages are found in multichain polypeptides. The term polypeptide may also apply to amino acid polymers in which one or more amino acid residues are an artificial chemical analogue of a corresponding naturally occurring amino acid.

The term “polypeptide variant" refers to molecules which differ in their amino acid sequence from a native or reference sequence. The amino acid sequence variants may possess substitutions, deletions, and/or insertions at certain positions within the amino acid sequence, as compared to a native or reference sequence. Ordinarily, variants will possess at least about 50% identity to a native or reference sequence, and preferably, they will be at least about 80%, more preferably at least about 90% identical to a native or reference sequence.

In some embodiments “variant mimics" are provided. As used herein, the term “variant mimic" is one which contains one or more amino acids which would mimic an activated sequence. For example, glutamate may serve as a mimic for phosphoro-threonine and/or phosphoro-serine. Alternatively, variant mimics may result in deactivation or in an inactivated product containing the mimic, *e.g.*, phenylalanine may act as an inactivating substitution for tyrosine; or alanine may act as an inactivating substitution for serine.

The present invention contemplates several types of compositions which are polypeptide based including variants and derivatives. These include substitutional, insertional, deletion and covalent variants and derivatives. The term “derivative” is used synonymously with the term “variant” but generally refers to a molecule that has been modified and/or changed in any way relative to a reference molecule or starting molecule.

As such, mRNA encoding polypeptides containing substitutions, insertions and/or additions, deletions and covalent modifications with respect to reference sequences, in particular the polypeptide sequences disclosed herein, are included within the scope of this invention. For example, sequence tags or amino acids, such as one or more lysines, can be added to the peptide sequences of the invention (*e.g.*, at the N-terminal or C-terminal ends). Sequence tags can be used for peptide purification or localization. Lysines can be used to

increase peptide solubility or to allow for biotinylation. Alternatively, amino acid residues located at the carboxy and amino terminal regions of the amino acid sequence of a peptide or protein may optionally be deleted providing for truncated sequences. Certain amino acids (*e.g.*, C-terminal or N-terminal residues) may alternatively be deleted depending on the use of the sequence, as for example, expression of the sequence as part of a larger sequence which is soluble, or linked to a solid support.

"Substitutional variants" when referring to polypeptides are those that have at least one amino acid residue in a native or starting sequence removed and a different amino acid inserted in its place at the same position. The substitutions may be single, where only one amino acid in the molecule has been substituted, or they may be multiple, where two or more amino acids have been substituted in the same molecule.

As used herein the term "conservative amino acid substitution" refers to the substitution of an amino acid that is normally present in the sequence with a different amino acid of similar size, charge, or polarity. Examples of conservative substitutions include the substitution of a non-polar (hydrophobic) residue such as isoleucine, valine and leucine for another non-polar residue. Likewise, examples of conservative substitutions include the substitution of one polar (hydrophilic) residue for another such as between arginine and lysine, between glutamine and asparagine, and between glycine and serine. Additionally, the substitution of a basic residue such as lysine, arginine or histidine for another, or the substitution of one acidic residue such as aspartic acid or glutamic acid for another acidic residue are additional examples of conservative substitutions. Examples of non-conservative substitutions include the substitution of a non-polar (hydrophobic) amino acid residue such as isoleucine, valine, leucine, alanine, methionine for a polar (hydrophilic) residue such as cysteine, glutamine, glutamic acid or lysine and/or a polar residue for a non-polar residue.

"Insertional variants" when referring to polypeptides are those with one or more amino acids inserted immediately adjacent to an amino acid at a particular position in a native or starting sequence. "Immediately adjacent" to an amino acid means connected to either the alpha-carboxy or alpha-amino functional group of the amino acid.

"Deletional variants" when referring to polypeptides are those with one or more amino acids in the native or starting amino acid sequence removed. Ordinarily, deletional variants will have one or more amino acids deleted in a particular region of the molecule.

"Covalent derivatives" when referring to polypeptides include modifications of a native or starting protein with an organic proteinaceous or non-proteinaceous derivatizing

agent, and/or post-translational modifications. Covalent modifications are traditionally introduced by reacting targeted amino acid residues of the protein with an organic derivatizing agent that is capable of reacting with selected side-chains or terminal residues, or by harnessing mechanisms of post-translational modifications that function in selected recombinant host cells. The resultant covalent derivatives are useful in programs directed at identifying residues important for biological activity, for immunoassays, or for the preparation of anti-protein antibodies for immunoaffinity purification of the recombinant glycoprotein. Such modifications are within the ordinary skill in the art and are performed without undue experimentation.

Certain post-translational modifications are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues may be present in the polypeptides produced in accordance with the present invention.

Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (T. E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)).

As used herein when referring to polypeptides the term "domain" refers to a motif of a polypeptide having one or more identifiable structural or functional characteristics or properties (*e.g.*, binding capacity, serving as a site for protein-protein interactions).

As used herein when referring to polypeptides the terms "site" as it pertains to amino acid based embodiments is used synonymously with "amino acid residue" and "amino acid side chain." A site represents a position within a peptide or polypeptide that may be modified, manipulated, altered, derivatized or varied within the polypeptide based molecules of the present invention.

As used herein the terms "termini" or "terminus" when referring to polypeptides refers to an extremity of a peptide or polypeptide. Such extremity is not limited only to the first or final site of the peptide or polypeptide but may include additional amino acids in the terminal regions. The polypeptide based molecules of the present invention may be characterized as having both an N-terminus (terminated by an amino acid with a free amino group (NH2)) and a C-terminus (terminated by an amino acid with a free carboxyl group

(COOH)). Proteins of the invention are in some cases made up of multiple polypeptide chains brought together by disulfide bonds or by non-covalent forces (multimers, oligomers). These sorts of proteins will have multiple N- and C-termini. Alternatively, the termini of the polypeptides may be modified such that they begin or end, as the case may be, with a non-polypeptide based moiety such as an organic conjugate.

Once any of the features have been identified or defined as a desired component of a polypeptide to be encoded by the mRNA of the invention, any of several manipulations and/or modifications of these features may be performed by moving, swapping, inverting, deleting, randomizing or duplicating. Furthermore, it is understood that manipulation of features may result in the same outcome as a modification to the molecules of the invention. For example, a manipulation which involved deleting a domain would result in the alteration of the length of a molecule just as modification of a nucleic acid to encode less than a full length molecule would.

Modifications and manipulations can be accomplished by methods known in the art such as, but not limited to, site directed mutagenesis. The resulting modified molecules may then be tested for activity using in vitro or in vivo assays such as those described herein or any other suitable screening assay known in the art.

The present invention provides multimeric structures and pharmaceutical compositions thereof optionally in combination with one or more pharmaceutically acceptable excipients. Pharmaceutical compositions may optionally comprise one or more additional active substances, *e.g.*, therapeutically and/or prophylactically active substances. Pharmaceutical compositions of the present invention may be sterile and/or pyrogen-free. General considerations in the formulation and/or manufacture of pharmaceutical agents may be found, for example, in Remington: The Science and Practice of Pharmacy 21st ed., Lippincott Williams & Wilkins, 2005 (incorporated herein by reference in its entirety).

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 the multimeric structures or the polynucleotides contained therein, *e.g.*, mRNA encoding polynucleotides to be delivered as described herein.

Formulations of the pharmaceutical compositions described herein may 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 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 invention will vary, depending upon the identity, size, and/or condition of the subject treated and further depending upon the route by which the composition is to be administered. By way of example, the composition may comprise between 0.1% and 100%, *e.g.*, between 0.5 and 50%, between 1-30%, between 5-80%, at least 80% (w/w) active ingredient.

The multimeric structures of the invention can be formulated using one or more excipients to: (1) increase stability; (2) increase cell transfection; (3) permit the sustained or delayed release (*e.g.*, from a depot formulation); (4) alter the biodistribution (*e.g.*, target 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 invention can include, without limitation, lipidoids, liposomes, lipid nanoparticles, polymers, lipoplexes, core-shell nanoparticles, peptides, proteins, cells transfected with multimeric structures, hyaluronidase, nanoparticle mimics and combinations thereof.

The instant invention is based, in part, on the surprising discovery that non-covalent bonding between untranslated regions of nucleic acids (*e.g.*, mRNAs, or IVT mRNAs) allows formation of multimeric molecules and efficient encapsulation of said molecules by lipid nanoparticles (LNPs). In some embodiments, multimeric nucleic acid molecules of the invention (*e.g.*, multimeric mRNA molecules) can be formulated using one or more liposomes, lipoplexes, or lipid nanoparticles. In one embodiment, pharmaceutical compositions of multimeric nucleic acid molecules include lipid nanoparticles (LNPs). In some embodiments, lipid nanoparticles are MC3-based lipid nanoparticles.

The number of multimeric molecules encapsulated by a lipid nanoparticle ranges from about 1 multimeric molecule to about 100 multimeric molecules. In some embodiments, he number of multimeric molecules encapsulated by a lipid nanoparticle ranges from about 50 multimeric molecules to about 500 multimeric molecules. In some embodiments, the number of multimeric molecules encapsulated by a lipid nanoparticle ranges from about 250 multimeric molecules to about 1000 multimeric molecules. In some embodiments, the

number of multimeric molecules encapsulated by a lipid nanoparticle is greater than 1000 multimeric molecules.

In one embodiment, the multimeric structures may be formulated in a lipid-polycation complex. The formation of the lipid-polycation complex may be accomplished by methods known in the art. As a non-limiting example, the polycation may include a cationic peptide or a polypeptide such as, but not limited to, polylysine, polyornithine and/or polyarginine. In another embodiment, the multimeric structures may be formulated in a lipid-polycation complex which may further include a non-cationic lipid such as, but not limited to, cholesterol or dioleoyl phosphatidylethanolamine (DOPE).

The liposome formulation may be influenced by, but not limited to, the selection of the cationic lipid component, the degree of cationic lipid saturation, the nature of the PEGylation, ratio of all components and biophysical parameters such as size. In one example by Semple et al. (Semple et al. Nature Biotech. 2010 28:172-176; herein incorporated by reference in its entirety), the liposome formulation was composed of 57.1 % cationic lipid, 7.1% dipalmitoylphosphatidylcholine, 34.3 % cholesterol, and 1.4% PEG-c-DMA. As another example, changing the composition of the cationic lipid could more effectively deliver siRNA to various antigen presenting cells (Basha et al. Mol Ther. 2011 19:2186-2200; herein incorporated by reference in its entirety). In some embodiments, liposome formulations may comprise from about 35 to about 45% cationic lipid, from about 40% to about 50% cationic lipid, from about 50% to about 60% cationic lipid and/or from about 55% to about 65% cationic lipid. In some embodiments, the ratio of lipid to mRNA in liposomes may be from about 5:1 to about 20:1, from about 10:1 to about 25:1, from about 15:1 to about 30:1 and/or at least 30:1.

In some embodiments, the ratio of PEG in the lipid nanoparticle (LNP) formulations may be increased or decreased and/or the carbon chain length of the PEG lipid may be modified from C14 to C18 to alter the pharmacokinetics and/or biodistribution of the LNP formulations. As a non-limiting example, LNP formulations may contain from about 0.5% to about 3.0%, from about 1.0% to about 3.5%, from about 1.5% to about 4.0%, from about 2.0% to about 4.5%, from about 2.5% to about 5.0% and/or from about 3.0% to about 6.0% of the lipid molar ratio of PEG-c-DOMG (R-3-[(ω-methoxy-poly(ethyleneglycol)2000)carbamoyl)]-1,2-dimyristyloxypropyl-3-amine) (also referred to herein as PEG-DOMG) as compared to the cationic lipid, DSPC and cholesterol. In another embodiment the PEG-c-DOMG may be replaced with a PEG lipid such as, but not limited to,

PEG- DSG (1,2-Distearoyl-sn-glycerol, methoxypolyethylene glycol), PEG-DMG (1,2-Dimyristoyl-sn-glycerol) and/or PEG-DPG (1,2-Dipalmitoyl-sn-glycerol, methoxypolyethylene glycol). The cationic lipid may be selected from any lipid known in the art such as, but not limited to, DLin-MC3-DMA, DLin-DMA, C12-200 and DLin-KC2-DMA.

In one embodiment, the multimeric structures is formulated in a nanoparticle which may comprise at least one lipid. The lipid may be selected from, but is not limited to, DLin-DMA, DLin-K-DMA, 98N12-5, C12-200, DLin-MC3-DMA, DLin-KC2-DMA, DODMA, PLGA, PEG, PEG-DMG, PEGylated lipids and amino alcohol lipids. In another aspect, the lipid may be a cationic lipid such as, but not limited to, DLin-DMA, DLin-D-DMA, DLin-MC3-DMA, DLin-KC2-DMA, DODMA and amino alcohol lipids. The amino alcohol cationic lipid may be the lipids described in and/or made by the methods described in US Patent Publication No. US20130150625, herein incorporated by reference in its entirety. As a non-limiting example, the cationic lipid may be 2-amino-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]-2-{[(9Z,2Z)-octadeca-9,12-dien-1-yloxy]methyl}propan-1-ol (Compound 1 in US20130150625); 2-amino-3-[(9Z)-octadec-9-en-1-yloxy]-2-{[(9Z)-octadec-9-en-1-yloxy]methyl}propan-1-ol (Compound 2 in US20130150625); 2-amino-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]-2-[(octyloxy)methyl]propan-1-ol (Compound 3 in US20130150625); and 2-(dimethylamino)-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]-2-{[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]methyl}propan-1-ol (Compound 4 in US20130150625); or any pharmaceutically acceptable salt or stereoisomer thereof.

Lipid nanoparticle formulations typically comprise a lipid, in particular, an ionizable cationic lipid, for example, 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), or di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), and further comprise a neutral lipid, a sterol and a molecule capable of reducing particle aggregation, for example a PEG or PEG-modified lipid.

In one embodiment, the lipid nanoparticle formulation consists essentially of (i) at least one lipid selected from the group consisting of 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319); (ii) a neutral lipid selected from DSPC, DPPC, POPC, DOPE and SM; (iii) a sterol,

e.g., cholesterol; and (iv) a PEG-lipid, *e.g.*, PEG-DMG or PEG-cDMA, in a molar ratio of about 20-60% cationic lipid: 5-25% neutral lipid: 25-55% sterol; 0.5-15% PEG-lipid.

In one embodiment, the formulation includes from about 25% to about 75% on a molar basis of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), *e.g.*, from about 35 to about 65%, from about 45 to about 65%, about 60%, about 57.5%, about 50% or about 40% on a molar basis.

In one embodiment, the formulation includes from about 0.5% to about 15% on a molar basis of the neutral lipid *e.g.*, from about 3 to about 12%, from about 5 to about 10% or about 15%, about 10%, or about 7.5% on a molar basis. Exemplary neutral lipids include, but are not limited to, DSPC, POPC, DPPC, DOPE and SM. In one embodiment, the formulation includes from about 5% to about 50% on a molar basis of the sterol (*e.g.*, about 15 to about 45%, about 20 to about 40%, about 40%, about 38.5%, about 35%, or about 31% on a molar basis. An exemplary sterol is cholesterol. In one embodiment, the formulation includes from about 0.5% to about 20% on a molar basis of the PEG or PEG-modified lipid (*e.g.*, about 0.5 to about 10%, about 0.5 to about 5%, about 1.5%, about 0.5%, about 1.5%, about 3.5%, or about 5% on a molar basis. In one embodiment, the PEG or PEG modified lipid comprises a PEG molecule of an average molecular weight of 2,000 Da. In other embodiments, the PEG or PEG modified lipid comprises a PEG molecule of an average molecular weight of less than 2,000, for example around 1,500 Da, around 1,000 Da, or around 500 Da. Exemplary PEG-modified lipids include, but are not limited to, PEG-distearoyl glycerol (PEG-DMG) (also referred herein as PEG-C14 or C14-PEG), PEG-cDMA.

In one embodiment, the formulations of the inventions include 25-75% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), 0.5-15% of the neutral lipid, 5-50% of the sterol, and 0.5-20% of the PEG or PEG-modified lipid on a molar basis.

In one embodiment, the formulations of the inventions include 35-65% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), 3-12% of the neutral lipid, 15-45% of the sterol, and 0.5-10% of the PEG or PEG-modified lipid on a molar basis.

In one embodiment, the formulations of the inventions include 45-65% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), 5-10% of the neutral lipid, 25-40% of the sterol, and 0.5-10% of the PEG or PEG-modified lipid on a molar basis.

In one embodiment, the formulations of the inventions include about 60% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), about 7.5% of the neutral lipid, about 31 % of the sterol, and about 1.5% of the PEG or PEG-modified lipid on a molar basis.

In one embodiment, the formulations of the inventions include about 50% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), about 10% of the neutral lipid, about 38.5 % of the sterol, and about 1.5% of the PEG or PEG-modified lipid on a molar basis.

In one embodiment, the formulations of the inventions include about 50% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), about 10% of the neutral lipid, about 35 % of the sterol, about 4.5% or about 5% of the PEG or PEG-modified lipid, and about 0.5% of the targeting lipid on a molar basis.

In one embodiment, the formulations of the inventions include about 40% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), about 15% of the neutral lipid, about 40% of the sterol, and about 5% of the PEG or PEG-modified lipid on a molar basis.

In one embodiment, the formulations of the inventions include about 57.2% of a cationic lipid selected from 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319), about 7.1% of the

neutral lipid, about 34.3% of the sterol, and about 1.4% of the PEG or PEG-modified lipid on a molar basis.

In one embodiment, the formulations of the inventions include about 57.5% of a cationic lipid selected from the PEG lipid is PEG-cDMA (PEG-cDMA is further discussed in Reyes et al. (J. Controlled Release, 107, 276-287 (2005), the contents of which are herein incorporated by reference in its entirety), about 7.5% of the neutral lipid, about 31.5 % of the sterol, and about 3.5% of the PEG or PEG-modified lipid on a molar basis.

