POLYNUCLEOTIDE COMPOSITIONS, RELATED FORMULATIONS, AND METHODS OF USE THEREOF

CROSS-REFERENCE

[0001] This application claims the benefits of U.S. Provisional Application No. 63/164,573 filed March 23, 2021, U.S. Provisional Application No. 63/208,966 filed June 9, 2021, and U.S. Provisional Application No. 63/274,912 filed November 2, 2021, each of which is entirely incorporated herein by reference for all purposes.

BACKGROUND

[0002] Nucleic acids, such as messenger ribonucleic acid(s) (mRNA(s)) may be used by cells to express proteins and polypeptides. Some cells may be deficient in a certain protein or nucleic acid and result in disease states. A cell can also take up and translate exogenous ribonucleic acid(s) (RNA(s)), but many factors influence efficient uptake and translation. For instance, the immune system recognizes many exogenous RNAs as foreign and triggers a response that is aimed at inactivating the RNAs.

SUMMARY

[0003] Provided here are composition and methods for delivery of nucleic acids. Nucleic acids may be used as a therapeutic. In particular, mRNA may be delivered to a cell of a subject. Upon delivery of a nucleic acid to a cell, the nucleic acid may be used to synthesize a polypeptide. In the case of a cell or subject with a disease or disorder, the nucleic acid may be effective at acting as a therapeutic by increasing the expression of a polypeptide. In cases, where a disorder or disease is caused or correlated to aberrant expression or activity of polypeptide, the increased in expression of the polypeptide may be beneficial. However, the cells may have limited uptake of exogenous nucleic acids and the delivery of the nucleic acids may benefit from compositions that allow for increase uptake of a nucleic acid.

[0004] Additionally, therapeutics may benefit from organ specific delivery. Many different types of compounds such as chemotherapeutic agents exhibit significant cytotoxicity. If these compounds were better directed towards delivery to the desired organs, then fewer off target effects will be seen.

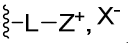
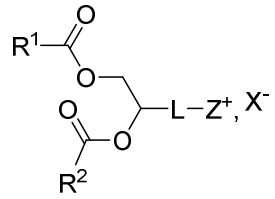
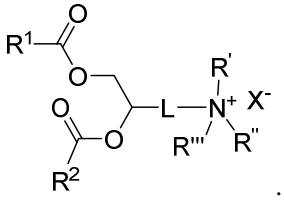
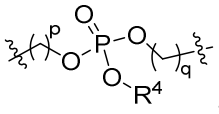
[0005] In an aspect, the present disclosure provides a synthetic polynucleotide encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said synthetic polynucleotide comprises one or more nucleoside analogue(s). In some embodiments, the synthetic polynucleotide comprises 1-methylpseudouridine.

[0006] In an aspect, the present disclosure provides A synthetic polynucleotide encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70%,75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 bases of a sequence selected from SEQ ID NOs: 1-4 and 23. In some embodiments, the said nucleic acid sequence comprises fewer than about 115, 110, 105, 100, 95, or 90 UU or TT dinucleotide. In some embodiments, the nucleic acid sequence

comprises at least two synonymous codons encoding arginine. In some embodiments, the nucleic acid sequence comprises at least three synonymous codons encoding arginine. In some embodiments, no more than about 70%, 65%, 60%, 55%, or 50% of all arginine encoding codons of said nucleic acid sequence is AGA codon. In some embodiments, the nucleic acid sequence encodes a polypeptide that comprises an amino acid sequence having at least 70%, 75%, 80%,81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 contiguous amino acid residues to SEQ ID NO: 5. In some embodiments, the synthetic polynucleotide is a messenger ribonucleic acid (mRNA). In some embodiments, the synthetic polynucleotide further comprises a 3'- or 5'-noncoding region. In some embodiments, the 3'- or 5'-noncoding region enhances an expression or activity of said CFTR protein encoded by said synthetic polynucleotide within a cell. In some embodiments, the synthetic polynucleotide further comprises a 5' cap structure. In some embodiments, the 3' noncoding region comprises a poly adenosine tail. In some embodiments, the poly adenosine tail comprises at most 200 adenosines. In some embodiments, the poly adenosine tail improves a pharmacokinetic characteristic of said synthetic polynucleotide in a cell. In some embodiments, the poly adenosine tail improves a prolonged half-life of said synthetic polynucleotide in a cell.

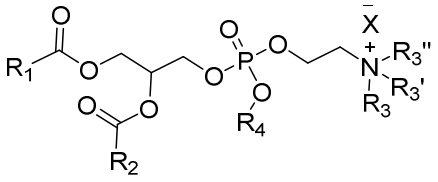
[0007] In an aspect, the present disclosure provides a pharmaceutical composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises: an ionizable cationic lipid; and a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid.

In some embodiments, the lipid composition comprises said ionizable cationic lipid at a molar percentage of about 5% to about 30% In some embodiments, a (e.g., mass or weight) ratio of said ionizable cationic lipid to said synthetic polynucleotide is of no more than about 50:1, 40:1, 30:1, 20:1, 15:1 or 10:1. In some embodiments, the SORT lipid is a permanently cationic lipid. In some embodiments, the SORT lipid is a second ionizable cationic lipid. In some embodiments, the lipid composition comprises said SORT lipid at a molar percentage of about 5% to about 65%. In some embodiments, the lipid composition comprises said SORT lipid at a molar percentage of about 5% to about 30%. In some embodiments, the lipid composition further comprises a zwitterionic lipid (e.g., a phospholipid). In some embodiments, the lipid composition comprises said zwitterionic lipid at a molar percentage of about 5% to about 25%. In some embodiments, a (e.g., mass or weight) ratio of said zwitterionic lipid to said synthetic polynucleotide is of no more than about 50:1, 40:1, 30:1, or 20:1. In some embodiments, the lipid composition further comprises a steroid or steroid derivative. In some embodiments, the lipid composition comprises said steroid or steroid derivative at a molar percentage of about 15% to about 46%. In some embodiments, the lipid composition further comprises a polymer-conjugated lipid (e.g., poly(ethylene glycol) (PEG)-conjugated lipid). In some embodiments, the lipid composition comprises said polymer-conjugated lipid at a molar percentage of about 0.5% to about 10% In some embodiments, a

molar ratio of nitrogen in said lipid composition to phosphate in said synthetic polynucleotide (N/P ratio) is of no more than about 50:1, 40:1, 30:1, or 20:1 In some embodiments, the N/P ratio is from about 5:1 to about 30:1. In some embodiments, the a (e.g., mass or weight) ratio of said synthetic polynucleotide to total lipids of said lipid composition is no more than about 1:20, 1:50, or 1:100. In some embodiments, the SORT lipid comprises a permanently positively charged moiety (e.g., a quaternary ammonium ion). In some embodiments, the SORT lipid comprises a counterion. In some embodiments, the SORT lipid is a phosphocholine lipid (e.g., saturated or unsaturated). In some embodiments, the SORT lipid is an ethylphosphocholine. In some embodiments, the SORT lipid comprises a headgroup having a structural formula: , wherein L is a (e.g., biodegradable) linker; Z^+ is positively charged moiety (e.g., a quaternary ammonium ion); and X^- is a counterion. In some embodiments, the SORT lipid has a structural formula: , wherein R^1 and R^2 are each independently an optionally substituted C_6-C_{24} alkyl, or an optionally substituted C_6-C_{24} alkenyl. In some embodiments, the SORT lipid has a structural formula: . In some embodiments, L is , wherein:

p and q are each independently 1, 2, or 3; and

R^4 is an optionally substituted C_1-C_6 alkyl. In some embodiments, the SORT lipid has a structural formula:



(IA), wherein:

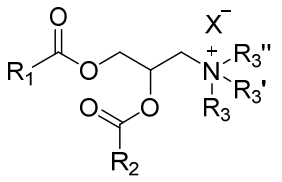
R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq 6)}$ or substituted $alkyl_{(C\leq 6)}$;

R_4 is $alkyl_{(C\leq 6)}$ or substituted $alkyl_{(C\leq 6)}$; and

X^- is a monovalent anion.

In some embodiments, the SORT lipid has a structural formula:



(S-I')

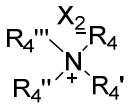
wherein:

R_1 and R_2 are each independently $alkyl_{(C8-C24)}$, $alkenyl_{(C8-C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$;

X^- is a monovalent anion.

In some embodiments, the SORT lipid has a structural formula:



(S-II')

wherein:

R_4 and R_4' are each independently $alkyl_{(C6-C24)}$, $alkenyl_{(C6-C24)}$, or a substituted version of either group;

R_4'' is $alkyl_{(C\leq24)}$, $alkenyl_{(C\leq24)}$, or a substituted version of either group;

R_4''' is $alkyl_{(C1-C8)}$, $alkenyl_{(C2-C8)}$, or a substituted version of either group; and

X_2 is a monovalent anion.

[0008] In some embodiments, the pharmaceutical composition is an aerosol composition. In some embodiments, the aerosol composition has a droplet size from 0.5 micron (μm) to 10 μm. In some embodiments, the aerosol composition has a median droplet size from 0.5 μm to 10 μm. In some embodiments, the aerosol composition has an average droplet size from 0.5 μm to 10 μm. In some embodiments, the pharmaceutical composition is formulated for aerosol administration. In some embodiments, the pharmaceutical composition is formulated for apical delivery. In some embodiments, the pharmaceutical composition is formulated for nebulization.

[0009] In another aspect, the present disclosure provides a method for enhancing an expression or activity of cystic fibrosis transmembrane conductance regulator (CFTR) protein in a cell, the method comprising: contacting said cell with a composition comprising a synthetic polynucleotide assembled with a lipid composition, wherein said synthetic polynucleotide encodes a CFTR protein; and wherein said lipid composition comprises: an ionizable cationic lipid; and a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid, thereby yielding a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after contacting, optionally wherein said therapeutically effective activity of said functional variant of CFTR protein is determined by measuring a change in a transepithelial ion transport characteristic of a plurality of cells comprising said cell as compared to that of a reference plurality of cells in absence of said contacting. In some embodiments, the contacting is repeated In some embodiments, the contacting is at least once a week. In some embodiments,

the contacting is at least twice a week. In some embodiments, the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after each contacting. In some embodiments, the contacting is a first contacting, and wherein the method comprises a second contacting, optionally, performed at least 1, 2, or 3 day(s) after said first contacting. In some embodiments, the method further comprises a third contacting, optionally wherein said third contacting is performed at least 1, 2, or 3 day(s) after said second contacting. In some embodiments, the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after said second contacting. In some embodiments, the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after said third contacting. In some embodiments, the contacting comprises administering to a subject said composition comprising said synthetic polynucleotide assembled with said lipid composition. In some embodiments, the subject is a mammal. In some embodiments, the subject is a human. In some embodiments, the administering comprises inhalation by nebulization. In some embodiments, the composition in each contacting is identical. In some embodiments, the cell is a lung airway cell. In some embodiments, the cell is a lung secretory cell or a lung basal cell. The lung basal cell may be a lung basal stem cell. In some embodiments, the cell is a bronchial epithelial cell In some embodiments, the cell is undifferentiated. In some embodiments, the cell is differentiated. In some embodiments, the cell is derived from said subject. In some embodiments, the contacting is *in vivo*. In some embodiments, the contacting is *in vitro*. In some embodiments, the contacting is *ex vivo*. In some embodiments, the functional variant of CFTR protein is a wild-type CFTR protein. In some embodiments, the functional variant of CFTR protein is a full-length CFTR protein. In some embodiments, the therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least about 5 micro-Ampere (μA), e.g., as determined in an *in vitro* assay. In some embodiments, the therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from about 5 micro-Ampere (μA) to about 30 μA. In some embodiments, the therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$), e.g., as determined in an *in vitro* assay. In some embodiments, the therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$) to about 20 $\mu A \cdot cm^{-2} \cdot min^{-1}$. In some embodiments, the method increases an amount or activity of said functional variant of CFTR protein in said cell (e.g., by at least about 1.1-fold) relative to a corresponding control (e.g., that of a corresponding cell absent said contacting). In some embodiments, the method enhances (e.g., chloride) ion transport in said cell (e.g., by at least about 1.1-fold) relative to a corresponding control (e.g., that of a corresponding cell absent said contacting). In some embodiments, the subject exhibits or is determined to exhibit a mutation in a cystic fibrosis transmembrane conductance regulator (CFTR) gene. In some embodiments, the mutation is a loss-of-function mutation. In some embodiments, the mutation is a nonsense or frameshift mutation. In some embodiments, the mutation is in one or more of exons 11-27 of CFTR gene.

In some embodiments, the mutation is R553X, G542X or F508del, or a combination thereof. In some embodiments, the mutation is R1162X. In some embodiments, the mutation is R553X, G542X, F508del, or R1162X, or a combination thereof.

[0010] In another aspect the present disclosure provides a method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises: an ionizable cationic lipid; and a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid, thereby yielding a therapeutically effective amount or activity of said synthetic polynucleotide in a lung secretory cell or lung basal cell of said subject, optionally wherein said therapeutically effective activity of said synthetic polynucleotide is determined by measuring a change in a transepithelial ion transport characteristic of a lung comprising said lung secretory cell or lung basal cell as compared to that of a reference lung in absence of said contacting. The lung basal cell may be a lung basal stem cell.