In preferred embodiments, lipid nanoparticle formulation consists essentially of a lipid mixture in molar ratios of about 20-70% cationic lipid: 5-45% neutral lipid: 20-55% cholesterol: 0.5-15% PEG-modified lipid; more preferably in a molar ratio of about 20-60% cationic lipid: 5-25% neutral lipid: 25-55% cholesterol: 0.5-15% PEG-modified lipid.

In particular embodiments, the molar lipid ratio is approximately 50/10/38.5/1.5 (mol% cationic lipid/neutral lipid, *e.g.*, DSPC/Chol/PEG-modified lipid, *e.g.*, PEG-DMG, PEG-DSG or PEG-DPG), 57.2/7.1134.3/1.4 (mol% cationic lipid/ neutral lipid, *e.g.*, DPPC/Chol/ PEG-modified lipid, *e.g.*, PEG-cDMA), 40/15/40/5 (mol% cationic lipid/ neutral lipid, *e.g.*, DSPC/Chol/ PEG-modified lipid, *e.g.*, PEG-DMG), 50/10/35/4.5/0.5 (mol% cationic lipid/ neutral lipid, *e.g.*, DSPC/Chol/ PEG-modified lipid, *e.g.*, PEG-DSG), 50/10/35/5 (cationic lipid/ neutral lipid, *e.g.*, DSPC/Chol/ PEG-modified lipid, *e.g.*, PEG-DMG), 40/10/40/10 (mol% cationic lipid/ neutral lipid, *e.g.*, DSPC/Chol/ PEG-modified lipid, *e.g.*, PEG-DMG or PEG-cDMA), 35/15/40/10 (mol% cationic lipid/ neutral lipid, *e.g.*, DSPC/Chol/ PEG-modified lipid, *e.g.*, PEG-DMG or PEG-cDMA) or 52/13/30/5 (mol% cationic lipid/ neutral lipid, *e.g.*, DSPC/Chol/ PEG-modified lipid, *e.g.*, PEG-DMG or PEG-cDMA).

Exemplary lipid nanoparticle compositions and methods of making same are described, for example, in Semple et al. (2010) Nat. Biotechnol. 28:172-176; Jayarama et al. (2012), Angew. Chem. Int. Ed., 51: 8529–8533; and Maier et al. (2013) Molecular Therapy 21, 1570-1578 (the contents of each of which are incorporated herein by reference in their entirety).

In one embodiment, the lipid nanoparticle formulations described herein may comprise a cationic lipid, a PEG lipid and a structural lipid and optionally comprise a non-cationic lipid. As a non-limiting example, the lipid nanoparticle may comprise about 40-60% of cationic lipid, about 5-15% of a non-cationic lipid, about 1-2% of a PEG lipid and about 30-50% of a structural lipid. As another non-limiting example, the lipid nanoparticle may

comprise about 50% cationic lipid, about 10% non-cationic lipid, about 1.5% PEG lipid and about 38.5% structural lipid. As yet another non-limiting example, the lipid nanoparticle may comprise about 55% cationic lipid, about 10% non-cationic lipid, about 2.5% PEG lipid and about 32.5% structural lipid. In one embodiment, the cationic lipid may be any cationic lipid described herein such as, but not limited to, DLin-KC2-DMA, DLin-MC3-DMA and L319.

In one embodiment, the lipid nanoparticle formulations described herein may be 4 component lipid nanoparticles. The lipid nanoparticle may comprise a cationic lipid, a non-cationic lipid, a PEG lipid and a structural lipid. As a non-limiting example, the lipid nanoparticle may comprise about 40-60% of cationic lipid, about 5-15% of a non-cationic lipid, about 1-2% of a PEG lipid and about 30-50% of a structural lipid. As another non-limiting example, the lipid nanoparticle may comprise about 50% cationic lipid, about 10% non-cationic lipid, about 1.5% PEG lipid and about 38.5% structural lipid. As yet another non-limiting example, the lipid nanoparticle may comprise about 55% cationic lipid, about 10% non-cationic lipid, about 2.5% PEG lipid and about 32.5% structural lipid. In one embodiment, the cationic lipid may be any cationic lipid described herein such as, but not limited to, DLin-KC2-DMA, DLin-MC3-DMA and L319.

In one embodiment, the lipid nanoparticle formulations described herein may comprise a cationic lipid, a non-cationic lipid, a PEG lipid and a structural lipid. As a non-limiting example, the lipid nanoparticle comprise about 50% of the cationic lipid DLin-KC2-DMA, about 10% of the non-cationic lipid DSPC, about 1.5% of the PEG lipid PEG-DOMG and about 38.5% of the structural lipid cholesterol. As a non-limiting example, the lipid nanoparticle comprise about 50% of the cationic lipid DLin-MC3-DMA, about 10% of the non-cationic lipid DSPC, about 1.5% of the PEG lipid PEG-DOMG and about 38.5% of the structural lipid cholesterol. As a non-limiting example, the lipid nanoparticle comprise about 50% of the cationic lipid DLin-MC3-DMA, about 10% of the non-cationic lipid DSPC, about 1.5% of the PEG lipid PEG-DMG and about 38.5% of the structural lipid cholesterol. As yet another non-limiting example, the lipid nanoparticle comprise about 55% of the cationic lipid L319, about 10% of the non-cationic lipid DSPC, about 2.5% of the PEG lipid PEG-DMG and about 32.5% of the structural lipid cholesterol.

In one embodiment, the multimeric molecules (*e.g.*, multimeric mRNA molecules) of the invention may be formulated in lipid nanoparticles having a diameter from about 10 to about 100 nm such as, but not limited to, about 10 to about 20 nm, about 10 to about 30 nm, about 10 to about 40 nm, about 10 to about 50 nm, about 10 to about 60 nm, about 10 to

about 70 nm, about 10 to about 80 nm, about 10 to about 90 nm, about 20 to about 30 nm, about 20 to about 40 nm, about 20 to about 50 nm, about 20 to about 60 nm, about 20 to about 70 nm, about 20 to about 80 nm, about 20 to about 90 nm, about 20 to about 100 nm, about 30 to about 40 nm, about 30 to about 50 nm, about 30 to about 60 nm, about 30 to about 70 nm, about 30 to about 80 nm, about 30 to about 90 nm, about 30 to about 100 nm, about 40 to about 50 nm, about 40 to about 60 nm, about 40 to about 70 nm, about 40 to about 80 nm, about 40 to about 90 nm, about 40 to about 100 nm, about 50 to about 60 nm, about 50 to about 70 nm about 50 to about 80 nm, about 50 to about 90 nm, about 50 to about 100 nm, about 60 to about 70 nm, about 60 to about 80 nm, about 60 to about 90 nm, about 60 to about 100 nm, about 70 to about 80 nm, about 70 to about 90 nm, about 70 to about 100 nm, about 80 to about 90 nm, about 80 to about 100 nm and/or about 90 to about 100 nm.

In one embodiment, the lipid nanoparticles may have a diameter from about 10 to 500 nm. In one embodiment, the lipid nanoparticle may have a diameter greater than 100 nm, greater than 150 nm, greater than 200 nm, greater than 250 nm, greater than 300 nm, greater than 350 nm, greater than 400 nm, greater than 450 nm, greater than 500 nm, greater than 550 nm, greater than 600 nm, greater than 650 nm, greater than 700 nm, greater than 750 nm, greater than 800 nm, greater than 850 nm, greater than 900 nm, greater than 950 nm or greater than 1000 nm. In some embodiments, the cationic lipid nanoparticle has a mean diameter of 50-150 nm. In some embodiments, the cationic lipid nanoparticle has a mean diameter of 80-100 nm.

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 may 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 may comprise between 0.1% and 99% (w/w) of the active ingredient. By way of example, the composition may comprise between 0.1% and 100%, *e.g.*, between .5 and 50%, between 1-30%, between 5-80%, at least 80% (w/w) active ingredient.

In one embodiment, the compositions containing the multimeric structures may comprise the multimeric polynucleotides described herein, formulated in a lipid nanoparticle comprising MC3, Cholesterol, DSPC and PEG2000-DMG, the buffer trisodium citrate, sucrose and water for injection. As a non-limiting example, the composition comprises: 2.0 mg/mL of drug substance (*e.g.*, multimeric polynucleotides), 21.8 mg/mL of MC3, 10.1

mg/mL of cholesterol, 5.4 mg/mL of DSPC, 2.7 mg/mL of PEG2000-DMG, 5.16 mg/mL of trisodium citrate, 71 mg/mL of sucrose and about 1.0 mL of water for injection.

The multimeric structures of the present invention may be administered by any route which results in a therapeutically effective outcome. The present invention provides methods comprising administering multimeric structures and in accordance with the invention to a subject in need thereof. The exact amount required will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the disease, the particular composition, its mode of administration, its mode of activity, and the like. Compositions in accordance with the invention are typically formulated in dosage unit form for ease of administration and uniformity of dosage. It will be understood, however, that the total daily usage of the compositions of the present invention may be decided by the attending physician within the scope of sound medical judgment. The specific therapeutically effective, prophylactically effective, or appropriate imaging dose level for any particular patient will depend upon a variety of factors including the disorder being treated and the severity of the disorder; the activity of the specific compound employed; the specific composition employed; the age, body weight, general health, sex and diet of the patient; the time of administration, route of administration, and rate of excretion of the specific compound employed; the duration of the treatment; drugs used in combination or coincidental with the specific compound employed; and like factors well known in the medical arts.

In certain embodiments, compositions in accordance with the present invention may be administered at dosage levels sufficient to deliver from about 0.0001 mg/kg to about 100 mg/kg, from about 0.001 mg/kg to about 0.05 mg/kg, from about 0.005 mg/kg to about 0.05 mg/kg, from about 0.001 mg/kg to about 0.005 mg/kg, from about 0.05 mg/kg to about 0.5 mg/kg, from about 0.01 mg/kg to about 50 mg/kg, from about 0.1 mg/kg to about 40 mg/kg, from about 0.5 mg/kg to about 30 mg/kg, from about 0.01 mg/kg to about 10 mg/kg, from about 0.1 mg/kg to about 10 mg/kg, or from about 1 mg/kg to about 25 mg/kg, of subject body weight per day, one or more times a day, to obtain the desired therapeutic, diagnostic, prophylactic, or imaging. The desired dosage may be delivered three times a day, two times a day, once a day, every other day, every third day, every week, every two weeks, every three weeks, or every four weeks. In certain embodiments, the desired dosage may be delivered using multiple administrations (*e.g.*, two, three, four, five, six, seven, eight, nine, ten, eleven,

twelve, thirteen, fourteen, or more administrations). When multiple administrations are employed, split dosing regimens such as those described herein may be used.

A multimeric structure pharmaceutical composition described herein can be formulated into a dosage form described herein, such as an intranasal, intratracheal, or injectable (*e.g.*, intravenous, intraocular, intravitreal, intramuscular, intradermal, intracardiac, intraperitoneal, and subcutaneous).

The present invention provides pharmaceutical compositions including multimeric molecules (*e.g.*, multimeric mRNA molecules) and multimeric molecule compositions and/or complexes optionally in combination with one or more pharmaceutically acceptable excipients.

The present invention provides multimeric molecules (*e.g.*, multimeric mRNA molecules) and related pharmaceutical compositions and complexes optionally in combination with one or more pharmaceutically acceptable excipients. Pharmaceutical compositions may optionally comprise one or more additional active substances, *e.g.*, therapeutically and/or prophylactically active substances. Pharmaceutical compositions of the present invention may be sterile and/or pyrogen-free. General considerations in the formulation and/or manufacture of pharmaceutical agents may be found, for example, in Remington: The Science and Practice of Pharmacy 21st ed., Lippincott Williams & Wilkins, 2005 (incorporated herein by reference in its entirety).

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 the multimeric molecules (*e.g.*, multimeric mRNA molecules), 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 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, including commercially relevant

mammals such as cattle, pigs, horses, sheep, cats, dogs, mice, and/or rats; and/or birds, including commercially relevant birds such as poultry, chickens, ducks, geese, and/or turkeys.

Formulations of the pharmaceutical compositions described herein may 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 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 invention will vary, depending upon the identity, size, and/or condition of the subject treated and further depending upon the route by which the composition is to be administered. By way of example, the composition may comprise between 0.1% and 100%, *e.g.*, between 0.5 and 50%, between 1-30%, between 5-80%, at least 80% (w/w) active ingredient.

This invention is not limited in its application to the details of construction and the arrangement of components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments and of being practiced or of being carried out in various ways. Also, the phraseology and terminology used herein is for the purpose of description and should not be regarded as limiting. The use of "including," "comprising," or "having," "containing," "involving," and variations thereof herein, is meant to encompass the items listed thereafter and equivalents thereof as well as additional items.

EXAMPLES

Example 1

Conventionally, when two or more nucleic acid molecules are formulated into lipid nanoparticles (LNP), the resultant LNPs are a heterogeneous population: some LNPs are empty, some LNPs contain just only a single mRNA species, and some contain both or all mRNA species (Fig. 1). This heterogeneous population of LNPs may be unpredictable; for example, in an equimolar amount of two differently-labeled mRNAs, 47% of the resulting LNPs showed both types of mRNA (Fig. 2).

To obtain a more predictable and uniform distribution of mRNA within an LNP, a splint-assisted multimeric mRNA molecule was first created (Fig. 3A). Two mRNA molecules were linked via a short nucleic acid "splint" that hybridizes to the 5' end of each

mRNA molecule via non-covalent bonding (*e.g.*, hydrogen bonding between complementary nucleotide bases). The formation of a multimeric molecule was confirmed via gel electrophoresis (Fig. 3B).

It was found that the two mRNAs could also be directly tethered via non-covalent bonds between the 5' untranslated region (UTRs) (Figs. 4A-4B). In Figure 4, one mRNA consisted of a 7-methylguanosine cap with a triphosphate link to 47 nucleotides (nt) of the 5'UTR followed by the coding region, a 119 nt 3'UTR and a polyadenosine (polyA) track of 100 nucleotides. To form the dimer, the 5'UTR were modified to allow for complementary pairing or hydrogen bond formation between the bases. In the dimer, the entire 47 nucleotide 5'UTR was designed as a reverse complement; the ratio of G-C- paring was 36% and A-U was 64%. See for example, Fig. 16, which shows the nucleic acid sequences of dimer-forming mRNA molecules (*e.g.*, mRNA #1 (SEQ ID NO:1) and mRNA #2 (SEQ ID NO: 2)). In another embodiment, 45 nucleotide complementary base pairing was shown to be sufficient to tether two different mRNAs to one another.

Self-assembled multimeric mRNA co-translation in JAWSII monocytes showed the successful co-translation of the two mRNAs (Fig. 5). The tethering process was also shown to have a much higher efficiency of co-translation compared to monomeric mRNAs (Fig. 6). This result was further confirmed by FACS (Figs. 10A-10C). Furthermore the lack of change in eGFP fluorescence between monomeric mRNA and multimeric mRNA indicates that formation of multimeric mRNA complex does not interfere with mRNA translation and protein expression, irrespective of multimeric complex dosage (Figs. 7-9 and Fig. 11).

To tether three different mRNAs together (trimer formation), the 5'UTR of the mRNAs was designed such that each mRNA hybridizes to two different mRNAs (Fig. 12A and Fig. 17). mRNAs #1 (SEQ ID NO: 1) and #2 (SEQ ID NO: 2) share 22 nts with G-C paring of 36% and A-U pairing of 64%; whereas mRNAs #1 (SEQ ID NO: 1) and #3 (SEQ ID NO: 3) share 20 nts with G-C pairing of 30% and A-U pairing of 70%. Finally, mRNAs #2 (SEQ ID NO: 2) and #3 (SEQ ID NO: 3) share 24 nts with 50% pairing for both G-C and A-U. Bioanalyzer gel analysis confirmed the formation of the trimer (Fig. 12B). Using fluorescence microscopy, co-translation of all three mRNAs from the trimeric molecule was confirmed (Fig. 13).

As with the trimer design, to tether four different mRNAs together (*i.e.*, tetramer formation), the 5'UTR of mRNAs was designed such that each mRNA hybridizes to two different mRNAs as shown in Figure 14A. mRNAs share between 19nts-24nts with G-C

pairing ranging from 48%-58% and A-U pairing of 42%-52% (see Fig. 18). Bioanalyzer gel analysis confirmed the formation of the self-assembling tetrameric mRNA molecule (Fig. 14B).

The multimeric mRNA complexes were formed under a heating and stepwise cooling protocol. A mixture of 5 μM of each mRNA desired to be incorporated into the multimeric complex was placed in a buffer containing 50 mM 2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris) pH 7.5, 150 mM sodium chloride (NaCl), and 1 mM ethylene-diamine-tetra-acetic acid (EDTA) and was heated to 65°C for 5 minutes, 60°C for 5 minutes, 40°C for 2 minutes, and then cooled to 4°C for 10 minutes, resulting in the multimeric complex.

The multimeric mRNA complexes were then formulated in LNPs and were shown to successfully co-localize in the LNP. An mRNA containing Alexa488 was tethered to an mRNA containing Alexa 647, forming a self-assembling dimeric mRNA molecule. The dimeric molecule was then formulated into an LNP and imaged (Fig. 15). The overlay image (right) illustrates the co-localization of the two mRNAs in the same LNP at a higher efficiency than LNPs loaded using the conventional method of an equimolar mixture of monomeric mRNAs.

Example 2

A multi-mRNA complex of eGFP and mCherry was synthesized. The diameter, PD index and percent encapsulation of each mRNA complex tested are shown in Table 1.