[0011] In another aspect, the present disclosure provides a method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises: an ionizable cationic lipid; and a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid, thereby yielding a greater therapeutic amount or activity of said synthetic polynucleotide in a lung secretory cell or lung basal cell of said subject as compared to that in a lung non-secretory cell or lung non-basal cell of said subject. In some embodiments of the method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), at least about 50%, 55%, or 60% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected or observed in lung secretory cells, lung basal cells, or a combination thereof, e.g., as determined by measuring an amount or activity of the corresponding polypeptide encoded by the synthetic polynucleotide. In some embodiments of the method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), no more than about 50%, 45%, or 40% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected or observed in lung non-secretory cells, lung non-basal cells, or a combination thereof, e.g., as determined by measuring an amount or activity of the corresponding polypeptide encoded by the synthetic polynucleotide. In some embodiments of the method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), no more than about 50%, 45%, or 40% of (e.g., pulmonary) expression of said synthetic polynucleotide is in lung ciliated cells, e.g., as determined by measuring an amount or activity of the corresponding polypeptide encoded by the synthetic polynucleotide. In some embodiments, the method for lung secretory cell or lung basal cell delivery yields an amount or activity of said synthetic

polynucleotide in lung secretory cell(s) or lung basal cell(s) that is at least 1.1-, 1.5-, or 2-fold greater than that in reference cell(s), which reference cell(s) are neither lung secretory cell(s) nor lung basal cell(s). The reference cell(s) may be lung ciliated cell(s). In some embodiments, the lung non-secretory cell or lung non-basal cell is a lung ciliated cell. In some embodiments, the lung non-secretory cell is a lung basal cell. The (e.g., lung) basal cell may be a (e.g., lung) basal stem cell.

[0012] In another aspect, the present disclosure provides a method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises: an ionizable cationic lipid; and a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid, thereby yielding a therapeutic amount or activity of said synthetic polynucleotide in at least 5% of lung secretory cells or lung basal cells of said subject. In some embodiments, the administering comprises administering to a lung of said subject said composition comprising said synthetic polynucleotide assembled with said lipid composition. In some embodiments, the lung secretory cell is a club cell or a goblet cell. The (e.g., lung) basal cell may be a (e.g., lung) basal stem cell.

[0013] In another aspect, the present disclosure provides a method for treating a subject having or suspected of having a cystic fibrosis transmembrane conductance regulator (CFTR)-associated condition, the method comprising administering to said subject a pharmaceutical composition disclosed elsewhere herein. In some embodiments, the CFTR-associated condition is cystic fibrosis, hereditary emphysema, or chronic obstructive pulmonary disease (COPD). In some embodiments, the administering comprises local administration. In some embodiments, the administering comprises nebulization.

[0014] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0015] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0017] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings (also "figure" and "FIG." herein), of which:

[0018] **FIG. 1** shows the chemical structures of example lipids.

[0019] **FIG. 2** shows the chemical structures of example dendrimer lipids.

[0020] **FIG. 3** shows a chart of cells type and expression levels of a delivered mRNA using different compositions of LNP.

[0021] **FIG. 4** illustrates images using *in vivo* imaging of bioluminescence of a mouse after inhaled aerosol delivery of a reporter Luc mRNA/LNP using multiple compositions of LNP.

[0022] **FIG. 5** shows a chart regarding cell toxicity of various LNP compositions in human bronchial epithelial (hBE) cells.

[0023] **FIG. 6** illustrates the stability and general characteristics of various LNP compositions.

[0024] **FIG. 7** shows a chart of tissue specific radiance over time in a mouse of an LNP composition (e.g., 5A2-SC8 DOTAP).

[0025] **FIG. 8** shows images of tissue specific radiance over time in a mouse of an LNP composition (e.g., 5A2-SC8 DOTAP).

[0026] **FIG. 9A** illustrates a structural design of CFTR mRNA described in the present application; and **FIG. 9B** illustrates production of CFTR mRNA described in the present application.

[0027] **FIG. 10A-10B** illustrate dose-dependent expression of CFTR protein in FRT cells.

[0028] **FIG. 11A** illustrates activity of CFTR in FRT cells.

[0029] **FIG. 11B-11C** show that does-dependent CFTR function was observed with CFTR mRNA described herein in FRT cells. **FIG. 11B** shows that 5 day-old confluent FRT cells grown on TransWell ® permeable support were transfected with ReCode-optimized mRNAs using Lipofectamine 2000. MTECC24 assay of the transepithelial conductance was performed 1 day after transfection with one dose of the CFTR mRNA described herein. **FIG. 11C** shows mRNA dose dependent transepithelial conductance (Gt) responses: bars were Gt area under the curve (AUC) per min between forskolin addition and Inhibitor-172 addition time points.

[0030] **FIG. 12** illustrates delivery of reporter mRNA into fully differentiated hBE cells.

[0031] **FIG. 13A-13C** show resistance and response of hBE cells to reference compounds and CFTR mRNA formulation of the present application w/tdTomato mRNA.

[0032] **FIG. 14A-14B** illustrates restoration of CFTR function in fully differentiated F508del/F508del hBEs by CFTR mRNA formulation of the present application.

[0033] **FIG. 15A-15B** illustrates restoration of CFTR function in R553X/F508del hBEs by CFTR mRNA formulation of the present application.

[0034] **FIG. 16A-16B** illustrates single dose aerosol treatment with CFTR mRNA LNPS described herein rescuing CFTR function in in primary CF hBE cells. **FIG. 16A** shows representative traces of Forskolin-induced G542X/F508 hBE cells using CFTR mRNA formulation of the present application after single dosing.

[0035] **FIG. 17A-17B** show repeat administration of the CFTR mRNA LNP formulation based on a twice weekly dosing schedule. Using a similar protocol to determine the CFTR function, the repeated administration showed CFTR function after each dose. FIG. 17A and FIG. 17B show that each dose was able to generate improved CFTR function over a negative control.

[0036] **FIG. 18** shows transfection of hBE cells with report mRNA reveals formulation-specific cell tropism signatures. Top graphs shows that well-differentiated human hBE cells were treated once with RTX0001 formulated Td Tomato mRNA (4 mg) using Vitrocell nebulization. % positive cells were determined by colocalization with the indicated markers. As used herein, "RTX001" refers to an example lipid composition tested herein. RTX0001 was a 5-component lipid composition comprising about 19.05% 4A3-SC7 (ionizable cationic lipid), about 20% DODAP (SORT lipid), about 19.05% DOPE, about 38.9% cholesterol, and about 3.81% DMG-PEG (PEG conjugated lipid), wherein each lipid component is defined as mol% of the total lipid composition.

[0037] **FIG. 19** shows an overview of a clinical trial in human subjects for treating cystic fibrosis with compositions disclosed herein.

DETAILED DESCRIPTION

[0038] While various embodiments of the invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions may occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed.

[0039] The term "polynucleotide" or "nucleic acid" as used herein generally refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides, that comprise purine and pyrimidine bases, purine and pyrimidine analogues, chemically or biochemically modified, natural or non-natural, or derivatized nucleotide bases. Polynucleotides include sequences of deoxyribonucleic acid (DNA), ribonucleic acid (RNA), or DNA copies of ribonucleic acid (cDNA), all of which can be recombinantly produced, artificially synthesized, or isolated and purified from natural sources. The polynucleotides and nucleic acids may exist as single-stranded or double-stranded. The backbone of the polynucleotide can comprise sugars and phosphate groups, as may typically be found in RNA or DNA, or analogues or substituted sugar or phosphate groups. A polynucleotide may comprise naturally occurring or non-naturally occurring nucleotides, such as methylated nucleotides and nucleotide analogues (or analogs).

[0040] The term "polyribonucleotide," as used herein, generally refers to polynucleotide polymers that comprise ribonucleic acids. The term also refers to polynucleotide polymers that comprise chemically modified ribonucleotides. A polyribonucleotide can be formed of D-ribose sugars, which can be found in nature.

[0041] The term "polypeptides," as used herein, generally refers to polymer chains comprised of amino acid residue monomers which are joined together through amide bonds (peptide bonds). A polypeptide can be a chain of at least three amino acids, a protein, a recombinant protein, an antigen, an epitope, an enzyme, a receptor, or a structure analogue or combinations thereof. As used herein, the abbreviations for the L-enantiomeric amino acids that form a polypeptide are as follows: alanine (A, Ala); arginine (R, Arg); asparagine (N, Asn); aspartic acid (D, Asp); cysteine (C, Cys); glutamic acid (E, Glu); glutamine (Q, Gln); glycine (G, Gly); histidine (H, His); isoleucine (I, Ile); leucine (L, Leu); lysine (K, Lys); methionine (M, Met); phenylalanine (F, Phe); proline (P, Pro); serine (S, Ser); threonine (T, Thr); tryptophan (W, Trp); tyrosine (Y, Tyr); valine (V, Val). X or Xaa can indicate any amino acid.

[0042] The term "engineered," as used herein, generally refers to polynucleotides, vectors, and nucleic acid constructs that have been genetically designed and manipulated to provide a polynucleotide intracellularly. An engineered polynucleotide can be partially or fully synthesized *in vitro*. An engineered polynucleotide can also be cloned. An engineered polyribonucleotide can contain one or more base or sugar analogues, such as ribonucleotides not naturally-found in messenger RNAs. An engineered polyribonucleotide can contain nucleotide analogues that exist in transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), guide RNAs (gRNAs), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), SmY RNA, spliced leader RNA (SL RNA), CRISPR RNA, long untranslated RNA (lncRNA), microRNA (miRNA), or another suitable RNA.

[0043] As used herein, the term "patient" or "subject" refers to a living mammalian organism, such as a human, monkey, cow, sheep, goat, dog, cat, mouse, rat, guinea pig, or transgenic species thereof. In certain embodiments, the patient or subject is a primate (e.g., non-human primate). In certain embodiments, the patient or subject is a human. Non-limiting examples of human subjects are adults, juveniles, infants and fetuses.

[0044] The term "assemble" or "assembled," as used herein, in context of delivery of a payload to target cell(s) generally refers to covalent or non-covalent interaction(s) or association(s), for example, such that a therapeutic or prophylactic agent be complexed with or encapsulated in a lipid composition.

[0045] As used herein, the term "lipid composition" generally refers to a composition comprising lipid compound(s), including but not limited to, a lipoplex, a liposome, a lipid particle. Example of lipid compositions include suspensions, emulsions, and vesicular compositions.

[0046] As used herein, the term "detectable" refers to an occurrence of, or a change in, a signal that is directly or indirectly detectable either by observation or by instrumentation. Typically, a detectable response is an occurrence of a signal wherein the fluorophore is inherently fluorescent and does not produce a change in signal upon binding to a metal ion or biological compound. Alternatively, the detectable

response is an optical response resulting in a change in the wavelength distribution patterns or intensity of absorbance or fluorescence or a change in light scatter, fluorescence lifetime, fluorescence polarization, or a combination of the above parameters. Other detectable responses include, for example, chemiluminescence, phosphorescence, radiation from radioisotopes, magnetic attraction, and electron density.

[0047] Unless otherwise indicated, all numbers expressing quantities, ranges, conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about." Accordingly, unless indicated to the contrary, the numerical parameters set forth in the present specification and attached claims are approximations that can vary depending upon the desired properties sought to be obtained by the present application. Generally the term "about", as used herein when referring to a measurable value such as an amount of weight, time, dose, etc. is meant to encompass in one example variations of ±20% or ±10%, in another example ±5%, in another example ±1%, and in yet another example ±0.1% from the specified amount, as such variations are appropriate to perform the disclosed method.

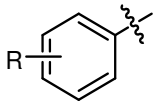
[0048] As used herein, the term "ratio" generally refers to the relative amount of one or more molecules to another molecule(s). Non-limiting examples of the ratio(s) include molar ratio(s), weight ratio(s), or mass ratio(s).

[0049] When used in the context of a chemical group: "hydrogen" means −H; "hydroxy" means −OH; "oxo" means =O; "carbonyl" means −C(=O)−; "carboxy" means −C(=O)OH (also written as −COOH or $-CO_2H$); "halo" means independently −F, −Cl, −Br or −I; "amino" means $-NH_2$; "hydroxyamino" means –NHOH; "nitro" means $-NO_2$; imino means =NH; "cyano" means −CN; "isocyanate" means −N=C=O; "azido" means $-N_3$; in a monovalent context "phosphate" means $-OP(O)(OH)_2$ or a deprotonated form thereof; in a divalent context "phosphate" means $-OP(O)(OH)O-$ or a deprotonated form thereof; "mercapto" means −SH; and "thio" means =S; "sulfonyl" means $-S(O)_2-$; "hydroxysulfonyl" means $-S(O)_2OH$; "sulfonamide" means $-S(O)_2NH_2$; and "sulfinyl" means –S(O)–.

[0050] In the context of chemical formulas, the symbol "–" means a single bond, "=" means a double bond, and "≡" means triple bond. The symbol "----" represents an optional bond, which if present is either single or double. The symbol "----" represents a single bond or a double bond. Thus, for example, the formula  includes , , ,  and . And it is understood that no one such ring atom forms part of more than one double bond. Furthermore, it is noted that the covalent bond symbol "–", when connecting one or two stereogenic atoms, does not indicate any preferred stereochemistry. Instead, it covers all stereoisomers as well as mixtures thereof. The symbol "", when drawn perpendicularly across a bond (*e.g.*, for methyl) indicates a point of attachment of the group. It is noted that the point of attachment is typically only identified in this manner for larger groups in order to assist the reader in unambiguously identifying a point of attachment. The symbol "" means a single bond where the group attached to the thick end of the wedge is "out of the page." The symbol "" means a single bond where the group attached to the thick end of the wedge is "into the page". The symbol "" means a single

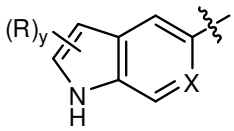
bond where the geometry around a double bond (*e.g.*, either *E* or *Z*) is undefined. Both options, as well as combinations thereof are therefore intended. Any undefined valency on an atom of a structure shown in this application implicitly represents a hydrogen atom bonded to that atom. A bold dot on a carbon atom indicates that the hydrogen attached to that carbon is oriented out of the plane of the paper.