Table 1

Name	**Group**	**Sample**	**Diameter (nm)**	**PD Index**	**Encapsulation (%)**
PBS	1	PBS			
eGFP (alone)	2	180082	109.4	0.15	96
mCherry (alone)	3	180089	91.1	0.11	99
eGFP+mCherry (nondimerized)	4	82+83	85.5	0.084	98
eGFP+mCherry (dimerized)	5	82+89	91.3	0.089	98

These complexes were used to assess GFP and mCherry protein expression in mouse livers using both immunohistochemistry (IHC) and immunofluorescence (IF) assays. GFP and

mCherry were tested in singleplex and co-localized to determine percent co-localization in multi-mRNA complexes. The dosing regimen used for the IHC and IF assays for each complex are given in Table 2.

Table 2

Group	Test/Control Material	Vehicle	Formulation	Route	Dosing Regimen	# of Doses	# of males	# of females	Dose Level (mg/kg)
1	PBS	Buffer	N/A	IV	Single	1	0	3	0
2	eGFP (alone)	MC3	100nm particle	IV	Single	1	0	3	2
3	mCherry (alone)	MC3	100nm particle	IV	Single	1	0	3	2
4	eGFP+ mCherry (non-dimerized) +83	MC3	100nm particle	IV	Single	1	0	3	2
5	eGFP+ mCherry (dimerized) +89	MC3	100nm particle	IV	Single	1	0	3	2

The formation of the self-assembling dimer multi-mRNA complexes comprising eGFP and mCherry was confirmed via gel electrophoresis (Fig. 19). Protein expression of both GFP and mCherry was confirmed through IHC after intravenous (IV) administration of PBS as a control or GFP (1:1500; Abcam, ab290) or mCherry (1:800; Abcam, 1C51). Both GFP and mCherry protein expression were observed in the cytoplasm of hepatocytes (Figs. 20 and 21).

GFP and mCherry antibodies for both IF and IHC assays in mouse livers were optimized. Representative microscopy data for GFP (Fig. 22A) and mCherry/GFP (Fig. 22B) are shown.

A dual mCherry and GFP IHC assay was performed to determine protein expression. The IHC results indicate that mCherry shows a 12-fold increase in protein expression over the vehicle group in the non-dimerized complex and a 6-fold increase with the dimerized

complex (Figs. 23 and 26B). The IHC results also indicate that GFP shows a 19-fold increase in protein expression over the vehicle group in the non-dimerized complex and a 14-fold increase with the dimerized complex (Figs. 24 and 26C). Together, the GFP-mCherry complex showed significantly increased expression over the vehicle in both the dimerized and non-dimerized forms (Figs. 25, 26D, and 26E).

Singleplex assays were also performed to examine protein expression in the different groups and to demonstrate the specificity of the binding. Staining for GFP was positive in the GFP group (Fig. 27B), the eGFP + mCherry (non-dimerized) group (Fig. 27D), and the eGFP + mCherry (dimerized) group (Fig. 27E) and negative in the mCherry group (Fig. 27C). Staining for mCherry was positive in the mCherry group (Fig. 28C) and negative in the GFP group (Fig. 28B).

Co-localization of mCherry and GFP was determined using an IF assay (FIG. 29 and FIG. 30). The non-dimerized complex shows 60% co-localization of mCherry and GFP protein expression and 23% co-localization in the dimerized complex.

EQUIVALENTS

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 claims.

All references, including patent documents, disclosed herein are incorporated by reference in their entirety.

What is claimed is:

CLAIMS

1. A lipid nanoparticle composition comprising a first nucleic acid and a second nucleic acid, wherein the first and second nucleic acids are uniformly distributed throughout the lipid nanoparticle and wherein the first and second nucleic acid molecules are not covalently linked to one another.

2 The lipid nanoparticle of claim 1, wherein the first nucleic acid is linked to the second nucleic acid by a non-covalent bond, and wherein the non-covalent bond is located between a first non-coding region of the first nucleic acid and a second non-coding region of the second nucleic acid.

3. The lipid nanoparticle of claim 1, wherein the first nucleic acid and the second nucleic acid are RNA molecules.

4. The lipid nanoparticle of claim 3, wherein the RNA molecules are mRNA molecules.

5. The lipid nanoparticle of claim 4, wherein the mRNA molecules are *in vitro* transcribed mRNA molecules (IVT mRNA).

6. The lipid nanoparticle of any one of claims 1 to 5, wherein the first non-coding region and/or the second non-coding region is an untranslated region (UTR).

7. The lipid nanoparticle of claim 6, wherein the UTR is a 5'UTR.

8. The lipid nanoparticle of any one of claims 1 to 7, wherein the first nucleic acid is linked to the second nucleic acid by 2, 3, 4, 5, or 6 non-covalent bonds.

9. The lipid nanoparticle of claim 8, wherein the first nucleic acid is linked to the second nucleic acid by at least 10 or at least 20 non-covalent bonds.

10. The lipid nanoparticle of claim 8 or 9, wherein the non-covalent bonds are formed between complementary nucleotide bases of the first nucleic acid and the second nucleic acid.

11. The lipid nanoparticle of claim 10, wherein the complementary nucleotide bases of the first nucleic acid and the second nucleic acid have a G-C pairing ratio in a range of from about 30% to about 60%.

12. The lipid nanoparticle of any one of claims 1 to 11, further comprising a third nucleic acid, wherein the third nucleic acid is linked to the first nucleic acid and/or the second nucleic acid by a non-covalent bond, wherein the non-covalent bond is located in a third non-coding region of the third nucleic acid.

13. The lipid nanoparticle of any one of claims 1 to 12, wherein the lipid nanoparticle comprises a cationic lipid, a PEG-modified lipid, a sterol and a non-cationic lipid.

14. The lipid nanoparticle of claim 13, wherein the cationic lipid is selected from the group consisting of 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), dilinoleyl-methyl-4-dimethylaminobutyrate (DLin-MC3-DMA), and di((Z)-non-2-en-1-yl) 9-((4-(dimethylamino)butanoyl)oxy)heptadecanedioate (L319).

15. The lipid nanoparticle of claim 13, wherein the lipid nanoparticle has a molar ratio of about 20-60% cationic lipid: about 5-25% non-cationic lipid: about 25-55% sterol; and about 0.5-15% PEG-modified lipid.

16. The lipid nanoparticle of claim 13, wherein the lipid nanoparticle comprises a molar ratio of about 50% cationic lipid, about 1.5% PEG-modified lipid, about 38.5% cholesterol and about 10% non-cationic lipid.

17. The lipid nanoparticle of any one of claims 13 to 16, wherein the lipid nanoparticle has a mean diameter of 50-150 nm, or 80-100 nm.

18. The lipid nanoparticle of any one of claims 1 to 17, further comprising 4-100 additional nucleic acids, each having a different nucleic acid sequence and wherein each additional nucleic acid is linked to at least one other nucleic acid a non-covalent bond.

19. The lipid nanoparticle of claim 18, wherein each of the nucleic acids is an mRNA molecule encoding a different protein.

20. The lipid nanoparticle of claim 1, wherein the first and second nucleic acids are non-covalently linked to one another through a splint.

21. The lipid nanoparticle of claim 20, wherein the splint is an oligonucleotide having a region of complementarity with the first nucleic acid and a region of complementarity with the second nucleic acid.

22. A self-assembling multimeric mRNA structure comprising a first mRNA having a first linking region comprised of a part A and a part B and a second mRNA having a second linking region comprised of a part C and a part D, wherein at least part A of the first and at least part C of the second linking regions are complementary to one another.

23. The mRNA structure of claim 22, wherein the first linking region is in a non-coding region of the mRNA.

24. The mRNA structure of claim 22, wherein the second linking region is in a non-coding region of the mRNA.

25. The mRNA structure of claim 23 or 24, wherein the non-coding region is a 5' untranslated region (UTR).

26. The mRNA structure of claim 22, wherein the first and second linking regions are 5-100 nucleotides in length.

27. The mRNA structure of claim 22, wherein the first and second linking regions are 10-25 nucleotides in length.

28. The mRNA structure of any one of claims 22-27, further comprising a third mRNA, having a third linking region comprised of a part E and a part F, wherein part B of the first and part D of the second linking regions are complementary to parts E and F of the third linking regions.

29. The mRNA structure of any one of claims 1-27, further comprising 3-100 additional mRNAs, each mRNA having a linking region, wherein each linking region is complementary at least in part to at least one other linking region.

30. A self-assembling multimeric mRNA structure comprising 2-100 mRNAs each mRNA having a linking region and a stabilizing nucleic acid, wherein the stabilizing nucleic acid has a nucleotide sequence with regions complementary to each linking region.

31. The mRNA structure of claim 30, wherein the stabilizing nucleic acid is an RNA.

32. The mRNA structure of claim 30, wherein the stabilizing nucleic acid is a DNA.

33. The mRNA structure of claim 30, wherein the stabilizing nucleic acid has the following structure:

$L_1X_1L_2X_2\ L_3X_3L_4X_4L_5X_5L_6X_6$

wherein L is a nucleic acid sequence complementary to a linking region and wherein X is any nucleic acid sequence 0-50 nucleotides in length.

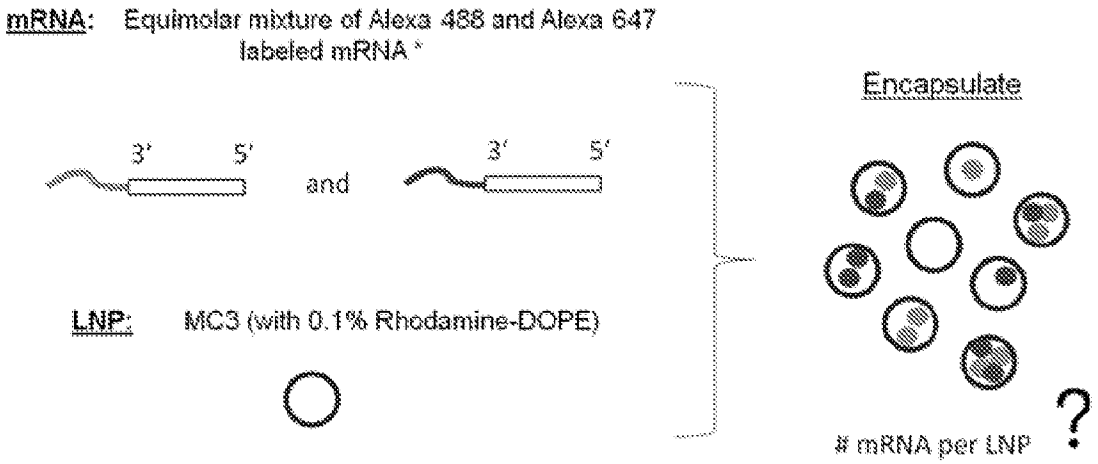
34. A multimeric mRNA structure comprising a first mRNA and a second mRNA, wherein the first mRNA and the second mRNA are non-covalently linked to one another through a splint.

35. The multimeric mRNA structure of claim 34, wherein the splint is an oligonucleotide having a first region of complementarity with the first nucleic acid and a second region of complementarity with the second nucleic acid.

36. The multimeric mRNA structure of claim 34 or 35, wherein the first region of complementarity is located in the 5'UTR of the first nucleic acid and the second region of complementarity is located in the 5'UTR of the second nucleic acid.

37. The multimeric mRNA structure of any one of claims 34 to 36, wherein each region of complementarity is at least 6 base pairs long.

FIG. 1


mRNA: Equimolar mixture of Alexa 488 and Alexa 647 labeled mRNA *
3' 5'
and
3' 5'
LNP: MC3 (with 0.1% Rhodamine-DOPE)
Encapsulate
mRNA per LNP
?

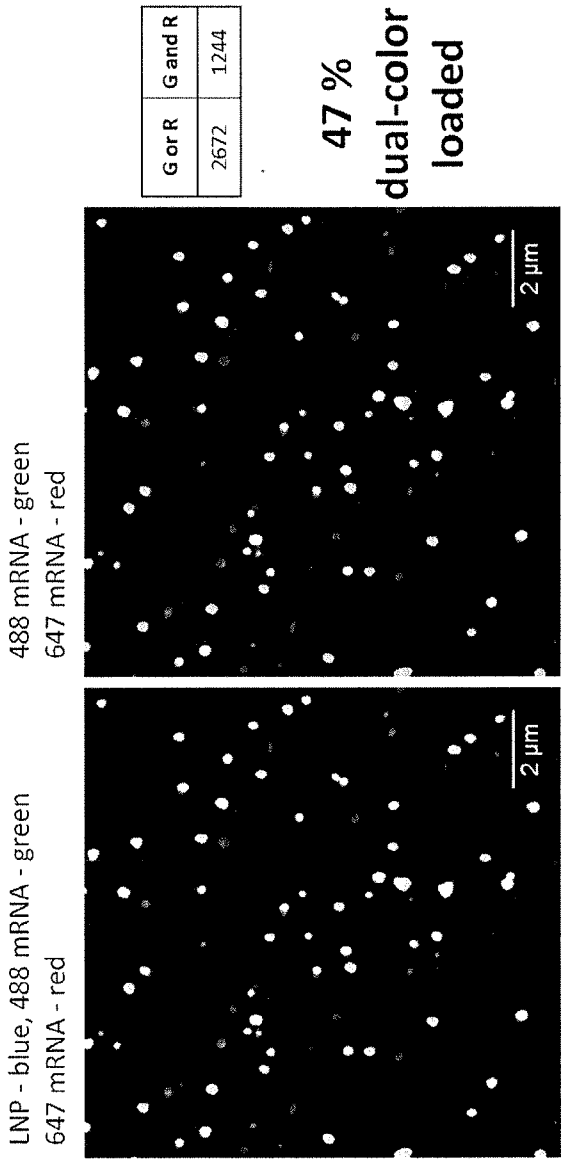

LNP - blue, 488 mRNA - green
647 mRNA - red
2 µm
488 mRNA - green
647 mRNA - red
2 µm
47 % dual-color loaded
G or R
G and R
2672
1244