[0051] When a group "R" is depicted as a "floating group" on a ring system, for example, in the formula:



,

then R may replace any hydrogen atom attached to any of the ring atoms, including a depicted, implied, or expressly defined hydrogen, so long as a stable structure is formed. When a group "R" is depicted as a "floating group" on a fused ring system, as for example in the formula:



,

then R may replace any hydrogen attached to any of the ring atoms of either of the fused rings unless specified otherwise. Replaceable hydrogens include depicted hydrogens (*e.g.*, the hydrogen attached to the nitrogen in the formula above), implied hydrogens (*e.g.*, a hydrogen of the formula above that is not shown but understood to be present), expressly defined hydrogens, and optional hydrogens whose presence depends on the identity of a ring atom (*e.g.,* a hydrogen attached to group X, when X equals –CH–), so long as a stable structure is formed. In the example depicted, R may reside on either the 5-membered or the 6-membered ring of the fused ring system. In the formula above, the subscript letter "y" immediately following the group "R" enclosed in parentheses, represents a numeric variable. Unless specified otherwise, this variable can be 0, 1, 2, or any integer greater than 2, only limited by the maximum number of replaceable hydrogen atoms of the ring or ring system.

[0052] For the chemical groups and compound classes, the number of carbon atoms in the group or class is as indicated as follows: "Cn" defines the exact number (n) of carbon atoms in the group/class. "C≤n" defines the maximum number (n) of carbon atoms that can be in the group/class, with the minimum number as small as possible for the group/class in question, *e.g.*, it is understood that the minimum number of carbon atoms in the group "alkenyl$_{(C\leq 8)}$" or the class "alkene$_{(C\leq 8)}$" is two. Compare with "alkoxy$_{(C\leq 10)}$", which designates alkoxy groups having from 1 to 10 carbon atoms. "Cn-n′" defines both the minimum (n) and maximum number (n′) of carbon atoms in the group. Thus, "alkyl$_{(C2\text{-}10)}$" designates those alkyl groups having from 2 to 10 carbon atoms. These carbon number indicators may precede or follow the chemical groups or class it modifies and it may or may not be enclosed in parenthesis, without signifying any change in meaning. Thus, the terms "C5 olefin", "C5-olefin", "olefin$_{(C5)}$", and "olefin$_{C5}$" are all synonymous.

[0053] The term "saturated" when used to modify a compound or chemical group means the compound or chemical group has no carbon-carbon double and no carbon-carbon triple bonds, except as noted below.

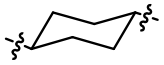
When the term is used to modify an atom, it means that the atom is not part of any double or triple bond. In the case of substituted versions of saturated groups, one or more carbon oxygen double bond or a carbon nitrogen double bond may be present. And when such a bond is present, then carbon-carbon double bonds that may occur as part of keto-enol tautomerism or imine/enamine tautomerism are not precluded. When the term "saturated" is used to modify a solution of a substance, it means that no more of that substance can dissolve in that solution.

[0054] The term "aliphatic" when used without the "substituted" modifier signifies that the compound or chemical group so modified is an acyclic or cyclic, but non-aromatic hydrocarbon compound or group. In aliphatic compounds/groups, the carbon atoms can be joined together in straight chains, branched chains, or non-aromatic rings (alicyclic). Aliphatic compounds/groups can be saturated, that is joined by single carbon-carbon bonds (alkanes/alkyl), or unsaturated, with one or more carbon-carbon double bonds (alkenes/alkenyl) or with one or more carbon-carbon triple bonds (alkynes/alkynyl).

[0055] The term "aromatic" when used to modify a compound or a chemical group atom means the compound or chemical group contains a planar unsaturated ring of atoms that is stabilized by an interaction of the bonds forming the ring.

[0056] The term "alkyl" when used without the "substituted" modifier refers to a monovalent saturated aliphatic group with a carbon atom as the point of attachment, a linear or branched acyclic structure, and no atoms other than carbon and hydrogen. The groups $-CH_3$ (Me), $-CH_2CH_3$ (Et), $-CH_2CH_2CH_3$ (*n*-Pr or propyl), $-CH(CH_3)_2$ (*i*-Pr, ^{i}Pr or isopropyl), $-CH_2CH_2CH_2CH_3$ (*n*-Bu), $-CH(CH_3)CH_2CH_3$ (*sec*-butyl), $-CH_2CH(CH_3)_2$ (isobutyl), $-C(CH_3)_3$ (*tert*-butyl, *t*-butyl, *t*-Bu or ^{t}Bu), and $-CH_2C(CH_3)_3$ (*neo*-pentyl) are non-limiting examples of alkyl groups. The term "alkanediyl" when used without the "substituted" modifier refers to a divalent saturated aliphatic group, with one or two saturated carbon atom(s) as the point(s) of attachment, a linear or branched acyclic structure, no carbon-carbon double or triple bonds, and no atoms other than carbon and hydrogen. The groups $-CH_2-$ (methylene), $-CH_2CH_2-$, $-CH_2C(CH_3)_2CH_2-$, and $-CH_2CH_2CH_2-$ are non-limiting examples of alkanediyl groups. An "alkane" refers to the class of compounds having the formula H–R, wherein R is alkyl as this term is defined above. When any of these terms is used with the "substituted" modifier one or more hydrogen atom has been independently replaced by $-OH$, $-F$, $-Cl$, $-Br$, $-I$, $-NH_2$, $-NO_2$, $-CO_2H$, $-CO_2CH_3$, $-CN$, $-SH$, $-OCH_3$, $-OCH_2CH_3$, $-C(O)CH_3$, $-NHCH_3$, $-NHCH_2CH_3$, $-N(CH_3)_2$, $-C(O)NH_2$, $-C(O)NHCH_3$, $-C(O)N(CH_3)_2$, $-OC(O)CH_3$, $-NHC(O)CH_3$, $-S(O)_2OH$, or $-S(O)_2NH_2$. The following groups are non-limiting examples of substituted alkyl groups: $-CH_2OH$, $-CH_2Cl$, $-CF_3$, $-CH_2CN$, $-CH_2C(O)OH$, $-CH_2C(O)OCH_3$, $-CH_2C(O)NH_2$, $-CH_2C(O)CH_3$, $-CH_2OCH_3$, $-CH_2OC(O)CH_3$, $-CH_2NH_2$, $-CH_2N(CH_3)_2$, and $-CH_2CH_2Cl$. The term "haloalkyl" is a subset of substituted alkyl, in which the hydrogen atom replacement is limited to halo (*i.e.* $-F$, $-Cl$, $-Br$, or $-I$) such that no other atoms aside from carbon, hydrogen and halogen are present. The group, $-CH_2Cl$ is a non-limiting example of a haloalkyl. The term "fluoroalkyl" is a subset of substituted alkyl, in which the hydrogen atom replacement is limited to fluoro such that no other atoms aside from

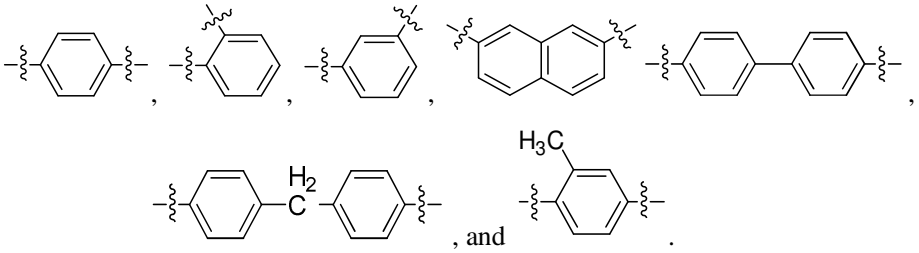
carbon, hydrogen and fluorine are present. The groups $-CH_2F$, $-CF_3$, and $-CH_2CF_3$ are non-limiting examples of fluoroalkyl groups.

[0057] The term "cycloalkyl" when used without the "substituted" modifier refers to a monovalent saturated aliphatic group with a carbon atom as the point of attachment, said carbon atom forming part of one or more non-aromatic ring structures, no carbon-carbon double or triple bonds, and no atoms other than carbon and hydrogen. Non-limiting examples include: $-CH(CH_2)_2$ (cyclopropyl), cyclobutyl, cyclopentyl, or cyclohexyl (Cy). The term "cycloalkanediyl" when used without the "substituted" modifier refers to a divalent saturated aliphatic group with two carbon atoms as points of attachment, no carbon-carbon double or triple bonds, and no atoms other than carbon and hydrogen. The group  is a non-limiting example of cycloalkanediyl group. A "cycloalkane" refers to the class of compounds having the formula H–R, wherein R is cycloalkyl as this term is defined above. When any of these terms is used with the "substituted" modifier one or more hydrogen atom has been independently replaced by $-OH$, $-F$, $-Cl$, $-Br$, $-I$, $-NH_2$, $-NO_2$, $-CO_2H$, $-CO_2CH_3$, $-CN$, $-SH$, $-OCH_3$, $-OCH_2CH_3$, $-C(O)CH_3$, $-NHCH_3$, $-NHCH_2CH_3$, $-N(CH_3)_2$, $-C(O)NH_2$, $-C(O)NHCH_3$, $-C(O)N(CH_3)_2$, $-OC(O)CH_3$, $-NHC(O)CH_3$, $-S(O)_2OH$, or $-S(O)_2NH_2$.

[0058] The term "alkenyl" when used without the "substituted" modifier refers to an monovalent unsaturated aliphatic group with a carbon atom as the point of attachment, a linear or branched, acyclic structure, at least one nonaromatic carbon-carbon double bond, no carbon-carbon triple bonds, and no atoms other than carbon and hydrogen. Non-limiting examples include: $-CH=CH_2$ (vinyl), $-CH=CHCH_3$, $-CH=CHCH_2CH_3$, $-CH_2CH=CH_2$ (allyl), $-CH_2CH=CHCH_3$, and $-CH=CHCH=CH_2$. The term "alkenediyl" when used without the "substituted" modifier refers to a divalent unsaturated aliphatic group, with two carbon atoms as points of attachment, a linear or branched, a linear or branched acyclic structure, at least one nonaromatic carbon-carbon double bond, no carbon-carbon triple bonds, and no atoms other than carbon and hydrogen. The groups $-CH=CH-$, $-CH=C(CH_3)CH_2-$, $-CH=CHCH_2-$, and $-CH_2CH=CHCH_2-$ are non-limiting examples of alkenediyl groups. It is noted that while the alkenediyl group is aliphatic, once connected at both ends, this group is not precluded from forming part of an aromatic structure. The terms "alkene" and "olefin" are synonymous and refer to the class of compounds having the formula H–R, wherein R is alkenyl as this term is defined above. Similarly, the terms "terminal alkene" and "α-olefin" are synonymous and refer to an alkene having just one carbon-carbon double bond, wherein that bond is part of a vinyl group at an end of the molecule. When any of these terms are used with the "substituted" modifier one or more hydrogen atom has been independently replaced by $-OH$, $-F$, $-Cl$, $-Br$, $-I$, $-NH_2$, $-NO_2$, $-CO_2H$, $-CO_2CH_3$, $-CN$, $-SH$, $-OCH_3$, $-OCH_2CH_3$, $-C(O)CH_3$, $-NHCH_3$, $-NHCH_2CH_3$, $-N(CH_3)_2$, $-C(O)NH_2$, $-C(O)NHCH_3$, $-C(O)N(CH_3)_2$, $-OC(O)CH_3$, $-NHC(O)CH_3$, $-S(O)_2OH$, or $-S(O)_2NH_2$. The groups $-CH=CHF$, $-CH=CHCl$ and $-CH=CHBr$ are non-limiting examples of substituted alkenyl groups.

[0059] The term "alkynyl" when used without the "substituted" modifier refers to a monovalent unsaturated aliphatic group with a carbon atom as the point of attachment, a linear or branched acyclic structure, at least one carbon-carbon triple bond, and no atoms other than carbon and hydrogen. As used herein, the term alkynyl does not preclude the presence of one or more non-aromatic carbon-carbon double bonds. The groups $-C\equiv CH$, $-C\equiv CCH_3$, and $-CH_2C\equiv CCH_3$ are non-limiting examples of alkynyl groups. An "alkyne" refers to the class of compounds having the formula H–R, wherein R is alkynyl. When any of these terms are used with the "substituted" modifier one or more hydrogen atom has been independently replaced by $-OH$, $-F$, $-Cl$, $-Br$, $-I$, $-NH_2$, $-NO_2$, $-CO_2H$, $-CO_2CH_3$, $-CN$, $-SH$, $-OCH_3$, $-OCH_2CH_3$, $-C(O)CH_3$, $-NHCH_3$, $-NHCH_2CH_3$, $-N(CH_3)_2$, $-C(O)NH_2$, $-C(O)NHCH_3$, $-C(O)N(CH_3)_2$, $-OC(O)CH_3$, $-NHC(O)CH_3$, $-S(O)_2OH$, or $-S(O)_2NH_2$.

[0060] The term "aryl" when used without the "substituted" modifier refers to a monovalent unsaturated aromatic group with an aromatic carbon atom as the point of attachment, said carbon atom forming part of a one or more six-membered aromatic ring structure, wherein the ring atoms are all carbon, and wherein the group consists of no atoms other than carbon and hydrogen. If more than one ring is present, the rings may be fused or unfused. As used herein, the term does not preclude the presence of one or more alkyl or aralkyl groups (carbon number limitation permitting) attached to the first aromatic ring or any additional aromatic ring present. Non-limiting examples of aryl groups include phenyl (Ph), methylphenyl, (dimethyl)phenyl, $-C_6H_4CH_2CH_3$ (ethylphenyl), naphthyl, and a monovalent group derived from biphenyl. The term "arenediyl" when used without the "substituted" modifier refers to a divalent aromatic group with two aromatic carbon atoms as points of attachment, said carbon atoms forming part of one or more six-membered aromatic ring structure(s) wherein the ring atoms are all carbon, and wherein the monovalent group consists of no atoms other than carbon and hydrogen. As used herein, the term does not preclude the presence of one or more alkyl, aryl or aralkyl groups (carbon number limitation permitting) attached to the first aromatic ring or any additional aromatic ring present. If more than one ring is present, the rings may be fused or unfused. Unfused rings may be connected via one or more of the following: a covalent bond, alkanediyl, or alkenediyl groups (carbon number limitation permitting). Non-limiting examples of arenediyl groups include:



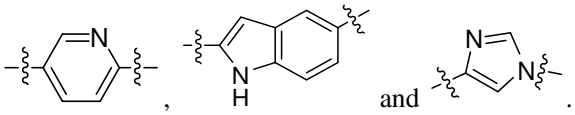
, and .

An "arene" refers to the class of compounds having the formula H–R, wherein R is aryl as that term is defined above. Benzene and toluene are non-limiting examples of arenes. When any of these terms are used with the "substituted" modifier one or more hydrogen atom has been independently replaced by $-OH$, $-F$, $-Cl$, $-Br$, $-I$, $-NH_2$, $-NO_2$, $-CO_2H$, $-CO_2CH_3$, $-CN$, $-SH$, $-OCH_3$, $-OCH_2CH_3$, $-C(O)CH_3$, $-NHCH_3$,

$-NHCH_2CH_3$, $-N(CH_3)_2$, $-C(O)NH_2$, $-C(O)NHCH_3$, $-C(O)N(CH_3)_2$, $-OC(O)CH_3$, $-NHC(O)CH_3$, $-S(O)_2OH$, or $-S(O)_2NH_2$.

[0061] The term "aralkyl" when used without the "substituted" modifier refers to the monovalent group −alkanediyl−aryl, in which the terms alkanediyl and aryl are each used in a manner consistent with the definitions provided above. Non-limiting examples are: phenylmethyl (benzyl, Bn) and 2-phenyl-ethyl. When the term aralkyl is used with the "substituted" modifier one or more hydrogen atom from the alkanediyl and/or the aryl group has been independently replaced by $-OH$, $-F$, $-Cl$, $-Br$, $-I$, $-NH_2$, $-NO_2$, $-CO_2H$, $-CO_2CH_3$, $-CN$, $-SH$, $-OCH_3$, $-OCH_2CH_3$, $-C(O)CH_3$, $-NHCH_3$, $-NHCH_2CH_3$, $-N(CH_3)_2$, $-C(O)NH_2$, $-C(O)NHCH_3$, $-C(O)N(CH_3)_2$, $-OC(O)CH_3$, $-NHC(O)CH_3$, $-S(O)_2OH$, or $-S(O)_2NH_2$. Non-limiting examples of substituted aralkyls are: (3-chlorophenyl)-methyl, and 2-chloro-2-phenyl-eth-1-yl.

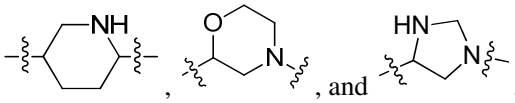
[0062] The term "heteroaryl" when used without the "substituted" modifier refers to a monovalent aromatic group with an aromatic carbon atom or nitrogen atom as the point of attachment, said carbon atom or nitrogen atom forming part of one or more aromatic ring structures wherein at least one of the ring atoms is nitrogen, oxygen or sulfur, and wherein the heteroaryl group consists of no atoms other than carbon, hydrogen, aromatic nitrogen, aromatic oxygen and aromatic sulfur. Heteroaryl rings may contain 1, 2, 3, or 4 ring atoms selected from are nitrogen, oxygen, and sulfur. If more than one ring is present, the rings may be fused or unfused. As used herein, the term does not preclude the presence of one or more alkyl, aryl, and/or aralkyl groups (carbon number limitation permitting) attached to the aromatic ring or aromatic ring system. Non-limiting examples of heteroaryl groups include furanyl, imidazolyl, indolyl, indazolyl (Im), isoxazolyl, methylpyridinyl, oxazolyl, phenylpyridinyl, pyridinyl (pyridyl), pyrrolyl, pyrimidinyl, pyrazinyl, quinolyl, quinazolyl, quinoxalinyl, triazinyl, tetrazolyl, thiazolyl, thienyl, and triazolyl. The term "*N*-heteroaryl" refers to a heteroaryl group with a nitrogen atom as the point of attachment. The term "heteroarenediyl" when used without the "substituted" modifier refers to an divalent aromatic group, with two aromatic carbon atoms, two aromatic nitrogen atoms, or one aromatic carbon atom and one aromatic nitrogen atom as the two points of attachment, said atoms forming part of one or more aromatic ring structure(s) wherein at least one of the ring atoms is nitrogen, oxygen or sulfur, and wherein the divalent group consists of no atoms other than carbon, hydrogen, aromatic nitrogen, aromatic oxygen and aromatic sulfur. If more than one ring is present, the rings may be fused or unfused. Unfused rings may be connected via one or more of the following: a covalent bond, alkanediyl, or alkenediyl groups (carbon number limitation permitting). As used herein, the term does not preclude the presence of one or more alkyl, aryl, and/or aralkyl groups (carbon number limitation permitting) attached to the aromatic ring or aromatic ring system. Non-limiting examples of heteroarenediyl groups include:



A "heteroarene" refers to the class of compounds having the formula H−R, wherein R is heteroaryl. Pyridine and quinoline are non-limiting examples of heteroarenes. When these terms are used with the "substituted"

modifier one or more hydrogen atom has been independently replaced by –OH, –F, –Cl, –Br, –I, $–NH_2$, $–NO_2$, $–CO_2H$, $–CO_2CH_3$, –CN, –SH, $–OCH_3$, $–OCH_2CH_3$, $–C(O)CH_3$, $–NHCH_3$, $–NHCH_2CH_3$, $–N(CH_3)_2$, $–C(O)NH_2$, $–C(O)NHCH_3$, $–C(O)N(CH_3)_2$, $–OC(O)CH_3$, $–NHC(O)CH_3$, $–S(O)_2OH$, or $–S(O)_2NH_2$.

[0063] The term "heterocycloalkyl" when used without the "substituted" modifier refers to a monovalent non-aromatic group with a carbon atom or nitrogen atom as the point of attachment, said carbon atom or nitrogen atom forming part of one or more non-aromatic ring structures wherein at least one of the ring atoms is nitrogen, oxygen or sulfur, and wherein the heterocycloalkyl group consists of no atoms other than carbon, hydrogen, nitrogen, oxygen and sulfur. Heterocycloalkyl rings may contain 1, 2, 3, or 4 ring atoms selected from nitrogen, oxygen, or sulfur. If more than one ring is present, the rings may be fused or unfused. As used herein, the term does not preclude the presence of one or more alkyl groups (carbon number limitation permitting) attached to the ring or ring system. Also, the term does not preclude the presence of one or more double bonds in the ring or ring system, provided that the resulting group remains non-aromatic. Non-limiting examples of heterocycloalkyl groups include aziridinyl, azetidinyl, pyrrolidinyl, piperidinyl, piperazinyl, morpholinyl, thiomorpholinyl, tetrahydrofuranyl, tetrahydrothiofuranyl, tetrahydropyranyl, pyranyl, oxiranyl, and oxetanyl. The term "*N*-heterocycloalkyl" refers to a heterocycloalkyl group with a nitrogen atom as the point of attachment. *N*-pyrrolidinyl is an example of such a group. The term "heterocycloalkanediyl" when used without the "substituted" modifier refers to an divalent cyclic group, with two carbon atoms, two nitrogen atoms, or one carbon atom and one nitrogen atom as the two points of attachment, said atoms forming part of one or more ring structure(s) wherein at least one of the ring atoms is nitrogen, oxygen or sulfur, and wherein the divalent group consists of no atoms other than carbon, hydrogen, nitrogen, oxygen and sulfur. If more than one ring is present, the rings may be fused or unfused. Unfused rings may be connected via one or more of the following: a covalent bond, alkanediyl, or alkenediyl groups (carbon number limitation permitting). As used herein, the term does not preclude the presence of one or more alkyl groups (carbon number limitation permitting) attached to the ring or ring system. Also, the term does not preclude the presence of one or more double bonds in the ring or ring system, provided that the resulting group remains non-aromatic. Non-limiting examples of heterocycloalkanediyl groups include:



, , and .

When these terms are used with the "substituted" modifier one or more hydrogen atom has been independently replaced by –OH, –F, –Cl, –Br, –I, $–NH_2$, $–NO_2$, $–CO_2H$, $–CO_2CH_3$, –CN, –SH, $–OCH_3$, $–OCH_2CH_3$, $–C(O)CH_3$, $–NHCH_3$, $–NHCH_2CH_3$, $–N(CH_3)_2$, $–C(O)NH_2$, $–C(O)NHCH_3$, $–C(O)N(CH_3)_2$, $–OC(O)CH_3$, $–NHC(O)CH_3$, $–S(O)_2OH$, or $–S(O)_2NH_2$.

[0064] The term "acyl" when used without the "substituted" modifier refers to the group –C(O)R, in which R is a hydrogen, alkyl, cycloalkyl, alkenyl, aryl, aralkyl or heteroaryl, as those terms are defined above. The groups, –CHO, $–C(O)CH_3$ (acetyl, Ac), $–C(O)CH_2CH_3$, $–C(O)CH_2CH_2CH_3$, $–C(O)CH(CH_3)_2$,

−C(O)CH(CH_2)$_2$, −C(O)C_6H_5, −C(O)$C_6H_4CH_3$, −C(O)$CH_2C_6H_5$, −C(O)(imidazolyl) are non-limiting examples of acyl groups. A "thioacyl" is defined in an analogous manner, except that the oxygen atom of the group −C(O)R has been replaced with a sulfur atom, −C(S)R. The term "aldehyde" corresponds to an alkane, as defined above, wherein at least one of the hydrogen atoms has been replaced with a −CHO group. When any of these terms are used with the "substituted" modifier one or more hydrogen atom (including a hydrogen atom directly attached to the carbon atom of the carbonyl or thiocarbonyl group, if any) has been independently replaced by −OH, −F, −Cl, −Br, −I, −NH_2, −NO_2, −CO_2H, −CO_2CH_3, −CN, −SH, −OCH_3, −OCH_2CH_3, −C(O)CH_3, −$NHCH_3$, −$NHCH_2CH_3$, −N(CH_3)$_2$, −C(O)NH_2, −C(O)$NHCH_3$, −C(O)N(CH_3)$_2$, −OC(O)CH_3, −$NHC(O)CH_3$, −S(O)$_2$OH, or −S(O)$_2NH_2$. The groups, −C(O)CH_2CF_3, −CO_2H (carboxyl), −CO_2CH_3 (methylcarboxyl), −$CO_2CH_2CH_3$, −C(O)NH_2 (carbamoyl), and −CON(CH_3)$_2$, are non-limiting examples of substituted acyl groups.

[0065] The term "alkoxy" when used without the "substituted" modifier refers to the group −OR, in which R is an alkyl, as that term is defined above. Non-limiting examples include: −OCH_3 (methoxy), −OCH_2CH_3 (ethoxy), −$OCH_2CH_2CH_3$, −OCH(CH_3)$_2$ (isopropoxy), −OC(CH_3)$_3$ (*tert*-butoxy), −OCH(CH_2)$_2$, −O−cyclopentyl, and −O−cyclohexyl. The terms "cycloalkoxy", "alkenyloxy", "alkynyloxy", "aryloxy", "aralkoxy", "heteroaryloxy", "heterocycloalkoxy", and "acyloxy", when used without the "substituted" modifier, refers to groups, defined as −OR, in which R is cycloalkyl, alkenyl, alkynyl, aryl, aralkyl, heteroaryl, heterocycloalkyl, and acyl, respectively. The term "alkoxydiyl" refers to the divalent group −O−alkanediyl−, −O−alkanediyl−O−, or −alkanediyl−O−alkanediyl−. The term "alkylthio" and "acylthio" when used without the "substituted" modifier refers to the group −SR, in which R is an alkyl and acyl, respectively. The term "alcohol" corresponds to an alkane, as defined above, wherein at least one of the hydrogen atoms has been replaced with a hydroxy group. The term "ether" corresponds to an alkane, as defined above, wherein at least one of the hydrogen atoms has been replaced with an alkoxy group. When any of these terms is used with the "substituted" modifier one or more hydrogen atom has been independently replaced by −OH, −F, −Cl, −Br, −I, −NH_2, −NO_2, −CO_2H, −CO_2CH_3, −CN, −SH, −OCH_3, −OCH_2CH_3, −C(O)CH_3, −$NHCH_3$, −$NHCH_2CH_3$, −N(CH_3)$_2$, −C(O)NH_2, −C(O)$NHCH_3$, −C(O)N(CH_3)$_2$, −OC(O)CH_3, −$NHC(O)CH_3$, −S(O)$_2$OH, or −S(O)$_2NH_2$.

[0066] The term "alkylamino" when used without the "substituted" modifier refers to the group −NHR, in which R is an alkyl, as that term is defined above. Non-limiting examples include: −$NHCH_3$ and −$NHCH_2CH_3$. The term "dialkylamino" when used without the "substituted" modifier refers to the group −NRR′, in which R and R′ can be the same or different alkyl groups, or R and R′ can be taken together to represent an alkanediyl. Non-limiting examples of dialkylamino groups include: −N(CH_3)$_2$ and −N(CH_3)(CH_2CH_3). The terms "cycloalkylamino", "alkenylamino", "alkynylamino", "arylamino", "aralkylamino", "heteroarylamino", "heterocycloalkylamino", "alkoxyamino", and "alkylsulfonylamino" when used without the "substituted" modifier, refers to groups, defined as −NHR, in which R is cycloalkyl, alkenyl, alkynyl, aryl, aralkyl, heteroaryl, heterocycloalkyl, alkoxy, and alkylsulfonyl, respectively. A non-limiting example of an arylamino group is −NHC_6H_5. The term "alkylaminodiyl" refers to the divalent

group –NH–alkanediyl–, –NH–alkanediyl–NH–, or –alkanediyl–NH–alkanediyl–. The term "amido" (acylamino), when used without the "substituted" modifier, refers to the group –NHR, in which R is acyl, as that term is defined above. A non-limiting example of an amido group is $–NHC(O)CH_3$. The term "alkylimino" when used without the "substituted" modifier refers to the divalent group =NR, in which R is an alkyl, as that term is defined above. When any of these terms is used with the "substituted" modifier one or more hydrogen atom attached to a carbon atom has been independently replaced by –OH, –F, –Cl, –Br, –I, $–NH_2$, $–NO_2$, $–CO_2H$, $–CO_2CH_3$, –CN, –SH, $–OCH_3$, $–OCH_2CH_3$, $–C(O)CH_3$, $–NHCH_3$, $–NHCH_2CH_3$, $–N(CH_3)_2$, $–C(O)NH_2$, $–C(O)NHCH_3$, $–C(O)N(CH_3)_2$, $–OC(O)CH_3$, $–NHC(O)CH_3$, $–S(O)_2OH$, or $–S(O)_2NH_2$. The groups $–NHC(O)OCH_3$ and $–NHC(O)NHCH_3$ are non-limiting examples of substituted amido groups.