FIG. 2

FIG. 3A

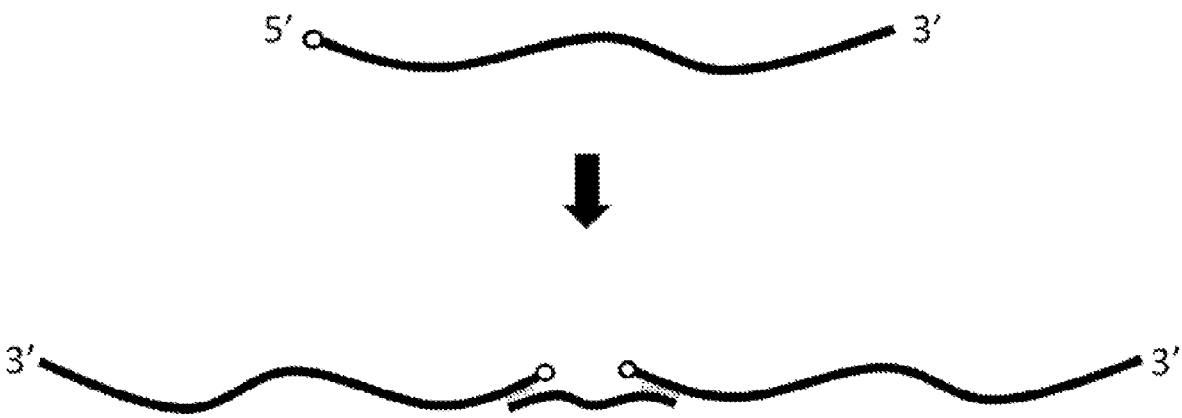


FIG. 3B

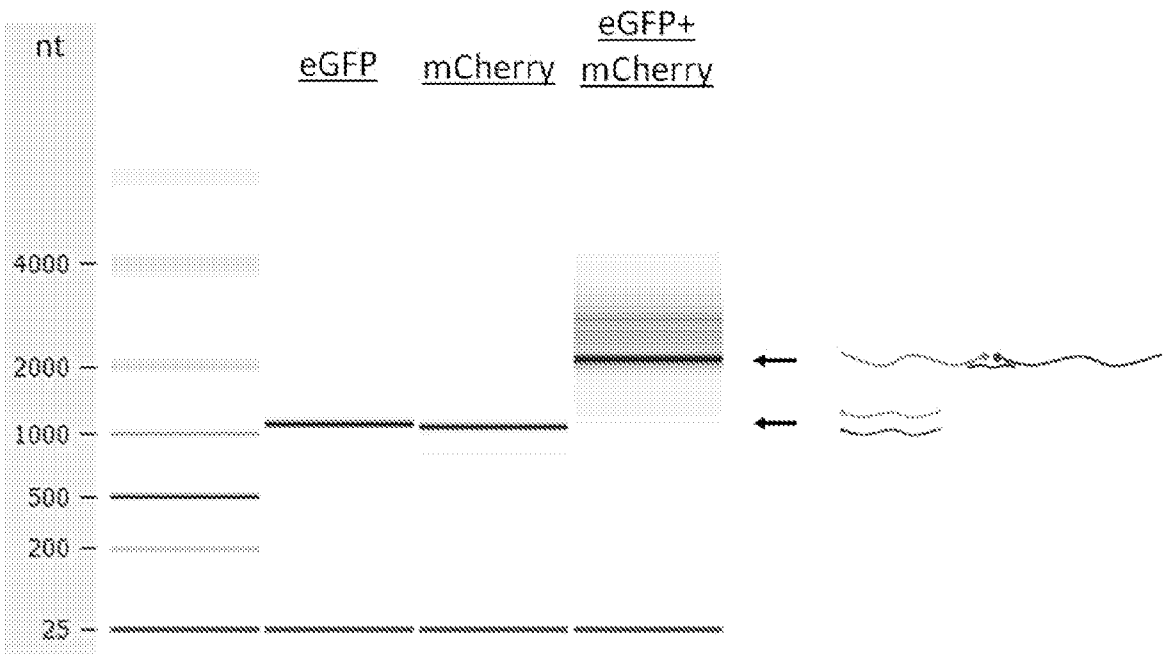


FIG. 4A

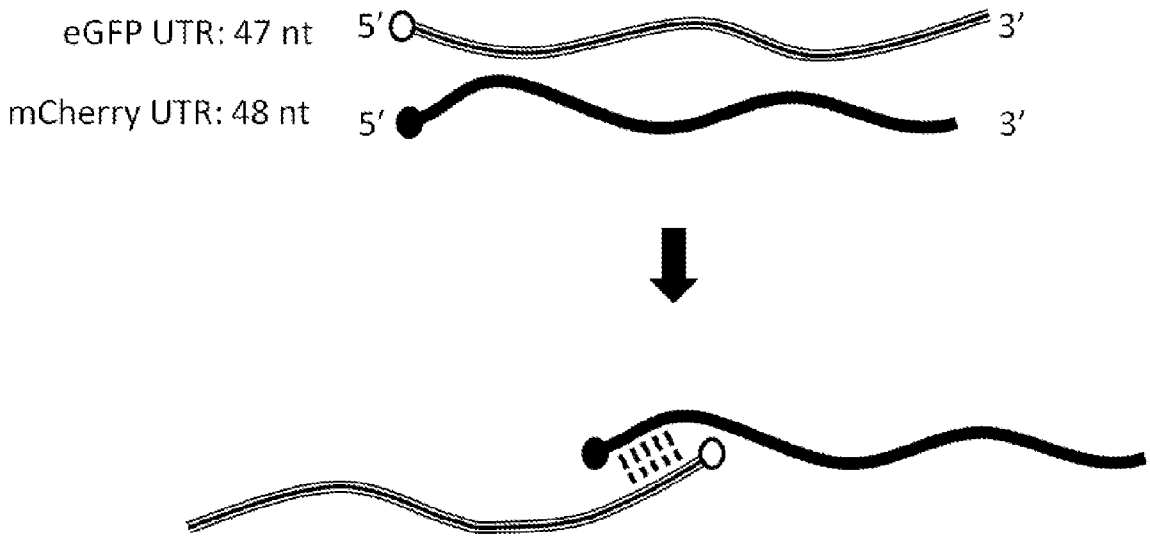
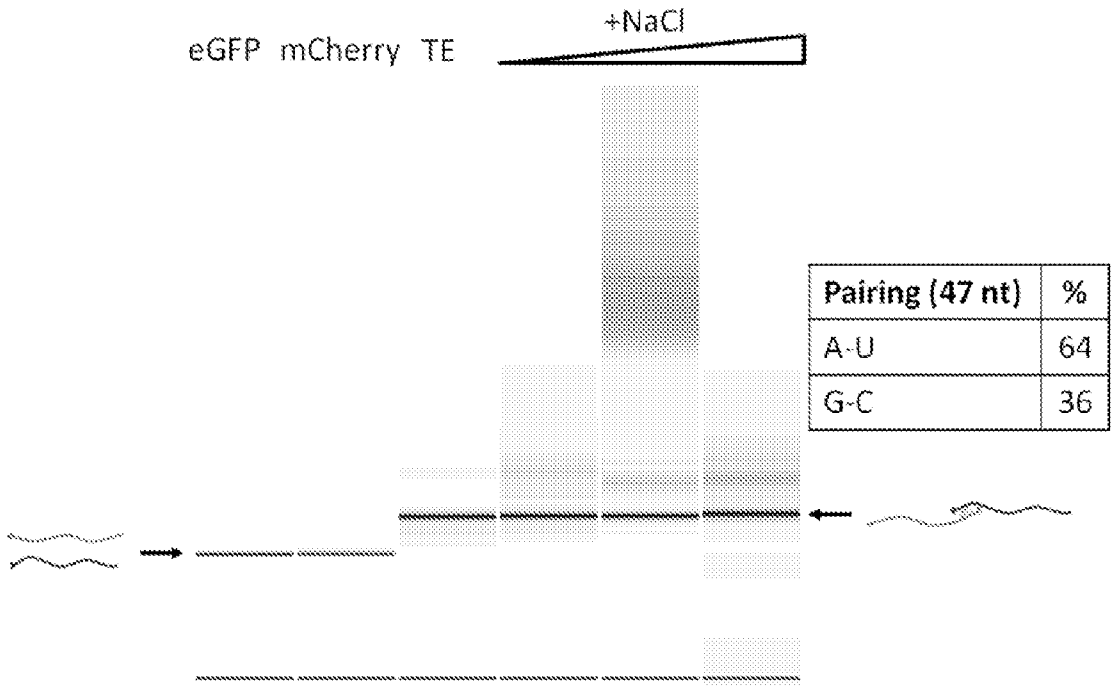


FIG. 4B



Pairing (47 nt)	%
A-U	64
G-C	36

FIG. 5

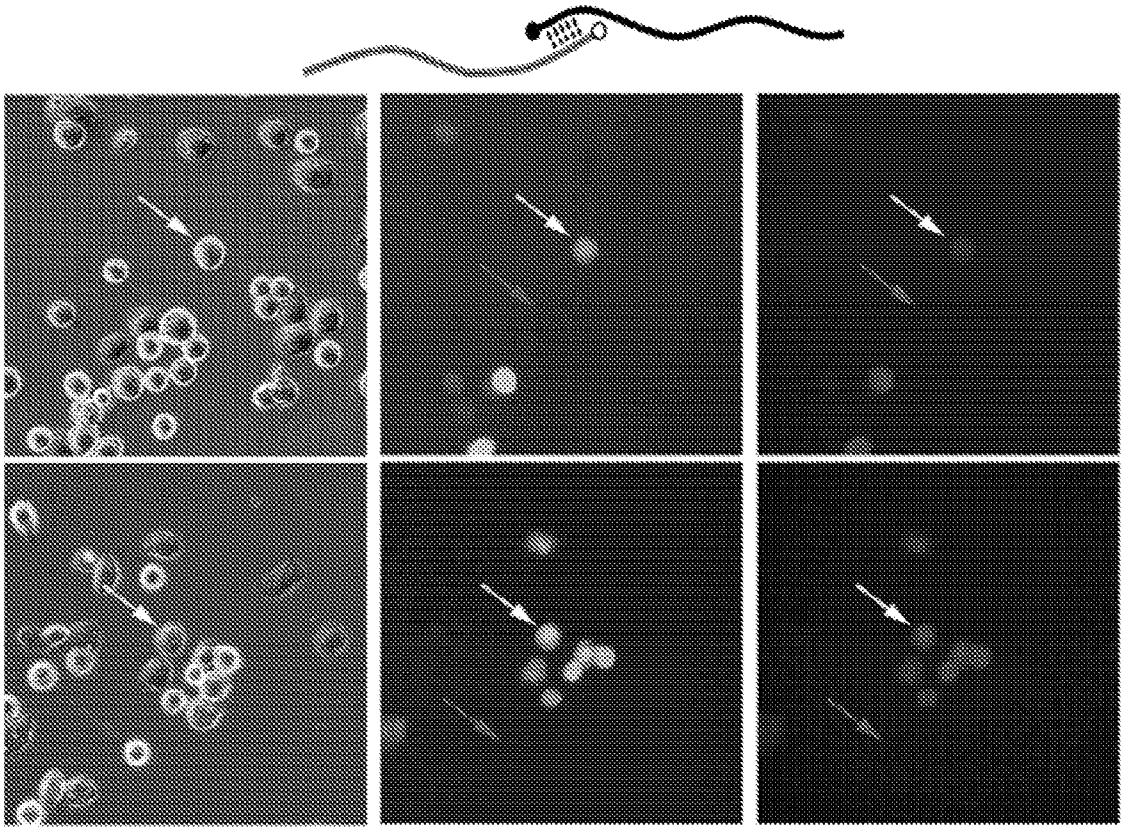
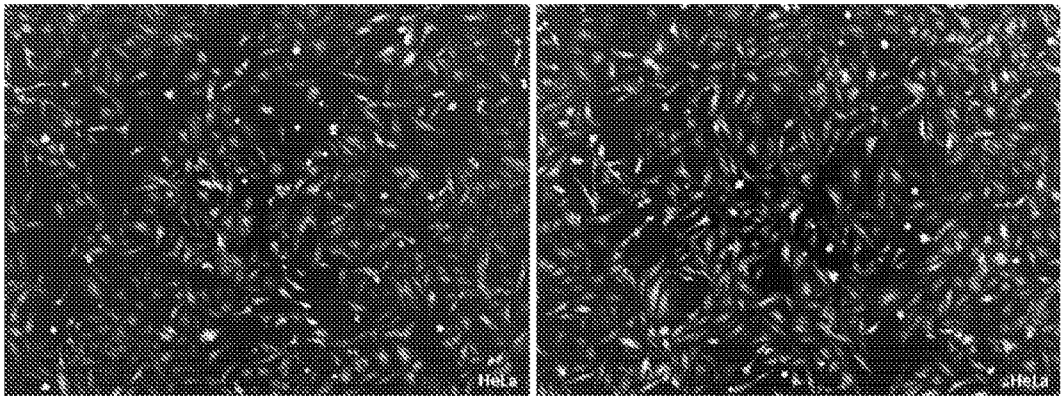


FIG. 6



No Multiplex **Multiplex**

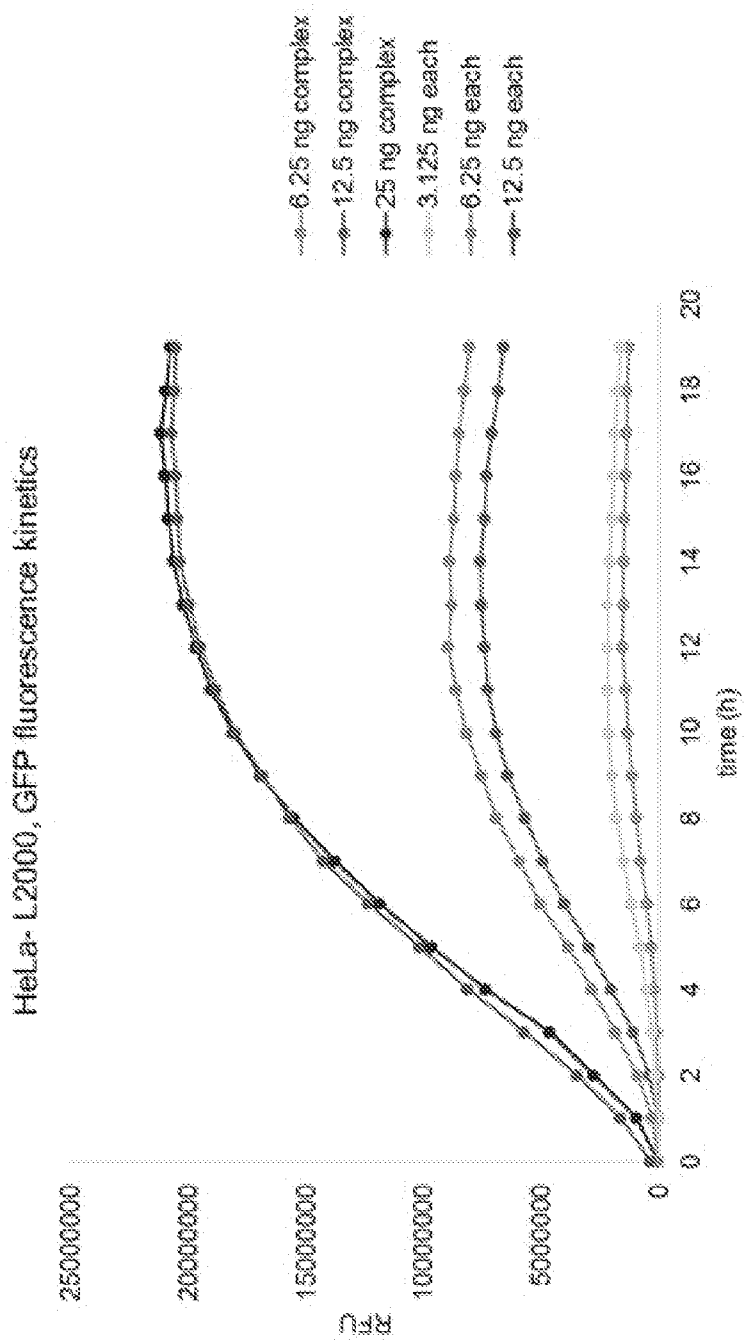

HeLa- L2000, GFP fluorescence kinetics
RFU
25000000
20000000
15000000
10000000
5000000
0
0
2
4
6
8
10
12
14
16
18
20
time (h)
6.25 ng complex
12.5 ng complex
25 ng complex
3.125 ng each
6.25 ng each
12.5 ng each

FIG. 7

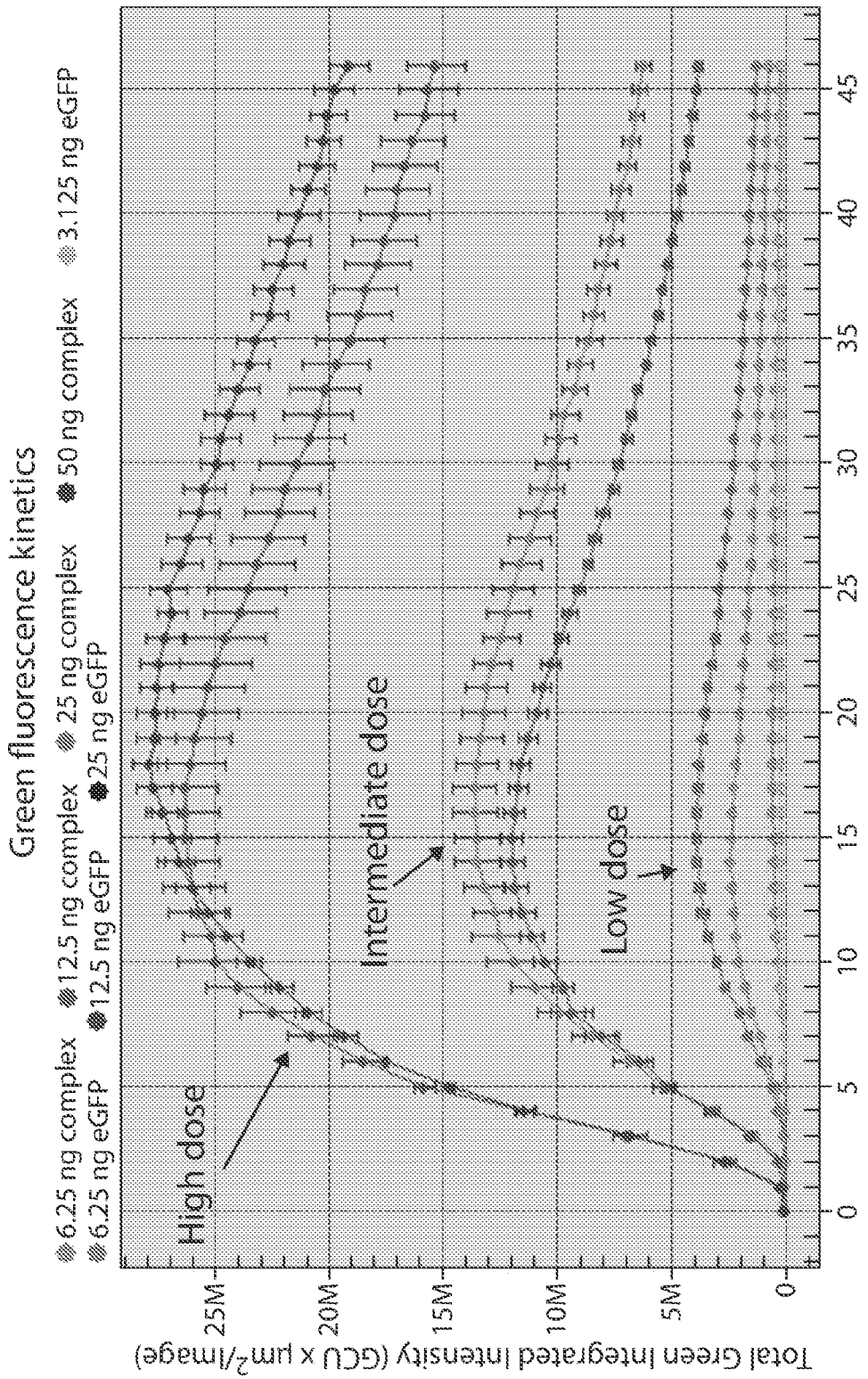

Green fluorescence kinetics
6.25 ng complex
12.5 ng complex
25 ng complex
50 ng complex
3.125 ng eGFP
6.25 ng eGFP
12.5 ng eGFP
25 ng eGFP
High dose
Intermediate dose
Low dose
Total Green Integrated Intensity (GCU x µm²/Image)
0
5M
10M
15M
20M
25M
Time (Hours)
0
5
10
15
20
25
30
35
40
45