[0067] The use of the word "a" or "an," when used in conjunction with the term "comprising" in the claims and/or the specification may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one."

[0068] As used in this application, the term "average molecular weight" refers to the relationship between the number of moles of each polymer species and the molar mass of that species. In particular, each polymer molecule may have different levels of polymerization and thus a different molar mass. The average molecular weight can be used to represent the molecular weight of a plurality of polymer molecules. Average molecular weight is typically synonymous with average molar mass. In particular, there are three major types of average molecular weight: number average molar mass, weight (mass) average molar mass, and Z-average molar mass. In the context of this application, unless otherwise specified, the average molecular weight represents either the number average molar mass or weight average molar mass of the formula. In some embodiments, the average molecular weight is the number average molar mass. In some embodiments, the average molecular weight may be used to describe a PEG component present in a lipid.

[0069] The terms "comprise," "have" and "include" are open-ended linking verbs. Any forms or tenses of one or more of these verbs, such as "comprises," "comprising," "has," "having," "includes" and "including," are also open-ended. For example, any method that "comprises," "has" or "includes" one or more steps is not limited to possessing only those one or more steps and also covers other unlisted steps.

[0070] The term "effective," as that term is used in the specification and/or claims, means adequate to accomplish a desired, expected, or intended result. "Effective amount," "Therapeutically effective amount" or "pharmaceutically effective amount" when used in the context of treating a patient or subject with a compound means that amount of the compound which, when administered to a subject or patient for treating a disease, is sufficient to effect such treatment for the disease.

[0071] As used herein, the term "IC_{50}" refers to an inhibitory dose which is 50% of the maximum response obtained. This quantitative measure indicates how much of a particular drug or other substance (inhibitor) is needed to inhibit a given biological, biochemical or chemical process (or component of a process, i.e. an enzyme, cell, cell receptor or microorganism) by half.

[0072] An "isomer" of a first compound is a separate compound in which each molecule contains the same constituent atoms as the first compound, but where the configuration of those atoms in three dimensions differs.

[0073] As used herein, the term "patient" or "subject" refers to a living mammalian organism, such as a human, monkey, cow, sheep, goat, dog, cat, mouse, rat, guinea pig, or transgenic species thereof. In certain embodiments, the patient or subject is a primate. Non-limiting examples of human subjects are adults, juveniles, infants and fetuses.

[0074] As generally used herein "pharmaceutically acceptable" refers to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues, organs, and/or bodily fluids of human beings and animals without excessive toxicity, irritation, allergic response, or other problems or complications commensurate with a reasonable benefit/risk ratio.

[0075] "Pharmaceutically acceptable salts" means salts of compounds of the present disclosure which are pharmaceutically acceptable, as defined above, and which possess the desired pharmacological activity. Such salts include acid addition salts formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid, and the like; or with organic acids such as 1,2-ethanedisulfonic acid, 2-hydroxyethanesulfonic acid, 2-naphthalenesulfonic acid, 3-phenylpropionic acid, 4,4'-methylenebis(3-hydroxy-2-ene-1-carboxylic acid), 4-methylbicyclo[2.2.2]oct-2-ene-1-carboxylic acid, acetic acid, aliphatic mono- and dicarboxylic acids, aliphatic sulfuric acids, aromatic sulfuric acids, benzenesulfonic acid, benzoic acid, camphorsulfonic acid, carbonic acid, cinnamic acid, citric acid, cyclopentanepropionic acid, ethanesulfonic acid, fumaric acid, glucoheptonic acid, gluconic acid, glutamic acid, glycolic acid, heptanoic acid, hexanoic acid, hydroxynaphthoic acid, lactic acid, laurylsulfuric acid, maleic acid, malic acid, malonic acid, mandelic acid, methanesulfonic acid, muconic acid, *o*-(4-hydroxybenzoyl)benzoic acid, oxalic acid, *p*-chlorobenzenesulfonic acid, phenyl-substituted alkanoic acids, propionic acid, *p*-toluenesulfonic acid, pyruvic acid, salicylic acid, stearic acid, succinic acid, tartaric acid, tertiarybutylacetic acid, trimethylacetic acid, and the like. Pharmaceutically acceptable salts also include base addition salts which may be formed when acidic protons present are capable of reacting with inorganic or organic bases. Acceptable inorganic bases include sodium hydroxide, sodium carbonate, potassium hydroxide, aluminum hydroxide and calcium hydroxide. Acceptable organic bases include ethanolamine, diethanolamine, triethanolamine, tromethamine, *N*-methylglucamine and the like. It should be recognized that the particular anion or cation forming a part of any salt of this disclosure is not critical, so long as the salt, as a whole, is pharmacologically acceptable. Additional examples of pharmaceutically acceptable salts and their methods of preparation and use are presented in *Handbook of Pharmaceutical Salts: Properties, and Use* (P. H. Stahl & C. G. Wermuth eds., Verlag Helvetica Chimica Acta, 2002).

[0076] "Prevention" or "preventing" includes: (1) inhibiting the onset of a disease in a subject or patient which may be at risk and/or predisposed to the disease but does not yet experience or display any or all of

the pathology or symptomatology of the disease, and/or (2) slowing the onset of the pathology or symptomatology of a disease in a subject or patient which may be at risk and/or predisposed to the disease but does not yet experience or display any or all of the pathology or symptomatology of the disease.

[0077] A "repeat unit" is the simplest structural entity of certain materials, for example, frameworks and/or polymers, whether organic, inorganic or metal-organic. In the case of a polymer chain, repeat units are linked together successively along the chain, like the beads of a necklace. For example, in polyethylene, -[-CH_2CH_2-]$_n$-, the repeat unit is –CH_2CH_2–. The subscript "n" denotes the degree of polymerization, that is, the number of repeat units linked together. When the value for "n" is left undefined or where "n" is absent, it simply designates repetition of the formula within the brackets as well as the polymeric nature of the material. The concept of a repeat unit applies equally to where the connectivity between the repeat units extends three dimensionally, such as in metal organic frameworks, modified polymers, thermosetting polymers, etc. Within the context of the dendrimer, the repeating unit may also be described as the branching unit, interior layers, or generations. Similarly, the terminating group may also be described as the surface group.

[0078] A "stereoisomer" or "optical isomer" is an isomer of a given compound in which the same atoms are bonded to the same other atoms, but where the configuration of those atoms in three dimensions differs. "Enantiomers" are stereoisomers of a given compound that are mirror images of each other, like left and right hands. "Diastereomers" are stereoisomers of a given compound that are not enantiomers. Chiral molecules contain a chiral center, also referred to as a stereocenter or stereogenic center, which is any point, though not necessarily an atom, in a molecule bearing groups such that an interchanging of any two groups leads to a stereoisomer. In organic compounds, the chiral center is typically a carbon, phosphorus or sulfur atom, though it is also possible for other atoms to be stereocenters in organic and inorganic compounds. A molecule can have multiple stereocenters, giving it many stereoisomers. In compounds whose stereoisomerism is due to tetrahedral stereogenic centers (*e.g.*, tetrahedral carbon), the total number of hypothetically possible stereoisomers will not exceed 2^n, where n is the number of tetrahedral stereocenters. Molecules with symmetry frequently have fewer than the maximum possible number of stereoisomers. A 50:50 mixture of enantiomers is referred to as a racemic mixture. Alternatively, a mixture of enantiomers can be enantiomerically enriched so that one enantiomer is present in an amount greater than 50%. Typically, enantiomers and/or diastereomers can be resolved or separated using techniques known in the art. It is contemplated that that for any stereocenter or axis of chirality for which stereochemistry has not been defined, that stereocenter or axis of chirality can be present in its *R* form, *S* form, or as a mixture of the *R* and *S* forms, including racemic and non-racemic mixtures. As used herein, the phrase "substantially free from other stereoisomers" means that the composition contains ≤ 15%, more preferably ≤ 10%, even more preferably ≤ 5%, or most preferably ≤ 1% of another stereoisomer(s).

[0079] "Treatment" or "treating" includes (1) inhibiting a disease in a subject or patient experiencing or displaying the pathology or symptomatology of the disease (*e.g.*, arresting further development of the pathology and/or symptomatology), (2) ameliorating a disease in a subject or patient that is experiencing

or displaying the pathology or symptomatology of the disease (*e.g.*, reversing the pathology and/or symptomatology), and/or (3) effecting any measurable decrease in a disease in a subject or patient that is experiencing or displaying the pathology or symptomatology of the disease.

[0080] The term "molar percentage" or "molar %" as used herein in connection with lipid composition(s) generally refers to the molar proportion of that component lipid relative to compared to all lipids formulated or present in the lipid composition.

[0081] The above definitions supersede any conflicting definition in any reference that is incorporated by reference herein. The fact that certain terms are defined, however, should not be considered as indicative that any term that is undefined is indefinite. Rather, all terms used are believed to describe the disclosure in terms such that one of ordinary skill can appreciate the scope and practice the present disclosure.

[0082] The present disclosure provides, in some embodiments, compositions and methods for the treatment of conditions associated with cilia maintenance and function, with nucleic acids encoding a protein or protein fragment(s). Numerous eukaryotic cells carry appendages, which are often referred to as cilia or flagella, whose inner core comprises a cytoskeletal structure called the axoneme. The axoneme can function as the skeleton of cellular cytoskeletal structures, both giving support to the structure and, In some embodiments, causing it to bend. Usually, the internal structure of the axoneme is common to both cilia and flagella. Cilia are often found in the linings of the airway, the reproductive system, and other organs and tissues. Flagella are tail-like structures that, similarly to cilia, can propel cells forward, such as sperm cells.

[0083] Without properly functioning cilia in the airway, bacteria can remain in the respiratory tract and cause infection. In the respiratory tract, cilia move back and forth in a coordinated way to move mucus towards the throat. This movement of mucus helps to eliminate fluid, bacteria, and particles from the lungs. Many infants afflicted with cilia and flagella malfunction experience breathing problems at birth, which suggests that cilia play an important role in clearing fetal fluid from the lungs. Beginning in early childhood, subjects afflicted with cilia malfunction can develop frequent respiratory tract infections.

Cystic fibrosis transmembrane conductance regulator (CFTR)

[0084] Cystic fibrosis transmembrane conductance regulator (CFTR) is a membrane protein and chloride channel in vertebrates encoded by the CFTR gene. CFTR gene is on the long arm of chromosome 7, at position q31.2. Mutations of the CFTR gene affecting chloride ion channel function led to dysregulation of epithelial fluid transport in the lung, pancreas and other organs, resulting in cystic fibrosis (CF).

[0085] Cystic fibrosis (CF) affects approximately one in every 2,500 infants in the United States. Within the general United States population, up to 10 million people carry a single copy of the defective gene without apparent ill effects. In contrast, individuals with two copies of the CF associated gene suffer from the debilitating and fatal effects of CF, including chronic lung disease. Complications of cystic fibrosis include thickened mucus in the lungs with frequent respiratory infections, and pancreatic insufficiency giving rise to malnutrition and diabetes. These conditions lead to chronic disability and reduced life expectancy. In male patients, the progressive obstruction and destruction of the developing vas deferens

(spermatic cord) and epididymis appear to result from abnormal intraluminal secretions, causing congenital absence of the vas deferens and male infertility.

[0086] So far, nearly 1000 cystic fibrosis-causing mutations have been described. Many mutations are infrequent. The distribution and frequency of mutations varies among different populations. Mutations consist of replacements, duplications, deletions, or shortenings in the CFTR gene. This may result in dysfunctional proteins which have less activity, are more quickly degraded or present in inadequate numbers. The most common mutation, DeltaF508 (ΔF508) results from a deletion (Δ) of three nucleotides which results in a loss of the amino acid phenylalanine (F) at the 508th position on the protein. As a result, the protein does not fold normally and is more quickly degraded.

COMPOSITIONS

[0087] In some embodiments, the present disclosure provides a (e.g., pharmaceutical) composition comprising a (e.g., synthetic) polynucleotide encoding CFTR protein as described herein. In some embodiments of the composition, the polynucleotide is assembled with a lipid composition (such as described herein).

POLYNUCLEOTIDES

[0088] In some embodiments, the synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein. In some embodiments, the synthetic polynucleotide is a ribonucleic acid (RNA), e.g., a messenger ribonucleic acid (mRNA), encoding a CFTR protein. In some embodiments, the synthetic polynucleotide is a deoxyribonucleic acid (DNA) encoding a CFTR protein.

[0089] In some embodiments of various aspects, the nucleic acid sequence encodes a polypeptide that comprises an amino acid sequence having at least 70%, 75%, 80%,81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 contiguous amino acid residues to **SEQ ID NO: 5**. In some embodiments of various aspects, the nucleic acid sequence encodes a polypeptide substantially identical to **SEQ ID NO: 5**. In some embodiments, said nucleic acid sequence encodes a polypeptide substantially identical over at least 1,000 contiguous amino acid residues to **SEQ ID NO. 5**. In some embodiments, said nucleic acid sequence encodes a polypeptide substantially identical to **SEQ ID NO. 5**. In some embodiments, said nucleic acid sequence encodes a polypeptide that comprises an amino acid sequence having at least 70%, 75%, 80%,81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 contiguous amino acid residues to **SEQ ID NO: 5**. In some embodiments, said nucleic acid sequence encodes a polypeptide substantially identical to **SEQ ID NO. 5**.