FIG. 8

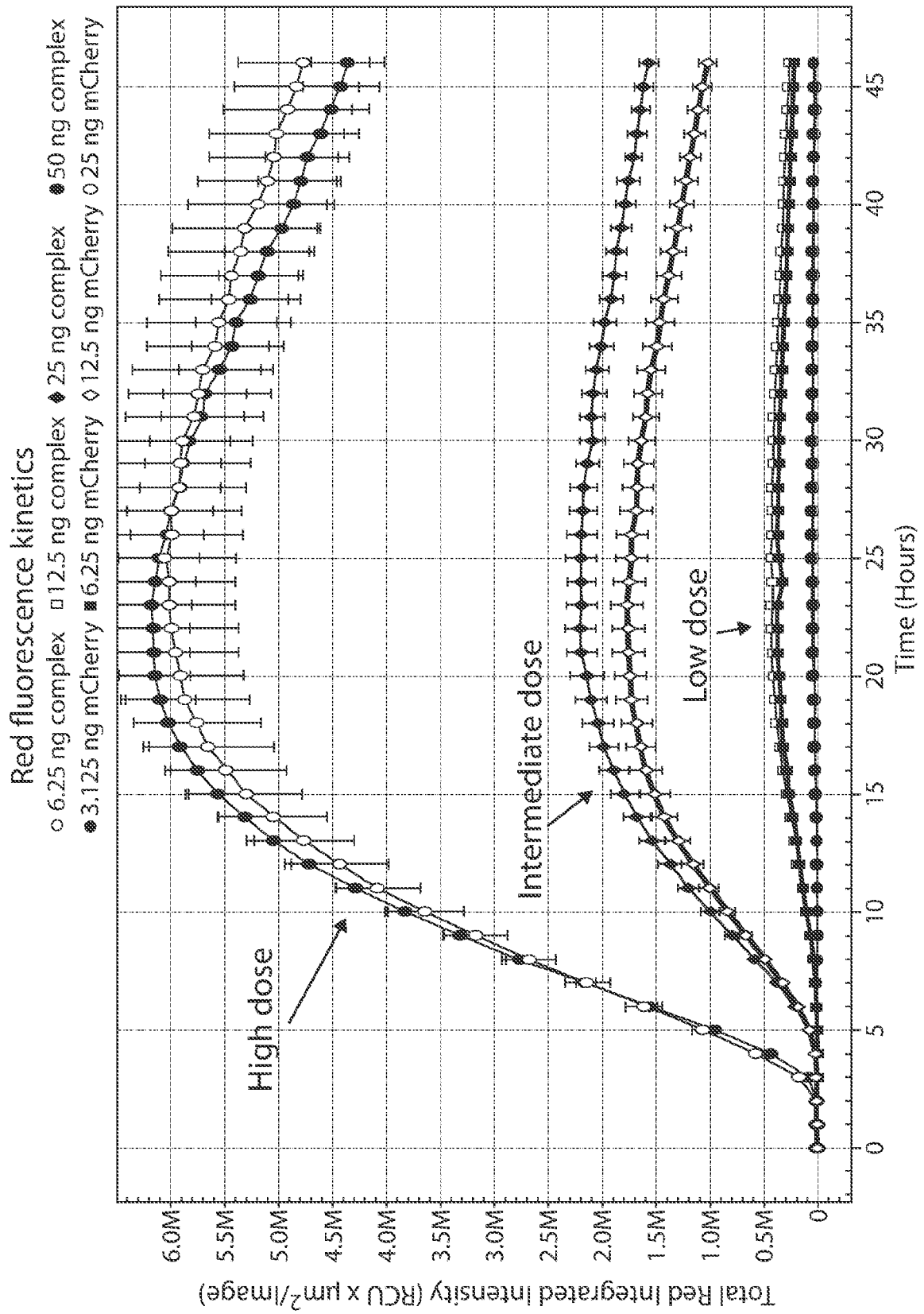
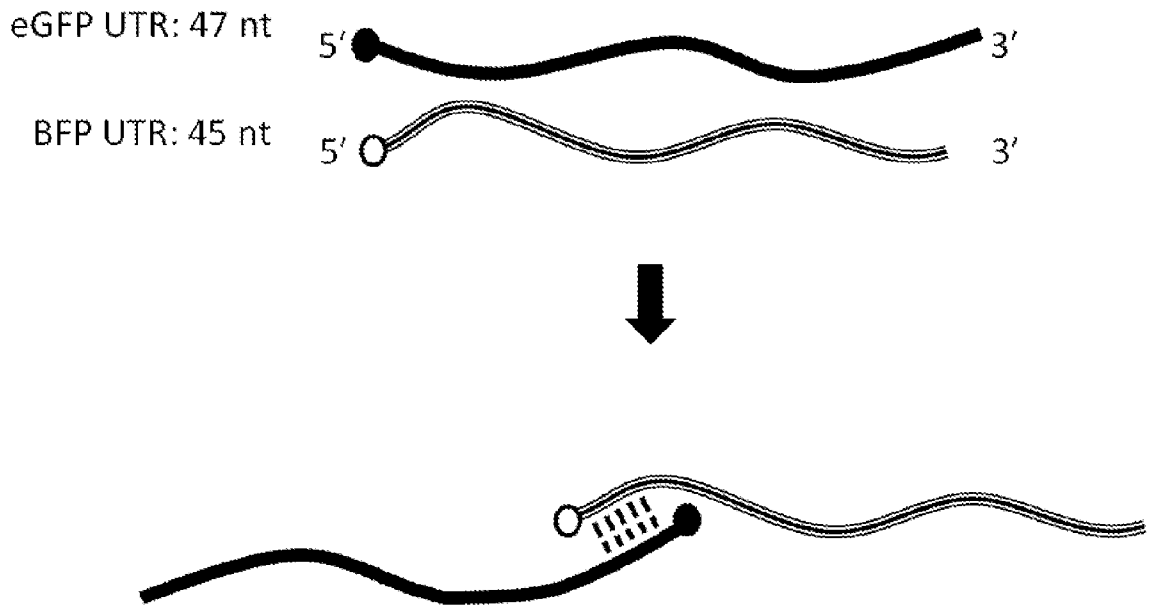

Red fluorescence kinetics
○ 6.25 ng complex □ 12.5 ng complex ◆ 25 ng complex ● 50 ng complex
● 3.125 ng mCherry ■ 6.25 ng mCherry ◇ 12.5 ng mCherry ○ 25 ng mCherry
High dose
Intermediate dose
Low dose
Total Red Integrated Intensity (RCU x µm²/Image)
Time (Hours)
0
0.5M
1.0M
1.5M
2.0M
2.5M
3.0M
3.5M
4.0M
4.5M
5.0M
5.5M
6.0M
0
5
10
15
20
25
30
35
40
45

FIG. 9

FIG. 10A



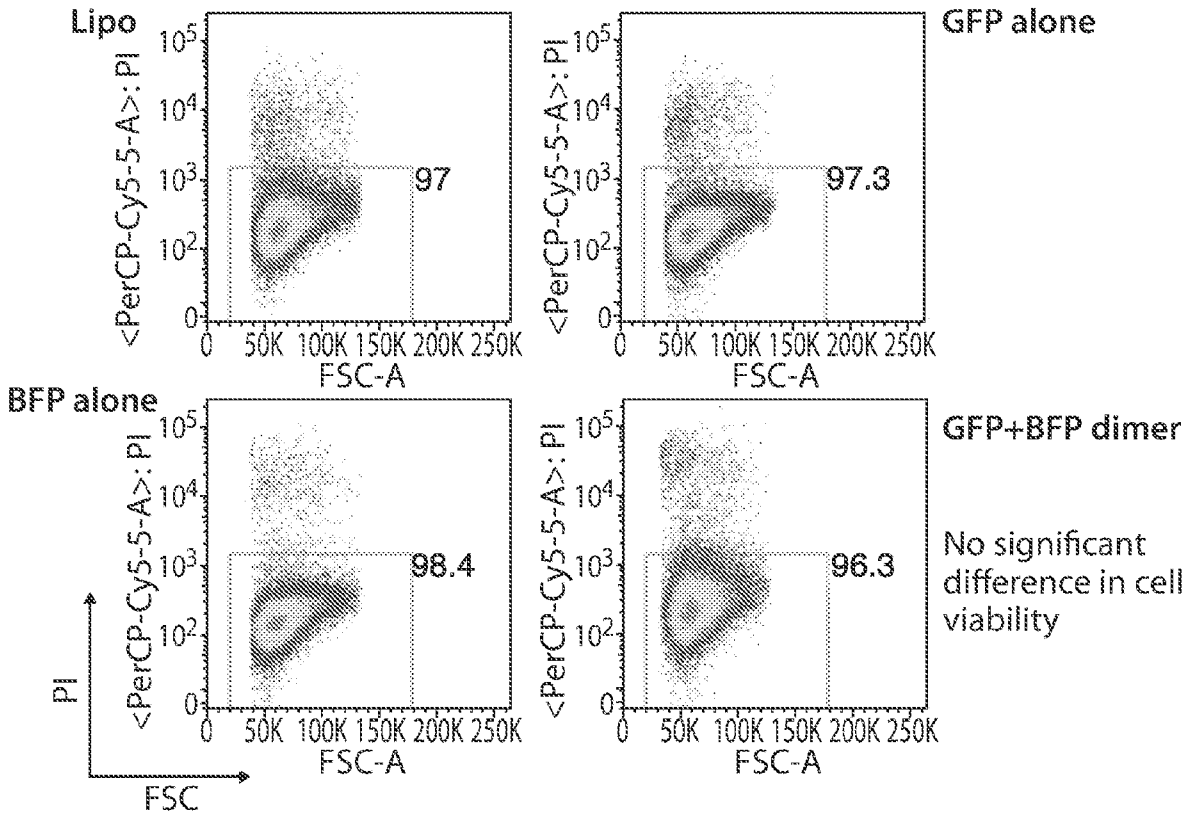


FIG. 10B

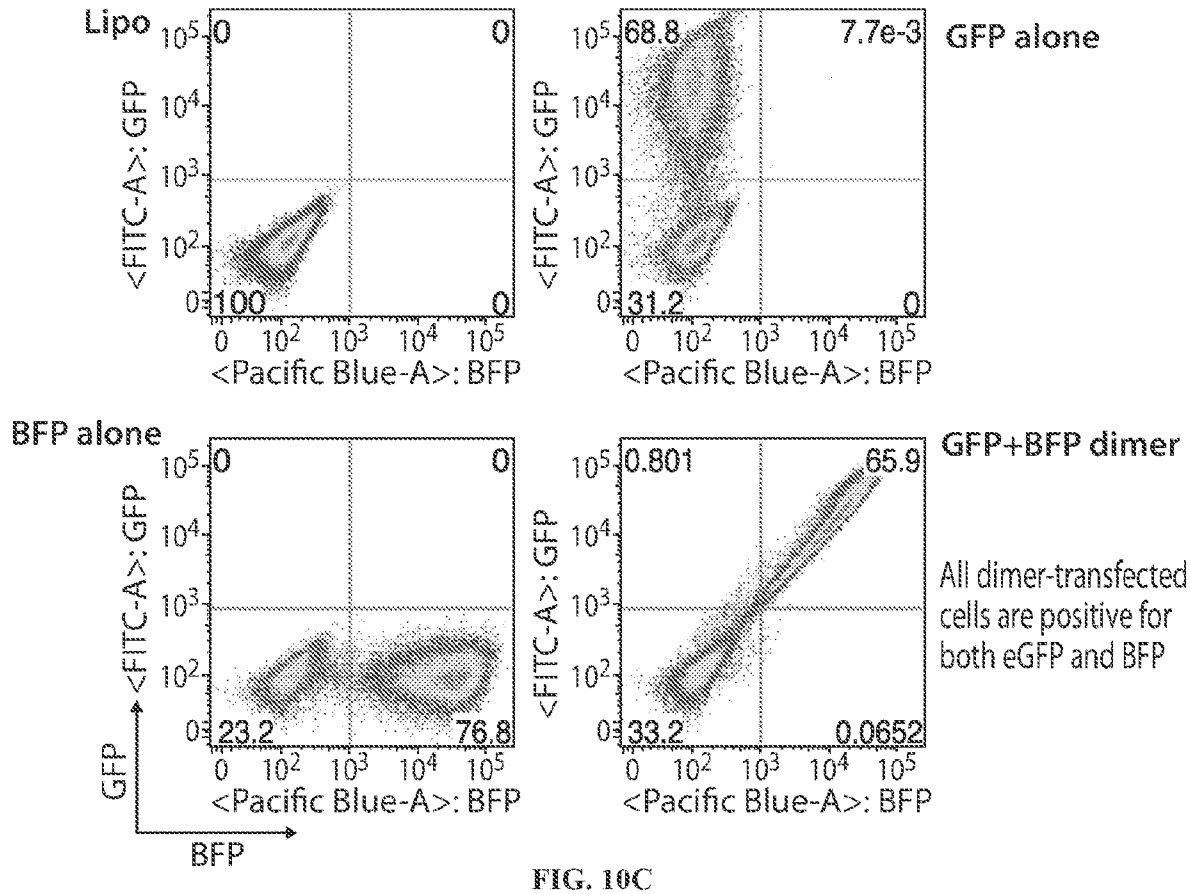

Lipo
0
0
100
0
GFP alone
68.8
7.7e-3
31.2
0
BFP alone
0
0
23.2
76.8
GFP+BFP dimer
0.801
65.9
33.2
0.0652
All dimer-transfected cells are positive for both eGFP and BFP
<FITC-A>: GFP
<Pacific Blue-A>: BFP
GFP
BFP

FIG. 10C

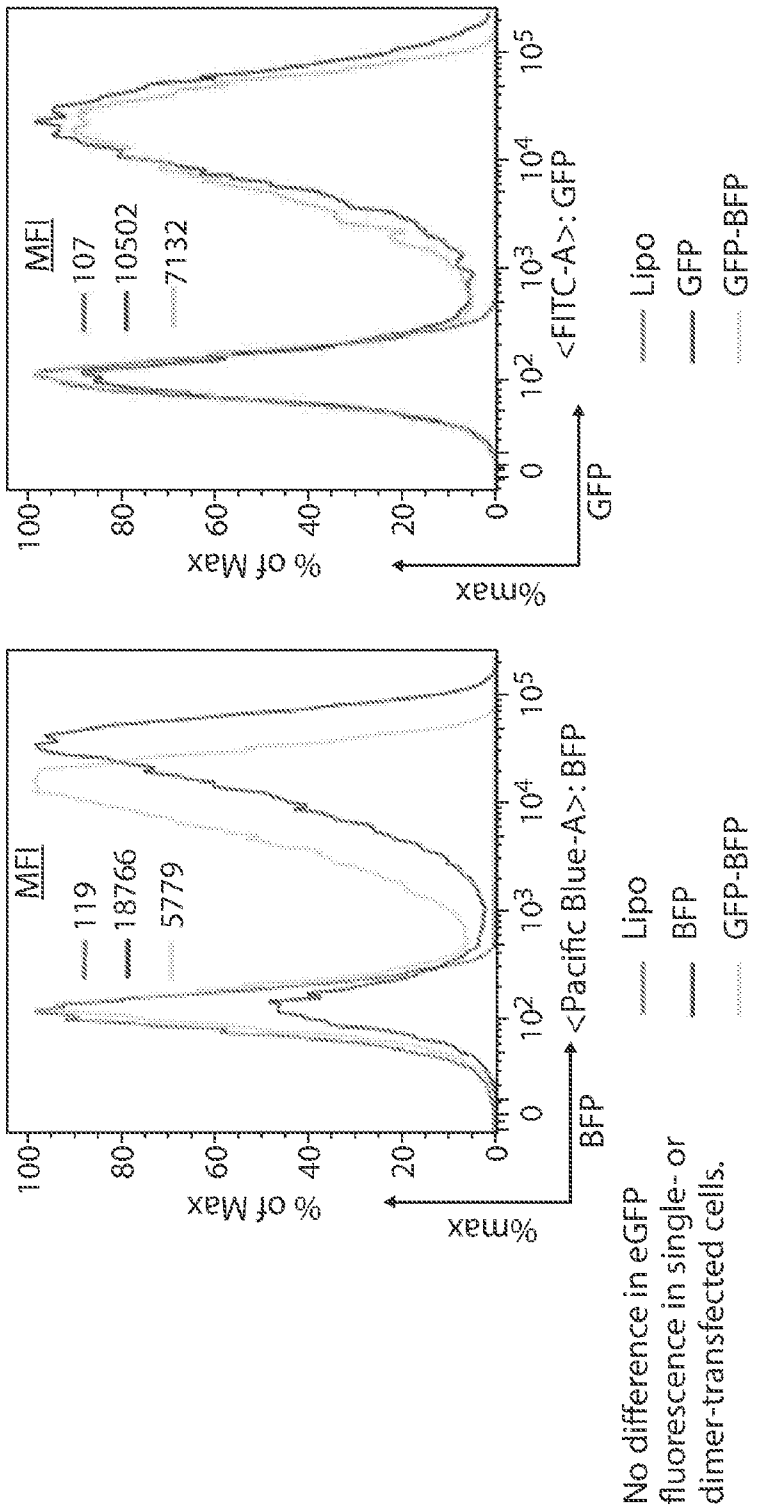

MFI
119
18766
5779
% of Max
%max
BFP
<Pacific Blue-A>: BFP
Lipo
BFP
GFP-BFP
No difference in eGFP fluorescence in single- or dimer-transfected cells.
MFI
107
10502
7132
% of Max
%max
GFP
<FITC-A>: GFP
Lipo
GFP
GFP-BFP
Difference in BFP could be due to higher amount transfected as a singlet.

FIG. 11

FIG 12A

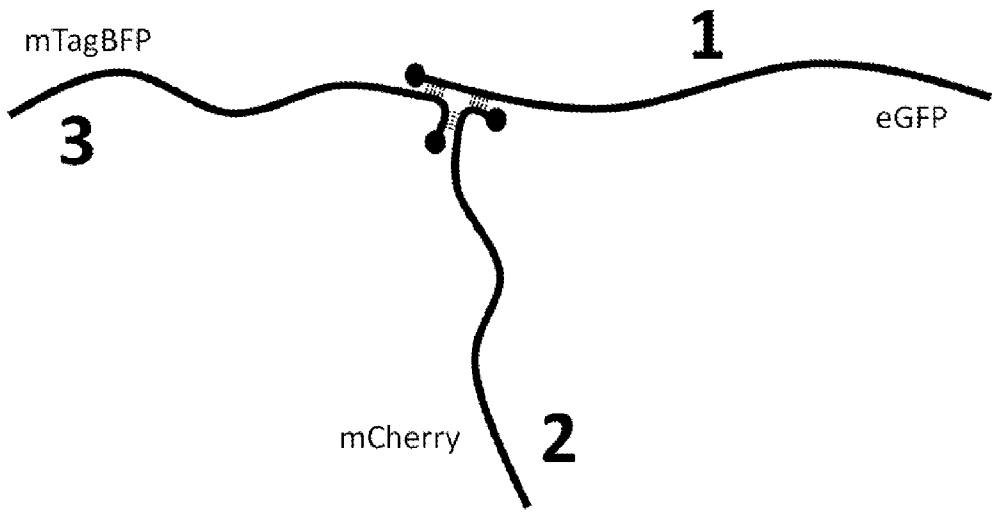
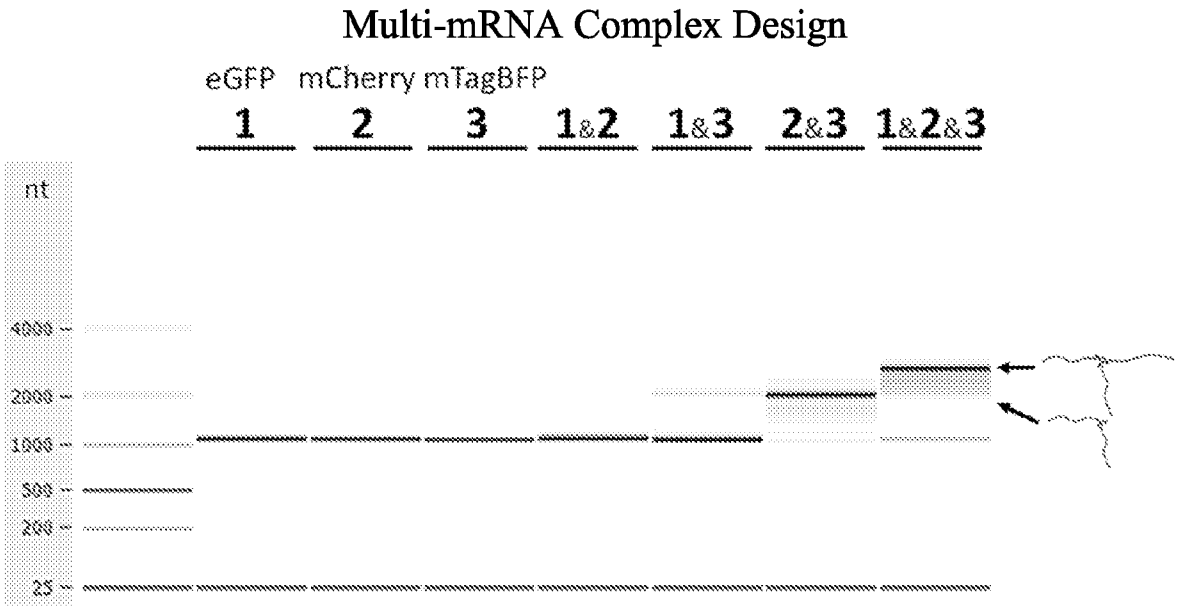


FIG. 12B



Pairing	1&2 (22nt)	1&3 (20nt)	2&3 (24nt)
A-U	64%	70%	50%
G-C	36%	30%	50%

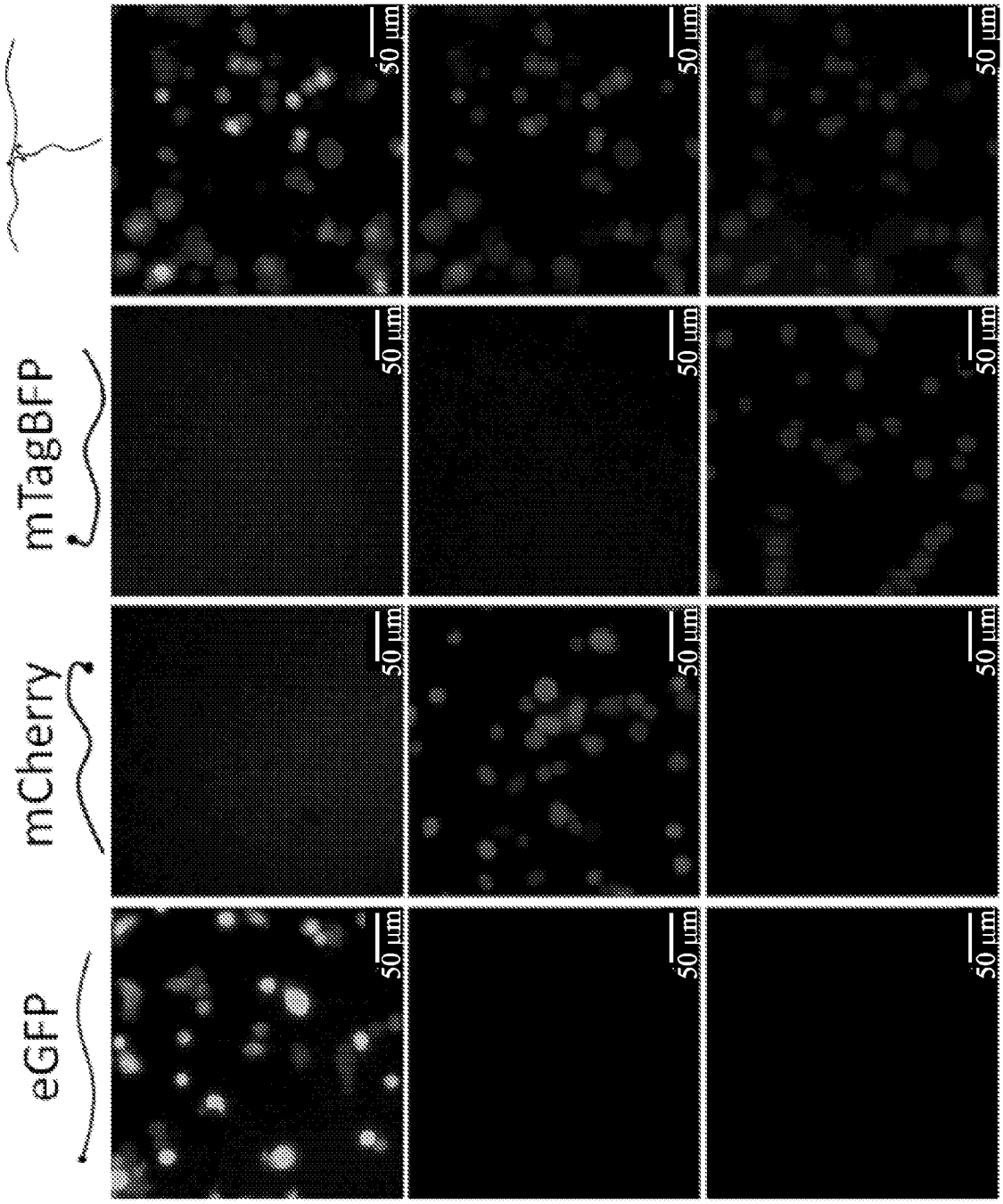


FIG. 13

FIG. 14A

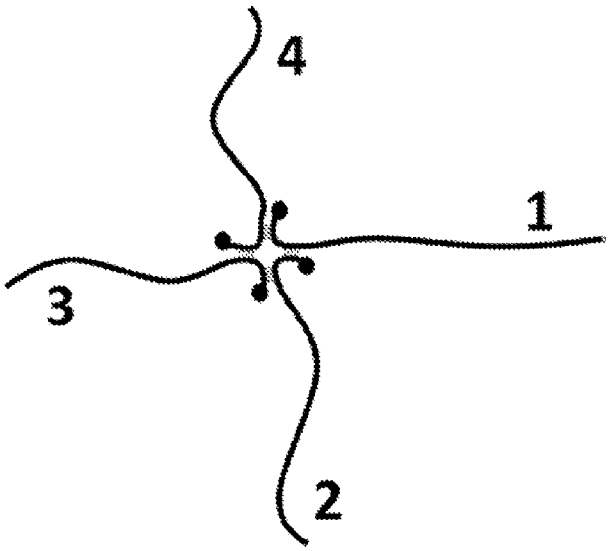


FIG. 14B

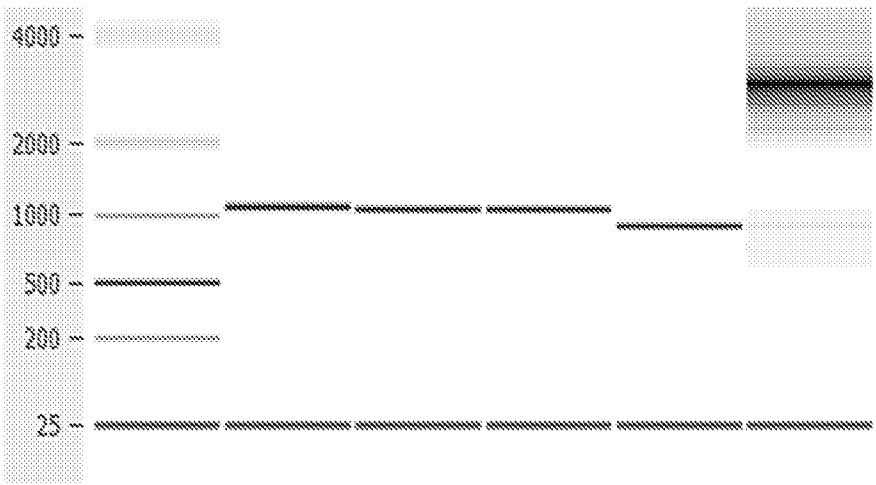


FIG. 15

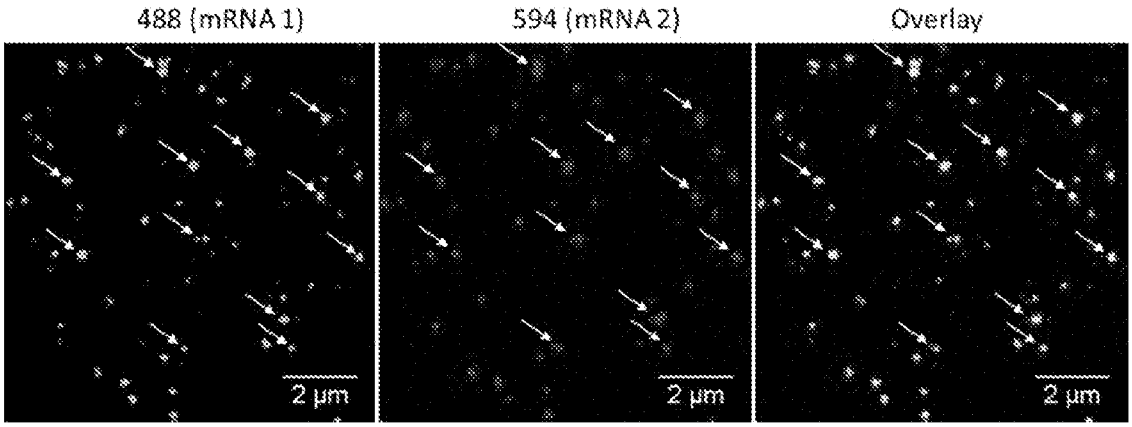


FIG. 16

<u>Dimer</u>

mRNA #1 (SEQ ID NO: 1)
GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCCACCNNNN
NN<u>TGATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCC</u>
<u>CCAGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTG</u>
<u>AGTGGGCGGCTCTAGA</u>aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

mRNA #2 (SEQ ID NO: 2)
GGGTGGCTCTTATATTTCTTCTTACTCTTCTTTTCTCTCTTATTTCCNNNNNN
<u>TGATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCCCC</u>
<u>AGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGA</u>
<u>GTGGGCGGCTCTAGA</u>aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

FIG. 17

Trimer

mRNA #1 (SEQ ID NO: 1)
GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCCACCNNNN
NNTGATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCC
CCAGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTG
AGTGGGCGGCTCTAGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

mRNA #2 (SEQ ID NO: 2)
GGGTGGCTCTTATATTTCTTCTTAGTCCGAATTCGAAGTACGGCTACNNNNN
NTGATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCCC
CAGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGA
GTGGGCGGCTCTAGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

mRNA #3 (SEQ ID NO: 3)
GGTAGCCGTACTTCGAATTCGGACTTTCTTTTCTCTCTTATTTCCNNNNNNTG
ATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCCCCAG
CCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTG
GGCGGCTCTAGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

FIG. 18

Tetramer

mRNA #1 (SEQ ID NO: 1)
GGGACAGACGAGAGAGAAGCACGCCAATTCTGCCTGCTTAAGCCNNNNNNT
GATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCCCCA
GCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGT
GGGCGGCTCTAGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

mRNA #2 (SEQ ID NO: 2)
GGGCTTAAGCAGGCAGAATTGGCCCAGAAGACAGATAAGAGCCTCCNNNNN
NTGATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCCC
CAGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGA
GTGGGCGGCTCTAGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

mRNA #3 (SEQ ID NO: 3)
GGGAGGCTCTTATCTGTCTTCTCAGTCCGAATTCGAAGTACGGCTACCNNNN
NNTGATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCC
CCAGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTG
AGTGGGCGGCTCTAGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

mRNA #4 (SEQ ID NO: 4)
GGGTAGCCGTACTTCGAATTCGGACAAGCTTCTCTCTCGTCTGTCCNNNNNN
TGATAATAGGCTGGAGCCTCGGTGGCCAAGCTTCTTGCCCCTTGGGCCTCCCCCC
AGCCCCTCCTCCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGA
GTGGGCGGCTCTAGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

FIG. 19

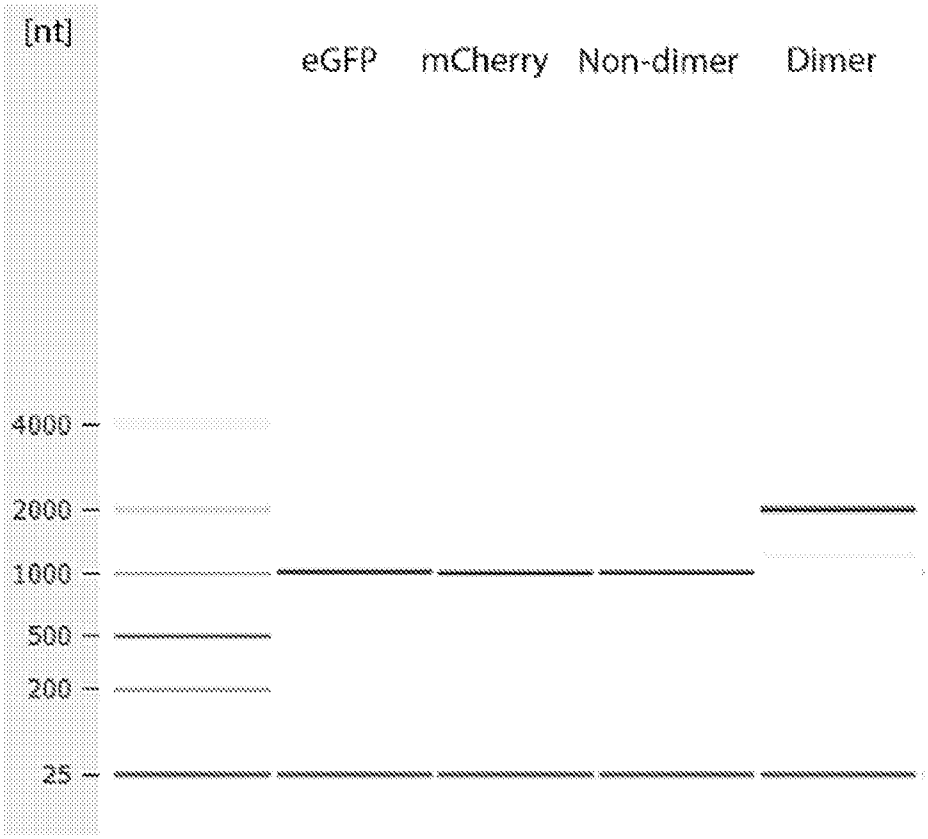


FIG. 20

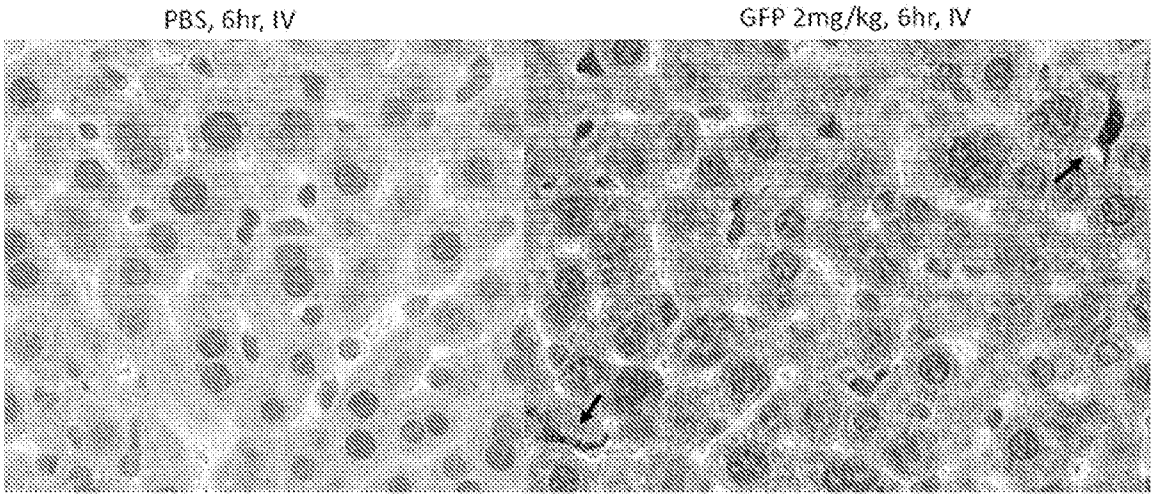
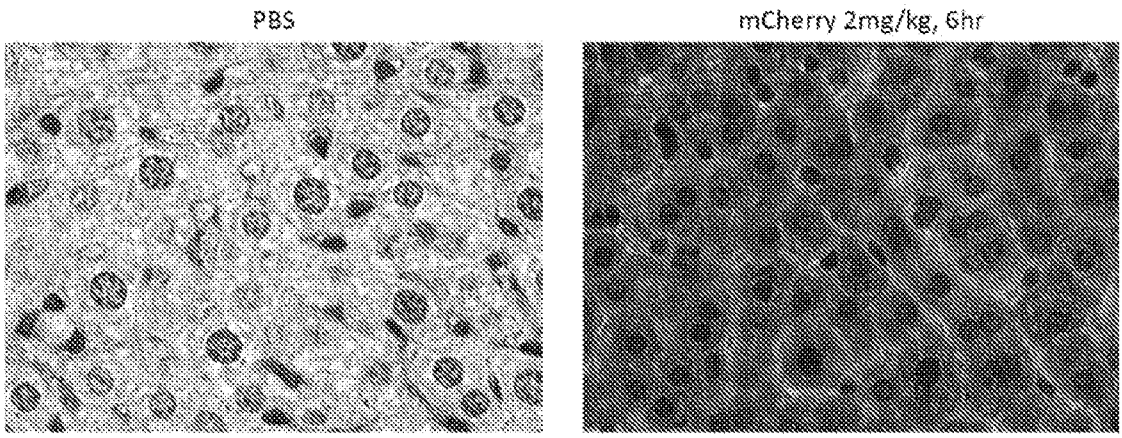