Hydrolysis hot spots

[0090] Altered nucleotide usage schemes aiming to reduce the number of more reactive 5'-U(U/A)-3' dinucleotides within codons as well as across codons of modified mRNAs partially alleviate limitations imposed by the inherent chemical instability of RNA. At the same time, lowering the U-content in RNA

transcripts renders them less immunogenic. The present disclosure relates to RNA transcripts comprising altered open reading frames (ORF). For example, the codon optimized or altered nucleotide usage may comprise a substantial reduction of 5'-U(U/A)-3' dinucleotides within protein coding regions leading to stabilized therapeutic mRNAs. The codon optimized polynucleotide may comprise a codon coding for a particular amino acid to be substituted or replaced of a with a synonymous codon. The codon optimized polynucleotide may encode a same or identical polypeptide as a corresponding wild type polynucleotide, with the polynucleotide comprising a different sequence of polynucleotide than the corresponding wild type. Multiple codons may encode for a same amino acid, however the qualities of a given codon are differ between even those that code for a same amino acid. Because multiple different codons may code for a same amino acid, a particular polynucleotide may encode for a same polypeptide and have advantageous features over another polynucleotide that codes for the same polypeptide. For example, a codon optimized polynucleotides may be translated faster, may comprise a higher stability (*in vivo* or *in vitro)*, may result in increased expression yield or full length or functional polypeptides, or may result in an increase of soluble polypeptide and a decrease in polypeptide aggregates. . Without being limited to a specific mechanism, the advantageous features of a codon optimized polynucleotides may be for example, a result of improved protein folding of the expressed product based on ribosomal interactions with the polynucleotides or may be result of decreased hydrolysis of reactive bonds in solution. For example, the codon optimization may alter or improve characteristics relating to ribosomal binding sites, Shine-Dalgarno sequences, or ribosomal or translational pausing. The advantageous features may be a result of decreased usage of “rare codons” which may have a lower concentration of cognate tRNAs, allowing for an improved translation reaction. The advantageous features may be a result of decreased usage of “rare codons” which may have a lower concentration of cognate tRNAs, allowing for an improved translation reaction. The advantageous features may be a result of decreasing degradation via enzymatic reaction. For example, hydrolysis of oligonucleotides suggests that the reactivity of the phosphodiester bond linking two ribonucleotides in single-stranded (ss)RNA depends on the nature of those nucleotides. At pH 8.5, dinucleotide cleavage susceptibility when embedded in ssRNA dodecamers may vary by an order of magnitude. Under near physiological conditions, hydrolysis of RNA usually involves an S_N2-type attack by the 2'-oxygen nucleophile on the adjacent phosphorus target center on the opposing side of the 5'-oxyanion leaving group, yielding two RNA fragments with 2',3'-cyclic phosphate and 5'-hydroxyl termini. More reactive scissile phosphodiester bonds may include 5'-UpA-3' ($R_1 = U_1$, $R_2 = A$) and 5'-CpA-3' ($R_1 = C$, $R_2 = A$) because the backbone at these steps can most easily adopt the "in-line" conformation that is required for S_N2-type nucleophilic attack by the 2'-OH on the adjacent phosphodiester linkage. In addition, interferon-regulated dsRNA-activated antiviral pathways produce 2'-5' oligoadenylates which bind to ankyrin repeats leading to activation of RNase L endoribonuclease. RNase L cleaves ssRNA efficiently at UA and UU dinucleotides. Lastly, U-rich sequences are potent activators of RNA sensors including Toll-like receptor 7 and 8 and RIG-I making global uridine content reduction a potentially attractive approach to reduce immunogenicity of therapeutic mRNAs.

[0091] In some cases, the number or percent of UU and UA sequences in the polynucleotide are below a certain threshold. For example, the percent of dinucleotide sequences comprising UU and UA may be less than 30%, 25%, 20%, 15%, 10%, 5% or less in the polynucleotides. In some case the number of UU or UA in a sequence may be less than 50, 45, 40, 35, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or less in the polynucleotide.

[0092] In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 115 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 110 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 105 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 100 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 95 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 90 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 85 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 80 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 75 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 70 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 65 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 60 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 55 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 50 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 45 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 40 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 35 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 45 UU or TT dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 30 UU or TT dinucleotide.

[0093] In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 115 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 110 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 105 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 100 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 95 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 90 UA or TA dinucleotide. In some embodiments of the

synthetic polynucleotide, the polynucleotide comprises fewer than about 85 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 80 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 75 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 70 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 65 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 60 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 55 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 50 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 45 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 40 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 35 UA or TA dinucleotide. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 30 UA or TA dinucleotide.

[0094] In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 200 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 195 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 190 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 185 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 180 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 175 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 170 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 165 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 160 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 155 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 150 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 145 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 140 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 135 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 130 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 125 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 120 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 115 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than

about 110 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 105 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 100 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 95 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 90 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 85 of UU and UA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 80 of UU and UA.

[0095] In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 200 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 195 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 190 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 185 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 180 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 175 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 170 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 165 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 160 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 155 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 150 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 145 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 140 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 135 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 130 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 125 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 120 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 115 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 110 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 105 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 100 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 95 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 90 of TT and TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 85 of TT and

TA. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises fewer than about 80 of TT and TA.

Codon usage

[0096] In some embodiments of the synthetic polynucleotide, the polynucleotide comprises at least two synonymous codons encoding arginine. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises at least two synonymous codons encoding arginine, and said codon is selected from the group consisting of AGG, AGA, CGG, CGA, CGT and CGC. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises at least three synonymous codons encoding arginine, and said codon is selected from the group consisting of AGG, AGA, CGG, CGA, CGT and CGC. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises at least four synonymous codons encoding arginine, and said codon is selected from the group consisting of AGG, AGA, CGG, CGA, CGT and CGC. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises at least five synonymous codons encoding arginine, and said codon is selected from the group consisting of AGG, AGA, CGG, CGA, CGT and CGC. In some embodiments of the synthetic polynucleotide, the polynucleotide comprises four synonymous codons encoding arginine, and said codon is selected from the group consisting of AGG, AGA, CGG and CGC.

[0097] In some embodiments of the synthetic polynucleotide, no more than about 70% of all arginine encoding codons of said nucleic acid sequence is AGG codon. In some embodiments of the synthetic polynucleotide, no more than about 65% of all arginine encoding codons of said nucleic acid sequence is AGG codon. In some embodiments of the synthetic polynucleotide, no more than about 60% of all arginine encoding codons of said nucleic acid sequence is AGG codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is AGG codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is AGG codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is AGG codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is AGG codon.

[0098] In some embodiments of the synthetic polynucleotide, no more than about 70% of all arginine encoding codons of said nucleic acid sequence is AGA codon. In some embodiments of the synthetic polynucleotide, no more than about 65% of all arginine encoding codons of said nucleic acid sequence is AGA codon. In some embodiments of the synthetic polynucleotide, no more than about 60% of all arginine encoding codons of said nucleic acid sequence is AGA codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is AGA codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is AGA codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is

AGA codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is AGA codon.

[0099] In some embodiments of the synthetic polynucleotide, no more than about 70% of all arginine encoding codons of said nucleic acid sequence is CGG codon. In some embodiments of the synthetic polynucleotide, no more than about 65% of all arginine encoding codons of said nucleic acid sequence is CGG codon. In some embodiments of the synthetic polynucleotide, no more than about 60% of all arginine encoding codons of said nucleic acid sequence is CGG codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGG codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGG codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGG codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGG codon.

[00100] In some embodiments of the synthetic polynucleotide, no more than about 70% of all arginine encoding codons of said nucleic acid sequence is CGA codon. In some embodiments of the synthetic polynucleotide, no more than about 65% of all arginine encoding codons of said nucleic acid sequence is CGA codon. In some embodiments of the synthetic polynucleotide, no more than about 60% of all arginine encoding codons of said nucleic acid sequence is CGA codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGA codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGA codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGA codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGA codon.

[00101] In some embodiments of the synthetic polynucleotide, no more than about 70% of all arginine encoding codons of said nucleic acid sequence is CGT (or CGU) codon. In some embodiments of the synthetic polynucleotide, no more than about 65% of all arginine encoding codons of said nucleic acid sequence is CGT (or CGU) codon. In some embodiments of the synthetic polynucleotide, no more than about 60% of all arginine encoding codons of said nucleic acid sequence is CGT (or CGU) codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGT (or CGU) codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGT (or CGU) codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGT (or CGU) codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGT (or CGU) codon.

[00102] In some embodiments of the synthetic polynucleotide, no more than about 70% of all arginine encoding codons of said nucleic acid sequence is CGC codon. In some embodiments of the synthetic polynucleotide, no more than about 65% of all arginine encoding codons of said nucleic acid sequence is CGC codon. In some embodiments of the synthetic polynucleotide, no more than about 60% of all arginine encoding codons of said nucleic acid sequence is CGC codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGC codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGC codon. In some embodiments of the synthetic polynucleotide, no more than about 55% of all arginine encoding codons of said nucleic acid sequence is CGC codon. In some embodiments of the synthetic polynucleotide, no more than about 50% of all arginine encoding codons of said nucleic acid sequence is CGC codon.

[00103] In some embodiments of the synthetic polynucleotide as described herein, the polynucleotides may comprise an open reading frame (ORF) sequence. The ORF sequence may be characterized by a codon usage profile comprising: (1) a total number of codons, (2) a species number of codons (e.g. a total number of different codon types), (3) a number of each (unique) codon, and (4) a (usage) frequency of each codon among all synonymous codons (if present). The codon usage profile may be altered or compared to a corresponding wild type sequence. For example, the frequency or number of particular codons may be reduced or increased compared to a wild type sequence. The change in codon frequency of the polynucleotide may provide benefits over the wild type sequence. For example, the altered codon frequency may result in a less immunogenic polynucleotide. The polynucleotide with an altered codon frequency may result in a polynucleotide that is more quickly expressed or results in a greater amount of expression product. The polynucleotide with an altered codon frequency may have increase stability, such as increased half-life in sera, or may be less susceptible to hydrolysis or other reactions that may result in the degradation of the polynucleotide.

[00104] In some embodiments, the polynucleotide comprises an altered nucleotide usage as compared to a corresponding wild type sequence. The altered nucleotide usage may also be referred to as a "codon optimized" sequence or be generated by way of "codon optimization".

[00105] In some cases, a codon coding for a particular amino acid in the polypeptide may be substituted or replaced with a synonymous codon. For example, a codon coding for leucine may be substituted for another codon coding for leucine. In this way, the resulting translation products may be identical with the polynucleotide differing in sequence. At least one type of an isoleucine-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence. At least one type of a valine-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence. At least one type of an alanine-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence. At least one type of a glycine-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence. At

least one type of a proline-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence. At least one type of a threonine-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence. At least one type of a leucine-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence. At least one type of an arginine-encoding codons in said corresponding wild-type sequence is substituted with a synonymous codon type in said nucleic acid sequence. At least one type of a serine-encoding codons in said corresponding wild-type sequence may be substituted with a synonymous codon type in said nucleic acid sequence.

[00106] In some embodiments, a particular codon of a particular amino acid comprises a percentage or amount of the total number of codons for that particular amino acid the polynucleotide. This may be referred to a "codon frequency". For example, at least 50% of the total codons encoding a particular amino acid in the polynucleotide may be encoded by a first codon sequence. For example, at least 55% of the total codons encoding a particular amino acid in the polynucleotide may be encoded by a first codon sequence. At least 5%, 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more of the total codons encoding a particular amino in the polynucleotide may be encoded by a first codon sequence. In some cases, no more than 5%, 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or less of the total codons encoding a particular amino in the polynucleotide are encoded by a first codon sequence. At least about 90% phenylalanine-encoding codons of said synthetic polynucleotide may be TTC (as opposed to TTT). At least about 60% cysteine-encoding codons of said synthetic polynucleotide may be TGC (as opposed to TGT). At least about 70% aspartic acid-encoding codons of said synthetic polynucleotide may be GAC (as opposed to GAT). At least about 50% glutamic acid-encoding codons of said synthetic polynucleotide may be GAG (as opposed to GAA). At least about 60% histidine-encoding codons of said synthetic polynucleotide may be CAC (as opposed to CAT). At least about 60% lysine-encoding codons of said synthetic polynucleotide may be AAG (as opposed to AAA). At least about 60% asparagine-encoding codons of said synthetic polynucleotide may be AAC (as opposed to AAT). At least about 70% glutamine-encoding codons of said synthetic polynucleotide may be CAG (as opposed to CAA). At least about 80% tyrosine-encoding codons of said synthetic polynucleotide may be TAC (as opposed to TAT). At least about 90% isoleucine-encoding codons of said synthetic polynucleotide may be ATC.

[00107] In some embodiments, a particular amino acid the polynucleotide may be encoded by a number of different codon sequences. For example, a particular amino acid in the polynucleotide may be encoded by no more than 2 different codon sequences. In some cases, the polynucleotide comprises no more than 2 types of isoleucine-encoding codons.

[00108] In some embodiments, a particular amino acid in the polynucleotide may be encoded by no more than 3 different codon sequences. The polynucleotide may comprise no more than 3 types of alanine (Ala)-encoding codons. The polynucleotide may comprise no more than 3 types of glycine (Gly)-encoding codons.

The polynucleotide may comprise no more than 3 types of proline (Pro)-encoding codons. The polynucleotide may comprise no more than 3 types of threonine (Thr)-encoding codons.