FIG. 21



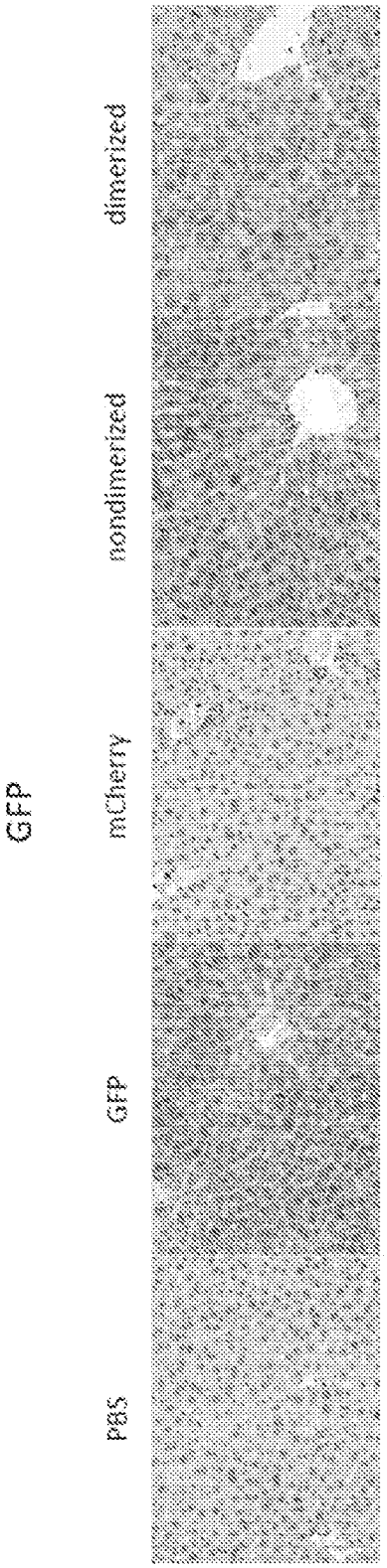


FIG. 22A

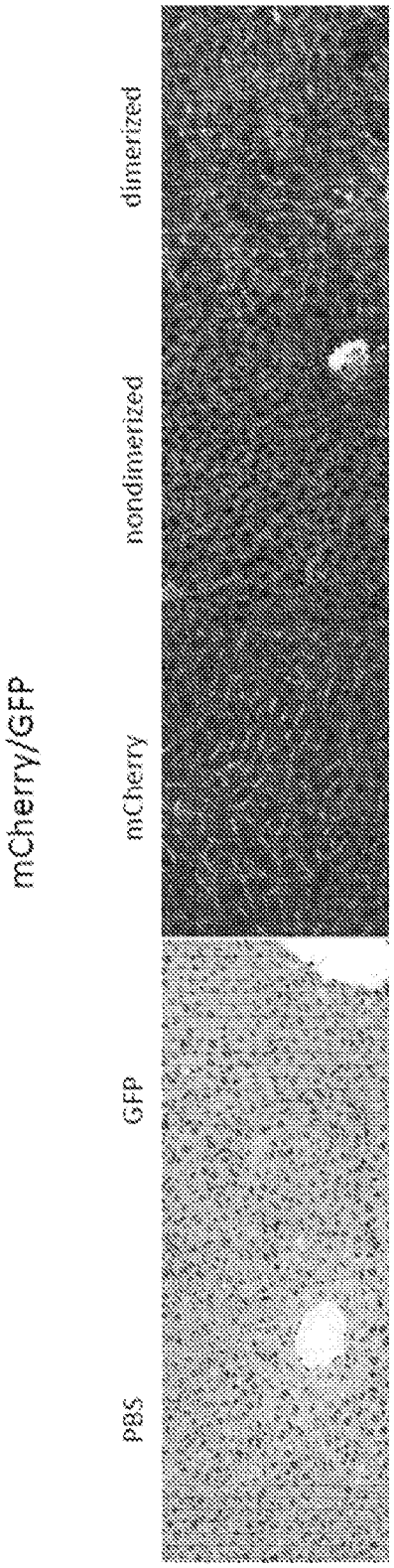


FIG. 22B

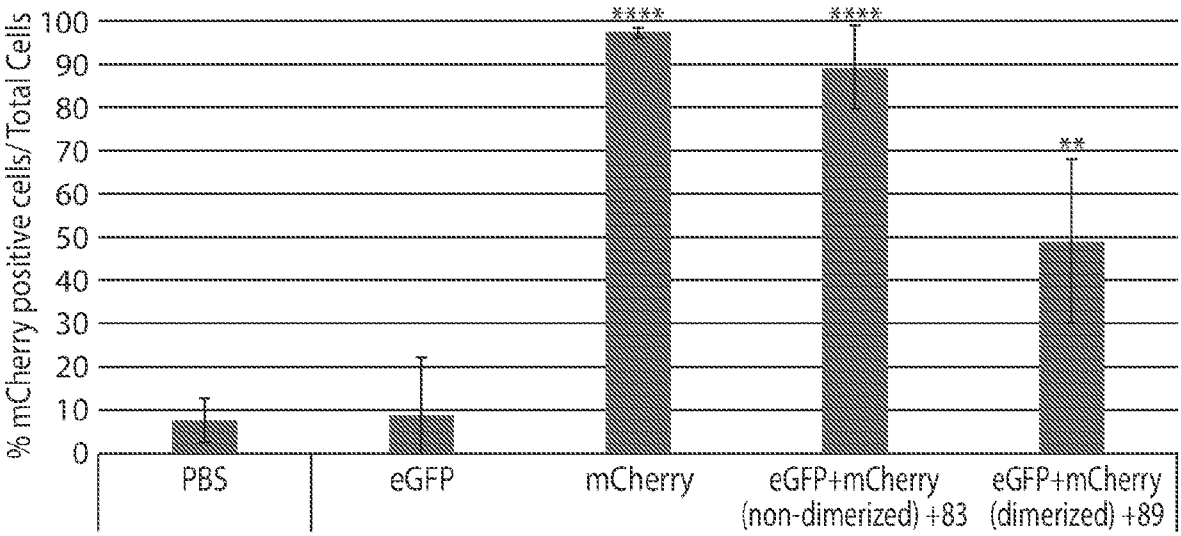

% mCherry positive cells/Total Cells
100
90
80
70
60
50
40
30
20
10
0
****
****
**
PBS
eGFP
mCherry
eGFP+mCherry (non-dimerized) +83
eGFP+mCherry (dimerized) +89

FIG. 23

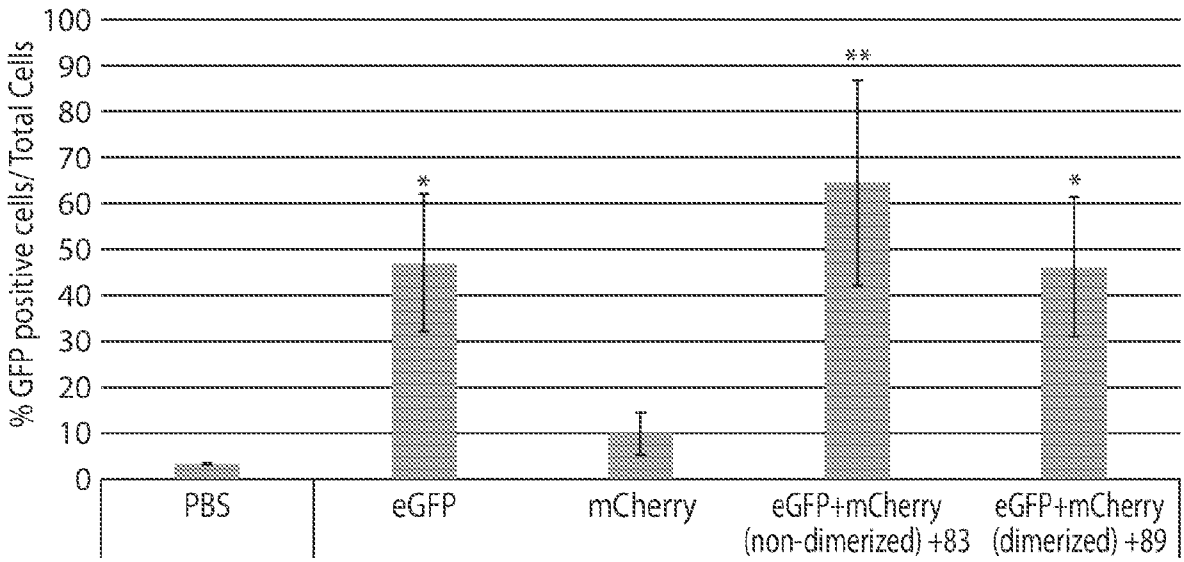

% GFP positive cells/ Total Cells
100
90
80
70
60
50
40
30
20
10
0
*
**
*
PBS
eGFP
mCherry
eGFP+mCherry (non-dimerized) +83
eGFP+mCherry (dimerized) +89

FIG. 24

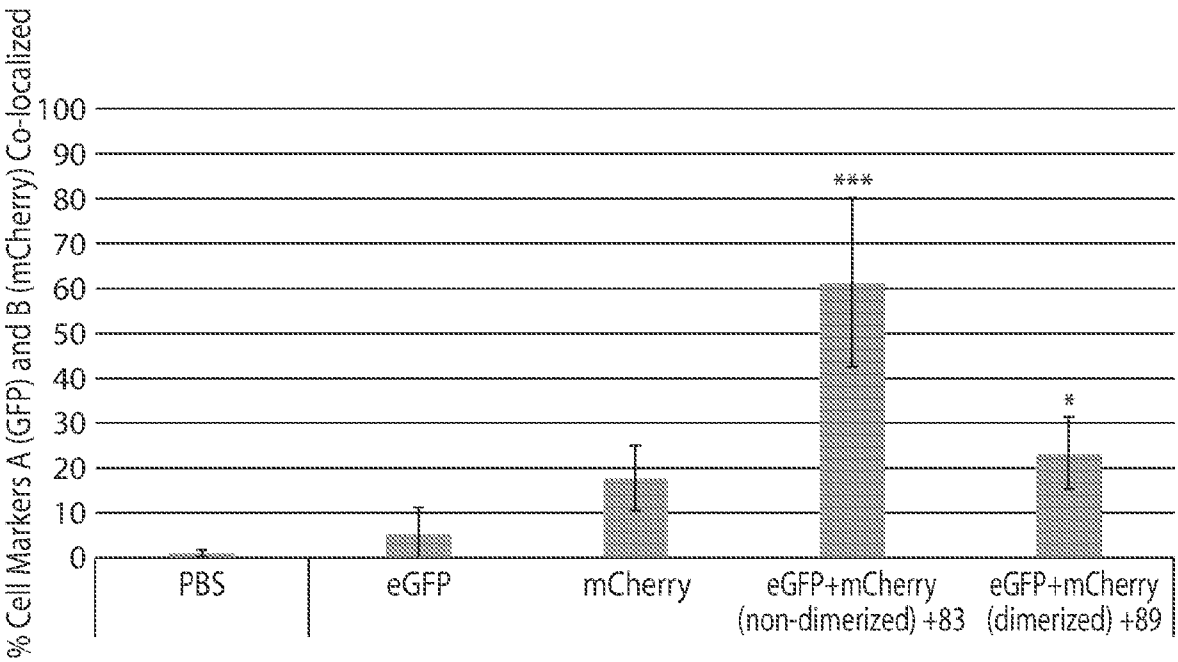

% Cell Markers A (GFP) and B (mCherry) Co-localized
100
90
80
70
60
50
40
30
20
10
0
***
*
PBS
eGFP
mCherry
eGFP+mCherry (non-dimerized) +83
eGFP+mCherry (dimerized) +89