[00109] In some embodiments, a particular amino acid in the polynucleotide may be encoded by no more than 4 different codon sequences. The polynucleotide may comprise no more than 4 types of arginine (Arg)-encoding codons. The polynucleotide may comprise no more than 4 types of serine (Ser)-encoding codons. In some embodiments, a particular amino acid in the polynucleotide may be encoded by no more than 5 different codon sequences. The polynucleotide may comprise no more than 5 types of arginine (Arg)-encoding codons. The polynucleotide may comprise no more than 5 types of serine (Ser)-encoding codons. In some embodiments, a particular amino acid in the polynucleotide may be encoded by no more than 6 different codon sequences. In some embodiments, a particular amino acid in the polynucleotide may be encoded by 1 or more different codon sequences. In some embodiments, a particular amino acid in the polynucleotide may be encoded by 2 or more different codon sequences. In some embodiments, a particular amino acid in the polynucleotide may be encoded by 3 or more different codon sequences. In some embodiments, a particular amino acid in the polynucleotide may be encoded by 4 or more different codon sequences. In some embodiments, a particular amino acid in the polynucleotide may be encoded by 5 or more different codon sequences. In some embodiments, a particular amino acid in the polynucleotide may be encoded by 6 or more different codon sequences.

[00110] In some cases, a frequency of a first codon sequence of a is higher, lower or the same as a frequency of a second codon sequence encoding for a particular amino acid in the polynucleotide. For example, a frequency of a first codon is higher than a frequency of second codon for a particular amino acid in the polynucleotide. The frequency of GCC codon may be higher than a frequency of GCT codon. The frequency of GCT codon may be lower than a frequency of GCA codon. The frequency of GCT codon may be higher than a frequency of GCA codon.

[00111] In some embodiments, the codon usage for alanine-encoding codons in the polynucleotide may have a particular parameter. For example, a frequency of GCG codon may be no more than about 10% or 5%. A frequency of GCA codon may be no more than about 20%. A frequency of GCT codon may be at least about 1%, 5%, 10%, 15%, 20%, or 25%. A frequency of GCT codon may be no more than about 30%, 25%, 20%, 15%, 10%, or 5%. A frequency of GCC codon may be at least about 60%, 70%, 80%, or 90%. A frequency of GCC codon is no more than about 95%, 90%, 85%, 80%, or 75%. The frequency of GCC codon may be higher than a frequency of GCT codon. The frequency of GCT codon may be lower than a frequency of GCA codon. The frequency of GCT codon may be higher than a frequency of GCA codon.

[00112] In some embodiments, the codon usage for glycine-encoding codons the polynucleotide may have a particular parameter. For example, a frequency of GGC codon may be lower than a frequency of GGA codon. For example, a frequency of GGC codon may be higher than a frequency of GGA codon. A frequency of GGG codon may be no more than about 10% or 5%. A frequency of GGG codon may be least about 1%. A frequency of GGA codon may be no more than about 30% or 20%. A frequency of GGA codon may be at least about 10% or 20%. A frequency of GGT codon may be more than about 10% or 5%.

A frequency of GGC codon may be no more than about 90%, 80%, or 70%. A frequency of GGC codon may be at least about 60%, 70%, or 80%.

[00113] In some embodiments, the codon usage for proline-encoding codons the polynucleotide may have a particular parameter. For example, a frequency of CCC codon may be lower than a frequency of CCT codon. A frequency of CCC codon may be higher than a frequency of CCT codon. A frequency of CCC codon may be lower than a frequency of CCA codon. A frequency of CCC codon may be higher than a frequency of CCA codon. A frequency of CCT codon may be lower than a frequency of CCA codon. A frequency of CCT codon may be higher than a frequency of CCA codon. A frequency of CCG codon may be no more than about 10% or 5%. frequency of CCA codon may be no more than about 30%, 20%, or 10%. A frequency of CCA codon may be at least about 5%, 10%, 15%, 20%, or 25%. A frequency of CCT codon may be no more than about 60%, 50%, 40%, or 30%. A frequency of CCT codon may be at least about 20%, 30%, 40%, or 50%. A frequency of CCC codon may be no more than about 60%, 50%, or 40%. A frequency of CCC codon may be at least about 30%, 40%, 50%, 60%, or 70%.

[00114] In some embodiments, the codon usage for threonine-encoding codons the polynucleotide may have a particular parameter. For example, a frequency of ACA codon is higher than a frequency of ACT codon. A frequency of ACC codon may be higher than a frequency of ACT codon. A frequency of ACC codon may be lower than a frequency of ACA codon. A frequency of ACC codon may be higher than a frequency of ACA codon. A frequency of ACG codon may be no more than about 10% or 5%. A frequency of ACA codon may be no more than about 60%, 50%, 40%, or 30%. A frequency of ACA codon may be at least about 10%, 20%, 30%, 40%, or 50%. A frequency of ACT codon may be no more than about 10% or 5%. A frequency of ACC codon may be no more than about 90%, 80%, 70%, 60%, or 50%. A frequency of ACC codon is at least about 40%, 50%, 60%, 70%, or 80%.

[00115] In some embodiments, the codon usage for arginine-encoding codons the polynucleotide may have a particular parameter. For example, a frequency of AGA codon may be lower than a frequency of AGG codon. A frequency of AGA codon may be higher than a frequency of AGG codon. A frequency of AGA codon may be lower than a frequency of CGG codon. A frequency of AGA codon may be higher than a frequency of CGG codon. A frequency of CGG codon may be higher than a frequency of CGA codon. A frequency of CGG codon is higher than a frequency of CGC codon. A frequency of AGG codon may be no more than about 10%. A frequency of AGG codon may be less than about 10%. A frequency of AGA codon may be no more than about 70%, 60%, or 50%. A frequency of AGA codon may be at least about 40%, 50%, 60%, or 70%. A frequency of CGG codon may be no more than about 50%, 40%, or 30%. A frequency of CGG codon may be at least about 20%, 30%, or 40%. A frequency of CGA codon may be at least about 1%. A frequency of CGA codon may be no more than about 10% or 5%. A frequency of CGT codon may be no more about 10% or 5%. A frequency of CGC codon may be no more than about 20%, 10%, or 5%. A frequency of CGC codon may be at least about 1%, 2%, 3%, 4%, or 5%.

[00116] In some embodiments, the codon usage for serine-encoding codons the polynucleotide may have a particular parameter. For example, a frequency of AGC codon may be higher than a frequency of TCT

codon. A frequency of TCT codon may be higher than a frequency of TCG codon. A frequency of TCT codon may be higher than a frequency of TCA codon. A frequency of TCT codon may be higher than a frequency of TCC codon. A frequency of AGT codon may be no more than about 10%. A frequency of AGT codon may be at least about 1%. A frequency of AGC codon may be no more about 95%, 90%, 85%, or 80%. A frequency of AGC codon may be at least about 70%, 80%, or 90%. A frequency of TCG codon may be no more than about 10% or 5%. A frequency of TCA codon may be no more than about 10% or 5%. A frequency of TCT codon may be no more than about 30%, 20%, or 10%. A frequency of TCT codon may be at least about 10%, or 20%. A frequency of TCC codon may be no more than about 10 % or 5%.

Example CFTR-encoding polynucleotides

[00117] In some embodiments of the synthetic polynucleotide, the synthetic polynucleotide is mRNA encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70% sequence identity over at least 100 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 75% sequence identity over at least 100 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity over at least 100 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**.

[00118] In some embodiments of the synthetic polynucleotide, the synthetic polynucleotide is mRNA encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70% sequence identity over at least 200 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 75% sequence identity over at least 200 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity over at least 200 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**.

[00119] In some embodiments of the synthetic polynucleotide, the synthetic polynucleotide is mRNA encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70% sequence identity with a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 75% sequence identity with a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity with a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) selected from **SEQ ID NOs: 1-4** and **23**.

[00120] In some embodiments of the synthetic polynucleotide, the synthetic polynucleotide is mRNA encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70% sequence identity over at least 100 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 75% sequence identity over at least 100 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity over at least 100 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**.

[00121] In some embodiments of the synthetic polynucleotide, the synthetic polynucleotide is mRNA encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70% sequence identity over at least 200 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 75% sequence identity over at least 200 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence)

having at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity over at least 200 bases of a sequence selected from **SEQ ID NOs: 1-4** and **23**.

[00122] In some embodiments of the synthetic polynucleotide, the synthetic polynucleotide is mRNA encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70% sequence identity with a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 75% sequence identity with a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity with a sequence selected from **SEQ ID NOs: 1-4** and **23**. In some embodiments, said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) selected from **SEQ ID NOs: 1-4** and **23**.

TABLE 1. Example CFTR ORF sequences

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
CFTR-001	ATGCAGAGAAGCCCTCTGGAAAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGCTGGAC CCGGCCCATCCTGCGGAAGGGCTACAGACAGAGACTGGAACTGAGCGACATCTATCAGA TCCCCAGCGTGGACAGCGCCGACAACCTGTCTGAGAAGCTGGAAAGAGAGTGGGACAGA GAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGCCCTGCGGCGGTGCTTCTTCTG GCGGTTCATGTTCTACGGCATCTTCCTGTACCTGGGCGAAGTGACCAAAGCCGTGCAGC CTCTGCTGCTGGGCAGAATCATCGCCAGCTACGACCCCGACAACAAAGAGGAACGGAGC ATCGCCATCTACCTCGGCATCGGCCTGTGCCTGCTGTTCATCGTCAGAACCCTGCTGCT GCACCCCGCCATCTTCGGACTGCACCACATCGGCATGCAGATGCGGATCGCCATGTTCA GCCTGATCTACAAGAAAACCCTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATC GGACAGCTGGTGAGCCTGCTGAGCAACAACCTGAACAAGTTCGACGAAGGCCTGGCTCT GGCCCACTTCGTGTGGATCGCTCCTCTGCAAGTGGCCCTGCTGATGGGCCTGATCTGGG AACTGCTGCAGGCCAGCGCCTTCTGCGGACTGGGATTCCTGATCGTGCTGGCCCTGTTC CAGGCCGGACTGGGGAGAATGATGATGAAGTACCGGGACCAGAGAGCCGGCAAGATCAG CGAGAGACTGGTCATCACCAGCGAGATGATCGAGAACATCCAGAGCGTGAAGGCCTACT GCTGGGAAGAGGCCATGGAAAAGATGATCGAGAATCTGCGGCAGACCGAGCTGAAGCTG ACAAGAAAGGCCGCCTACGTGCGCTACTTCAACAGCAGCGCCTTCTTCTTCAGCGGCTT CTTCGTGGTGTTCCTGAGCGTGCTGCCCTACGCTCTGATCAAGGGCATCATCCTGAGAA AGATCTTCACCACCATCAGCTTCTGCATCGTGCTGCGGATGGCCGTGACCAGACAGTTC CCCTGGGCTGTGCAGACCTGGTACGACAGCCTGGGCGCCATCAACAAGATCCAGGACTT CCTGCAGAAGCAAGAGTACAAGACCCTCGAGTACAACCTGACCACCACCGAGGTGGTCA TGGAAAACGTGACCGCCTTCTGGGAGGAAGGCTTCGGCGAGCTGTTCGAGAAGGCCAAG	**1**