FIG. 25

FIG. 26A

PBS M1 M2 M3

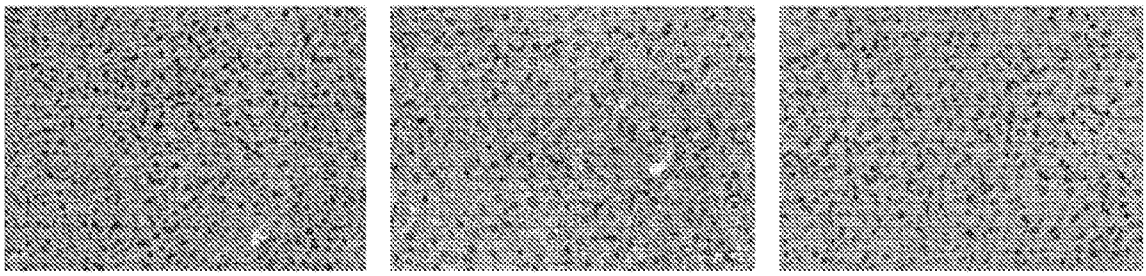


FIG. 26B

GFP mRNA 2mg/kg

M1 M2 M3

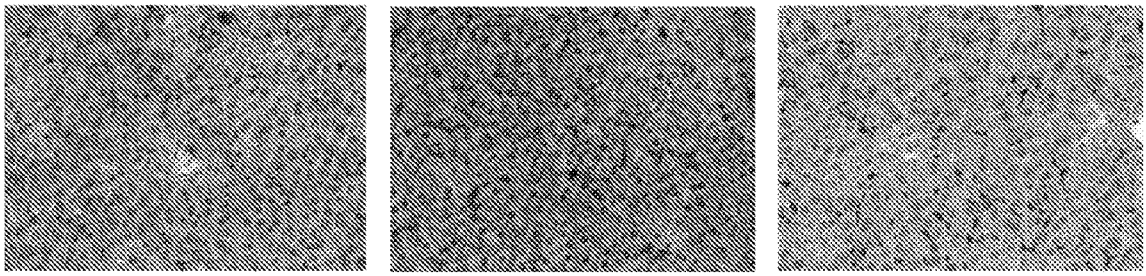


FIG. 26C

mCherry mRNA 2mg/kg

M1 M2 M3

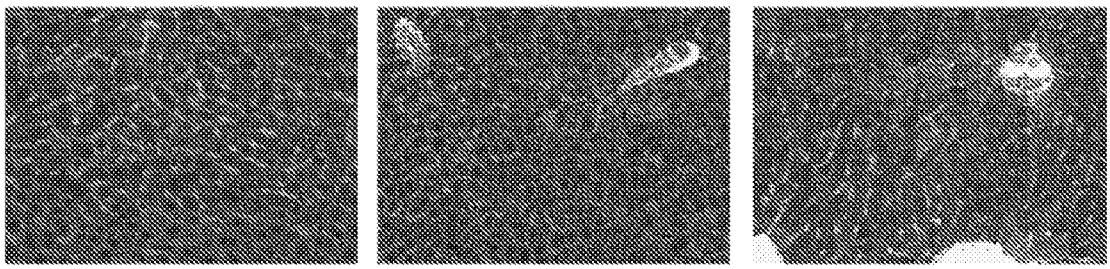


FIG. 26D

eGFP+mCherry (non-dimerized) +83 2mg/kg

M1 M2 M3

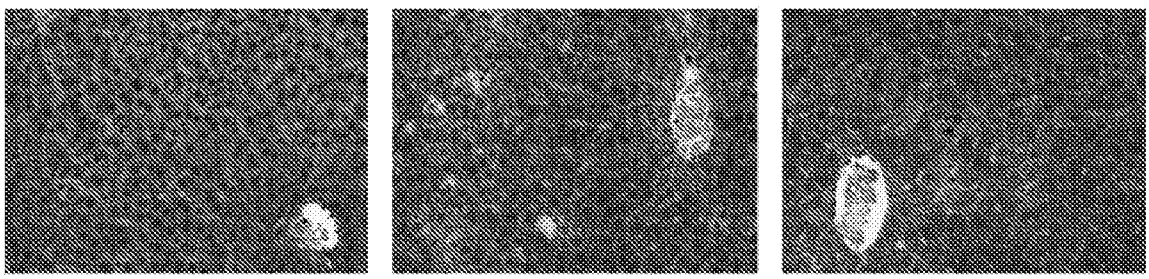


FIG. 26E

eGFP+mCherry (dimerized) +89
2mg/kg

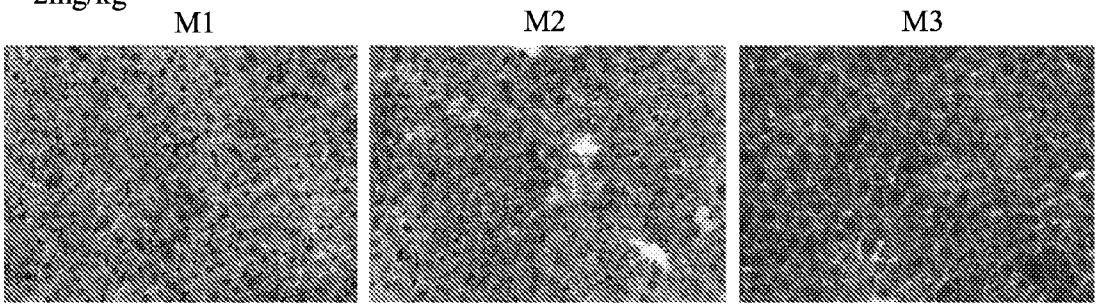


FIG. 27A

PBS

M1 M2 M3

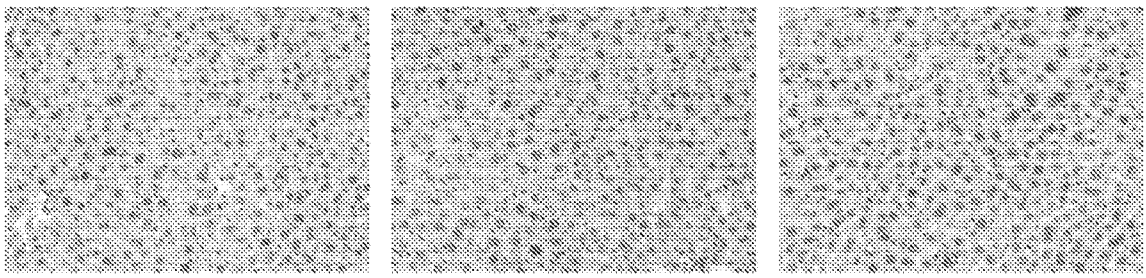


FIG. 27B

GFP mRNA 2mg/kg

M1 M2 M3

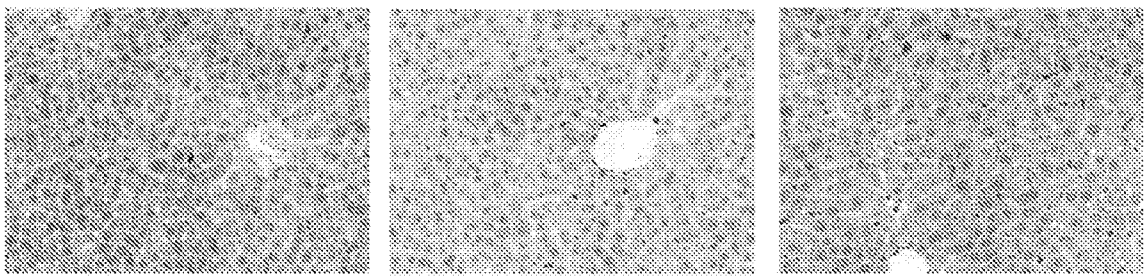
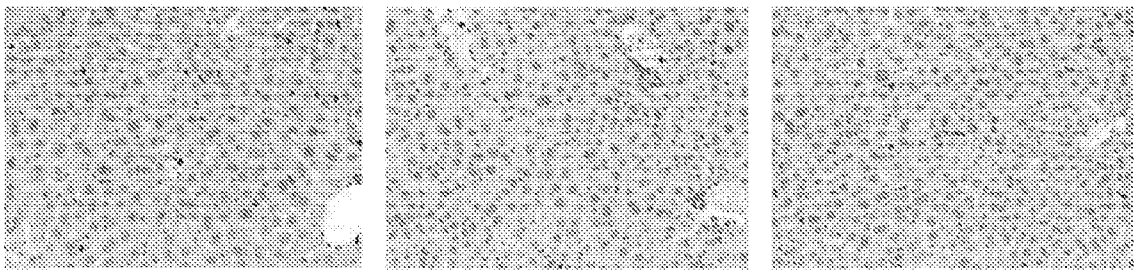


FIG. 27C

mCherry mRNA 2mg/kg

M1 M2 M3



Non-specific staining

FIG. 27D

eGFP+mCherry (non-dimerized) +83 2mg/kg

M1 M2 M3

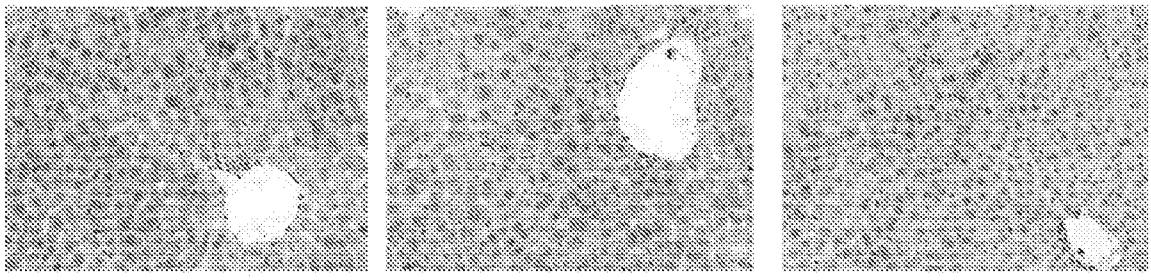


FIG. 27E

eGFP+mCherry (dimerized) +89
2mg/kg

M1 M2 M3

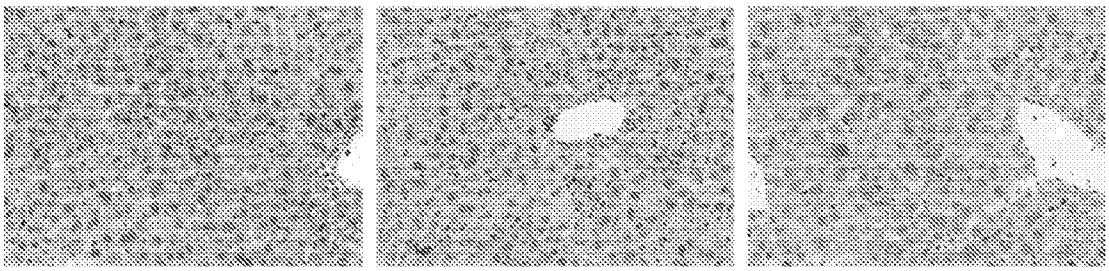


FIG. 28A

PBS

M1 M2 M3

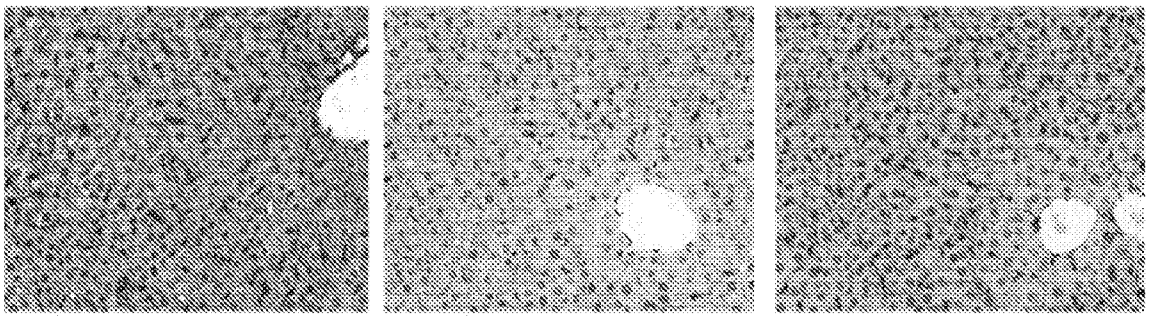


FIG. 28B

GFP mRNA 2mg/kg

M1 M2 M3

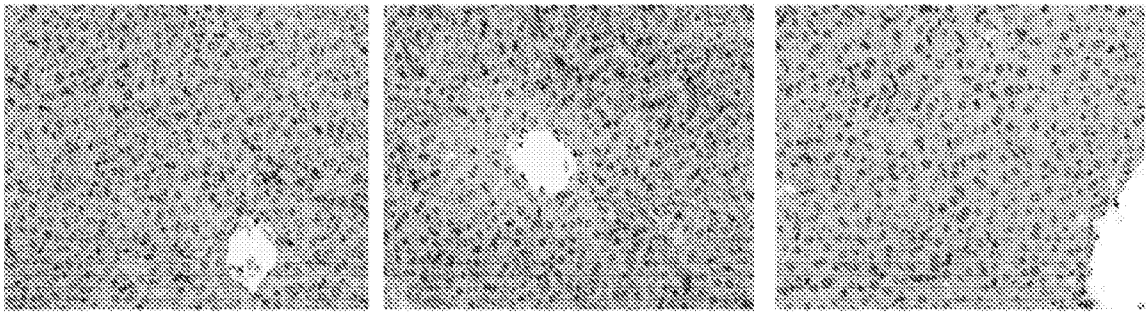


FIG. 28C

mCherry mRNA 2mg/kg

M1 M2 M3

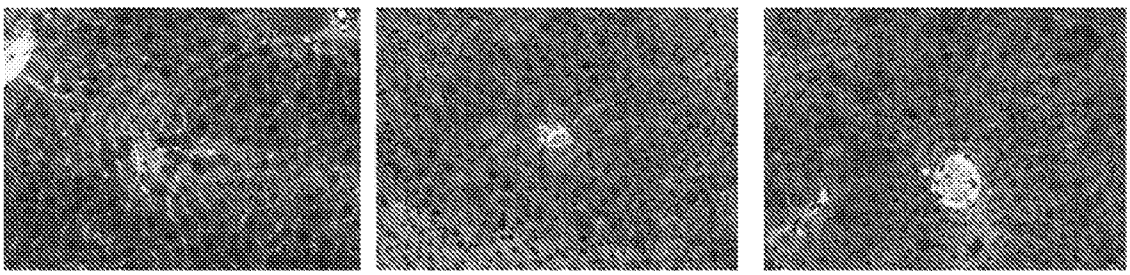


FIG. 29

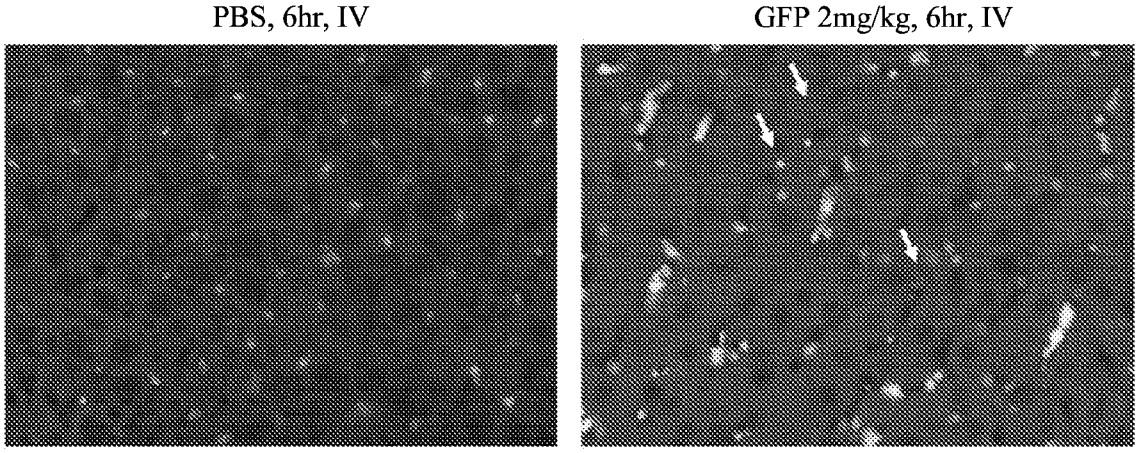
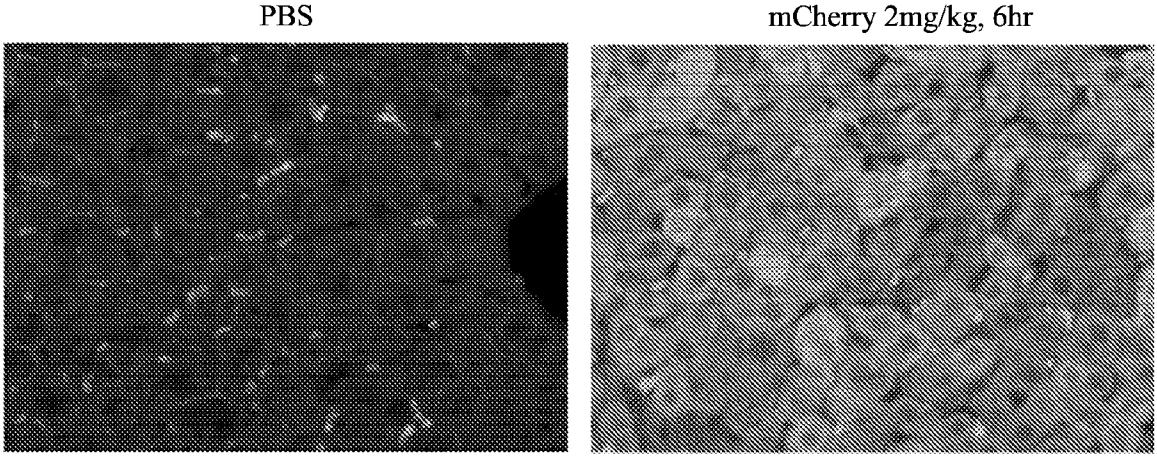


FIG. 30



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/044638

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

   a. (means)

   ☐ on paper

   ☐ in electronic form

   b. (time)

   ☐ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

   An electronic sequence listing was filed but was not used for the purposes of the search and opinion.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/044638

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
the subject matter listed in Rule 39 on which, under Article 17(2)(a)(i), an international search is not required to be carried out, including

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

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

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

See Supplemental Box for Details

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/044638

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/11 (2006.01) C12N 15/88 (2006.01) A61K 31/7105 (2006.01) A61K 8/11 (2006.01) A61K 8/14 (2006.01)

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

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

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

WPIAP, EPODOC, MEDLINE, CAPLUS, BIOSIS, EMBASE & keywords: mRNA, nucleic acid, lipid nanoparticle, splint, self-assembly, uniform distribution, CPC marks: A61K47/488, A61K9/1075, A61K47/48815, A61K9/127 and like terms.

AusPat, The Lens, Google, PubMed, Internal databases: Applicant and Inventor names.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

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

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 October 2016	12 October 2016
Name and mailing address of the ISA/AU	**Authorised officer**
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaustralia.gov.au	Richard Filmer AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. +61 2 6283 2735

INTERNATIONAL SEARCH REPORT C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		International application No. **PCT/US2016/044638**
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/034798 A1 (ALNYLAM PHARMACEUTICALS, INC.) 24 March 2011 page 14 line 16 – page 15 line 2; page 37 lines 26-30; page 40 lines 11-16	1-3, 8-18
X	WO 2014/152774 A1 (SHIRE HUMAN GENETIC THERAPIES INC.) 25 September 2014 [0016], [0055], [0092], [0093], [0096], [0177], [0122]	1, 3-7, 13-17
X	GONG, C. et al., "mRNA–mRNA duplexes that auto-elicit Staufen1-mediated mRNA decay", Nature Structural and Molecular Biology, 2013, vol. 20, no. 10, pages 1214-1220 Abstract; Fig. 1a, 2a, 4a	22-24, 26, 27, 30, 31, 33

Supplemental Box

Continuation of: **Box III**

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

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

- Invention 1 is defined by claims 1-21. The feature of a lipid nanoparticle composition comprising a uniformly distributed first and second nucleic acid is specific to this group of claims.
- Invention 2 is defined by claims 22-33. The feature of a multimeric mRNA structure comprising a first and second mRNA that comprise complementary regions to one another is specific to this group of claims.
- Invention 3 is defined by claims 34-37. The feature of a multimeric mRNA structure comprising a first mRNA and a second mRNA that are non-covalently linked to one another through a splint is specific to this group of claims.

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

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

In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention between Invention 1 and Inventions 2 or 3 are consequently not satisfied *a priori*.

The only common feature to Invention 2 and Invention 3 is a multimeric structure comprising two or more mRNA molecules. However this feature does not make a contribution over the prior art because it is disclosed in:

> GONG, C. et al., "mRNA–mRNA duplexes that auto-elicit Staufen1-mediated mRNA decay", Nature Structural and Molecular Biology, 2013, vol. 20, no. 10, pages 1214-1220

Therefore in the light of this document this common feature cannot be a special technical feature. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a posteriori*.

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. **PCT/US2016/044638**

This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document/s Cited in Search Report		**Patent Family Member/s**	
Publication Number	**Publication Date**	**Publication Number**	**Publication Date**
WO 2011/034798 A1	24 March 2011	WO 2011034798 A1	24 Mar 2011
		CN 102639115 A	15 Aug 2012
		US 2012282341 A1	08 Nov 2012
		US 8859516 B2	14 Oct 2014
		US 2014336243 A1	13 Nov 2014
WO 2014/152774 A1	25 September 2014	WO 2014152774 A1	25 Sep 2014
		AU 2014239184 A1	27 Aug 2015
		CA 2903880 A1	25 Sep 2014
		CN 105209490 A	30 Dec 2015
		EA 201591293 A1	29 Feb 2016
		EP 2970456 A1	20 Jan 2016
		JP 2016513710 A	16 May 2016
		MX 2015011947 A	01 Dec 2015
		US 2016031981 A1	04 Feb 2016

End of Annex

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.