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	CAGAACAACAACAACCGCAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTT CAGCCTGCTGGGGACCCCTGTGCTGAAGGACATCAACTTCAAGATCGAGCGGGGACAGC TGCTGGCCGTGGCTGGAAGCACAGGCGCCGGAAAAACCAGCCTGCTCATGGTCATCATG GGCGAGCTGGAACCCAGCGAGGGCAAGATCAAGCACAGCGGCAGGATCAGCTTCTGCAG CCAGTTCAGCTGGATCATGCCCGGCACCATCAAAGAGAACATCATCTTCGGCGTGAGCT ACGACGAGTACAGATACCGCAGCGTGATCAAGGCCTGCCAGCTGGAAGAGGACATCAGC AAGTTCGCCGAGAAGGACAACATCGTGCTCGGCGAAGGCGGCATCACACTGTCTGGCGG ACAGAGGGCCAGAATCTCTCTGGCCAGAGCCGTGTACAAGGACGCCGATCTGTACCTGC TGGACAGCCCCTTCGGCTACCTGGATGTGCTGACCGAGAAAGAGATCTTCGAGAGCTGC GTGTGCAAGCTGATGGCCAACAAGACCCGGATCCTGGTCACCAGCAAGATGGAACACCT GAAGAAGGCCGACAAGATCCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCT TCAGCGAGCTGCAGAACCTGCAGCCTGACTTCAGCAGCAAACTGATGGGCTGCGACAGC TTCGACCAGTTCAGCGCCGAGCGGAGAAACAGCATCCTGACAGAGACACTGCACCGGTT CAGCCTGGAAGGCGACGCTCCTGTGAGCTGGACCGAGACAAAGAAGCAGAGCTTCAAGC AGACCGGCGAGTTCGGCGAGAAGCGGAAGAACAGCATCCTGAACCCCATCAACAGCATC CGGAAGTTCAGCATCGTCCAGAAAACCCCTCTGCAGATGAACGGCATCGAAGAGGACAG CGACGAGCCCCTGGAAAGACGGCTGTCTCTGGTGCCTGACAGCGAACAGGGCGAAGCCA TCCTGCCTCGGATCAGCGTGATCAGCACAGGCCCCACACTGCAGGCTCGGAGAAGGCAG AGTGTGCTGAACCTGATGACCCACAGCGTGAACCAGGGACAGAACATCCACAGAAAGAC CACCGCCAGCACACGGAAAGTGAGCCTGGCCCCTCAGGCCAACCTGACTGAGCTGGACA TCTACAGCAGACGGCTGAGCCAAGAGACAGGCCTGGAAATCAGCGAGGAAATCAACGAA GAGGACCTGAAAGAGTGCTTCTTCGACGACATGGAAAGCATCCCCGCCGTGACAACCTG GAACACCTACCTGCGGTACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGGT GTCTCGTGATCTTCCTGGCCGAAGTGGCCGCCTCTCTGGTGGTGCTGTGGCTGCTCGGA AACACCCCACTGCAGGACAAGGGCAACAGCACCCACAGCCGGAACAACAGCTACGCCGT GATCATCACCAGCACCAGCAGCTACTACGTGTTCTACATCTACGTGGGCGTCGCCGACA CTCTGCTCGCCATGGGCTTCTTCAGAGGACTGCCCCTGGTGCACACCCTGATCACCGTG AGCAAGATCCTGCACCACAAGATGCTGCACAGCGTCCTGCAGGCCCCCATGAGCACACT GAACACCCTGAAAGCCGGCGGAATCCTGAACAGATTCAGCAAGGACATCGCCATCCTGG ACGACCTGCTGCCTCTGACCATCTTCGACTTCATCCAGCTGCTGCTGATCGTGATCGGC GCCATCGCTGTGGTGGCTGTGCTGCAGCCCTACATCTTCGTGGCCACCGTGCCTGTGAT CGTGGCCTTCATCATGCTGCGGGCCTACTTCCTGCAGACCTCTCAGCAGCTGAAGCAGC TCGAGTCTGAGGGCAGAAGCCCCATCTTCACCCACCTCGTGACCAGCCTGAAAGGCCTG TGGACCCTGAGAGCCTTCGGCAGACAGCCCTACTTCGAGACACTGTTCCACAAGGCCCT GAACCTGCACACCGCCAACTGGTTCCTGTATCTGAGCACCCTGCGGTGGTTCCAGATGA GGATCGAGATGATCTTCGTCATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTCACC ACTGGCGAAGGCGAGGGCAGAGTGGGAATCATCCTGACCCTGGCCATGAACATCATGAG CACACTCCAGTGGGCCGTGAACAGCAGCATCGATGTGGACAGCCTGATGCGGAGCGTGA GCCGGGTGTTCAAGTTCATCGACATGCCCACAGAGGGCAAGCCCACCAAGAGCACCAAG CCCTACAAGAACGGCCAGCTGAGCAAAGTCATGATCATCGAGAACAGCCACGTCAAGAA GGACGACATCTGGCCCAGCGGAGGCCAGATGACCGTGAAGGATCTGACCGCCAAGTACA CCGAAGGCGGAAACGCCATCCTGGAAAACATCAGCTTCAGCATCAGCCCTGGCCAGCGC GTGGGACTCCTGGGAAGAACCGGAAGCGGCAAGAGCACTCTGCTGAGCGCCTTCCTGAG ACTGCTGAACACCGAGGGCGAGATCCAGATCGATGGGGTGAGCTGGGACAGCATCACCC TGCAACAATGGCGGAAGGCCTTCGGCGTGATCCCTCAGAAGGTGTTCATCTTCAGCGGC ACGTTCCGGAAGAATCTGGACCCCTACGAGCAGTGGAGCGACCAAGAGATCTGGAAGGT GGCCGATGAAGTGGGACTGAGAAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGACTTCG TGCTGGTGGATGGCGGCTGTGTGCTGTCTCACGGACACAAGCAGCTGATGTGCCTGGCC AGAAGCGTGCTGAGCAAGGCCAAGATCCTGCTGCTCGACGAGCCCAGCGCTCACCTGGA TCCTGTGACCTACCAGATCATCCGGCGGACACTGAAGCAGGCCTTCGCCGACTGCACCG TGATCCTGTGCGAGCACAGAATCGAGGCCATGCTGGAATGCCAGCAGTTCCTGGTGATC GAAGAGAACAAAGTGCGGCAGTACGACAGCATCCAGAAGCTGCTGAACGAGCGGAGCCT GTTCAGACAGGCCATCTCTCCCAGCGACAGAGTGAAGCTGTTCCCTCACCGGAACAGCA GCAAGTGCAAGAGCAAGCCTCAGATCGCCGCTCTGAAAGAAGAAACCGAGGAAGAGGTG CAGGACACACGGCTGGCGGCCGTTTACCCATACGATGTTCCTGACTATGCGTGA	
CFTR-003	ATGCAGAGAAGCCCTCTGGAAAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGCTGGAC CCGGCCCATCCTGCGGAAGGGCTACAGACAGAGACTGGAACTGAGCGACATCTATCAGA TCCCCAGCGTGGACAGCGCCGACAACCTGTCTGAGAAGCTGGAAAGAGAGTGGGACAGA GAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGCCCTGCGGCGGTGCTTCTTCTG GCGGTTCATGTTCTACGGCATCTTCCTGTACCTGGGCGAAGTGACCAAAGCCGTGCAGC	**2**

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	CTCTGCTGCTGGGCAGAATCATCGCCAGCTACGACCCCGACAACAAAGAGGAACGGAGC ATCGCCATCTACCTCGGCATCGGCCTGTGCCTGCTGTTCATCGTCAGAACCCTGCTGCT GCACCCCGCCATCTTCGGACTGCACCACATCGGCATGCAGATGCGGATCGCCATGTTCA GCCTGATCTACAAGAAAACCCTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATC GGACAGCTGGTGAGCCTGCTGAGCAACAACCTGAACAAGTTCGACGAAGGCCTGGCTCT GGCCCACTTCGTGTGGATCGCTCCTCTGCAAGTGGCCCTGCTGATGGGCCTGATCTGGG AACTGCTGCAGGCCAGCGCCTTCTGCGGACTGGGATTCCTGATCGTGCTGGCCCTGTTC CAGGCCGGACTGGGGAGAATGATGATGAAGTACCGGGACCAGAGAGCCGGCAAGATCAG CGAGAGACTGGTCATCACCAGCGAGATGATCGAGAACATCCAGAGCGTGAAGGCCTACT GCTGGGAAGAGGCCATGGAAAAGATGATCGAGAATCTGCGGCAGACCGAGCTGAAGCTG ACAAGAAAGGCCGCCTACGTGCGCTACTTCAACAGCAGCGCCTTCTTCTTCAGCGGCTT CTTCGTGGTGTTCCTGAGCGTGCTGCCCTACGCTCTGATCAAGGGCATCATCCTGAGAA AGATCTTCACCACCATCAGCTTCTGCATCGTGCTGCGGATGGCCGTGACCAGACAGTTC CCCTGGGCTGTGCAGACCTGGTACGACAGCCTGGGCGCCATCAACAAGATCCAGGACTT CCTGCAGAAGCAAGAGTACAAGACCCTCGAGTACAACCTGACCACCACCGAGGTGGTCA TGGAAAACGTGACCGCCTTCTGGGAGGAAGGCTTCGGCGAGCTGTTCGAGAAGGCCAAG CAGAACAACAACAACCGCAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTT CAGCCTGCTGGGGACCCCTGTGCTGAAGGACATCAACTTCAAGATCGAGCGGGGACAGC TGCTGGCCGTGGCTGGAAGCACAGGCGCCGGAAAAACCAGCCTGCTCATG**GTC**ATCATG GGCGAGCTGGAACCCAGCGAGGGCAAGATCAAGCACAGCGGCAGGATCAGCTTCTGCAG CCAGTTCAGCTGGATCATGCCCGGCACCATCAAAGAGAACATCATCTTCGGCGTGAGCT ACGACGAGTACAGATACCGCAGCGTGATCAAGGCCTGCCAGCTGGAAGAGGACATCAGC AAGTTCGCCGAGAAGGACAACATCGTGCTCGGCGAAGGCGGCATCACACTGTCTGGCGG ACAGAGGGCCAGAATCTCTCTGGCCAGAGCCGTGTACAAGGACGCCGATCTGTACCTGC TGGACAGCCCCTTCGGCTACCTGGATGTGCTGACCGAGAAAGAGATCTTCGAGAGCTGC GTGTGCAAGCTGATGGCCAACAAGACCCGGATCCTGGTCACCAGCAAGATGGAACACCT GAAGAAGGCCGACAAGATCCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCT TCAGCGAGCTGCAGAACCTGCAGCCTGACTTCAGCAGCAAACTGATGGGCTGCGACAGC TTCGACCAGTTCAGCGCCGAGCGGAGAAACAGCATCCTGACAGAGACACTGCACCGGTT CAGCCTGGAAGGCGACGCTCCTGTGAGCTGGACCGAGACAAAGAAGCAGAGCTTCAAGC AGACCGGCGAGTTCGGCGAGAAGCGGAAGAACAGCATCCTGAACCCCATCAACAGCATC CGGAAGTTCAGCATCGTCCAGAAAACCCCTCTGCAGATGAACGGCATCGAAGAGGACAG CGACGAGCCCCTGGAAAGACGGCTGTCTCTGGTGCCTGACAGCGAACAGGGCGAAGCCA TCCTGCCTCGGATCAGCGTGATCAGCACAGGCCCCACACTGCAGGCTCGGAGAAGGCAG AGTGTGCTGAACCTGATGACCCACAGCGTGAACCAGGGACAGAACATCCACAGAAAGAC CACCGCCAGCACACGGAAAGTGAGCCTGGCCCCTCAGGCCAACCTGACTGAGCTGGACA TCTACAGCAGACGGCTGAGCCAAGAGACAGGCCTGGAAATCAGCGAGGAAATCAACGAA GAGGACCTGAAAGAGTGCTTCTTCGACGACATGGAAAGCATCCCCGCCGTGACAACCTG GAACACCTACCTGCGGTACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGGT GTCTCGTGATCTTCCTGGCCGAAGTGGCCGCCTCTCTGGTGGTGCTGTGGCTGCTCGGA AACACCCCACTGCAGGACAAGGGCAACAGCACCCACAGCCGGAACAACAGCTACGCCGT GATCATCACCAGCACCAGCAGCTACTACGTGTTCTACATCTACGTGGGCGTCGCCGACA CTCTGCTCGCCATGGGCTTCTTCAGAGGACTGCCCCTGGTGCACACCCTGATCACCGTG AGCAAGATCCTGCACCACAAGATGCTGCACAGCGTCCTGCAGGCCCCCATGAGCACACT GAACACCCTGAAAGCCGGCGGAATCCTGAACAGATTCAGCAAGGACATCGCCATCCTGG ACGACCTGCTGCCTCTGACCATCTTCGACTTCATCCAGCTGCTGCTGATCGTGATCGGC GCCATCGCTGTGGTGGCTGTGCTGCAGCCCTACATCTTCGTGGCCACCGTGCCTGTGAT CGTGGCCTTCATCATGCTGCGGGCCTACTTCCTGCAGACCTCTCAGCAGCTGAAGCAGC TCGAGTCTGAGGGCAGAAGCCCCATCTTCACCCACCTCGTGACCAGCCTGAAAGGCCTG TGGACCCTGAGAGCCTTCGGCAGACAGCCCTACTTCGAGACACTGTTCCACAAGGCCCT GAACCTGCACACCGCCAACTGGTTCCTGTATCTGAGCACCCTGCGGTGGTTCCAGATGA GGATCGAGATGATCTTCGTCATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTCACC ACTGGCGAAGGCGAGGGCAGAGTGGGAATCATCCTGACCCTGGCCATGAACATCATGAG CACACTCCAGTGGGCCGTGAACAGCAGCATCGATGTGGACAGCCTGATGCGGAGCGTGA GCCGGGTGTTCAAGTTCATCGACATGCCCACAGAGGGCAAGCCCACCAAGAGCACCAAG CCCTACAAGAACGGCCAGCTGAGCAAAGTCATGATCATCGAGAACAGCCACGTCAAGAA GGACGACATCTGGCCCAGCGGAGGCCAGATGACCGTGAAGGATCTGACCGCCAAGTACA CCGAAGGCGGAAACGCCATCCTGGAAAACATCAGCTTCAGCATCAGCCCTGGCCAGCGC GTGGGACTCCTGGGAAGAACCGGAAGCGGCAAGAGCACTCTGCTGAGCGCCTTCCTGAG ACTGCTGAACACCGAGGGCGAGATCCAGATCGATGGGGTGAGCTGGGACAGCATCACCC	

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	TGCAACAATGGCGGAAGGCCTTCGGCGTGATCCCTCAGAAGGTGTTCATCTTCAGCGGC ACGTTCCGGAAGAATCTGGACCCCTACGAGCAGTGGAGCGACCAAGAGATCTGGAAGGT GGCCGATGAAGTGGGACTGAGAAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGACTTCG TGCTGGTGGATGGCGGCTGTGTGCTGTCTCACGGACACAAGCAGCTGATGTGCCTGGCC AGAAGCGTGCTGAGCAAGGCCAAGATCCTGCTGCTCGACGAGCCCAGCGCTCACCTGGA TCCTGTGACCTACCAGATCATCCGGCGGACACTGAAGCAGGCCTTCGCCGACTGCACCG TGATCCTGTGCGAGCACAGAATCGAGGCCATGCTGGAATGCCAGCAGTTCCTGGTGATC GAAGAGAACAAAGTGCGGCAGTACGACAGCATCCAGAAGCTGCTGAACGAGCGGAGCCT GTTCAGACAGGCCATCTCTCCCAGCGACAGAGTGAAGCTGTTCCCTCACCGGAACAGCA GCAAGTGCAAGAGCAAGCCTCAGATCGCCGCTCTGAAAGAAGAAACCGAGGAAGAGGTG CAGGACACACGGCTGTGA	
CFTR-004	ATGCAGAGAAGCCCCCTGGAAAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGCTGGAC CCGGCCCATCCTGCGGAAGGGCTACAGACAGAGACTGGAACTGAGCGACATCTACCAGA TCCCCAGCGTGGACAGCGCCGACAACCTGAGCGAGAAGCTGGAAAGAGAGTGGGACAGA GAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGCCCTGCGGCGGTGCTTCTTCTG GCGGTTCATGTTCTACGGCATCTTCCTGTACCTGGGCGAAGTGACCAAAGCCGTGCAGC CCCTGCTGCTGGGCAGAATCATCGCCAGCTACGACCCCGACAACAAAGAGGAACGGAGC ATCGCCATCTACCTCGGCATCGGCCTGTGCCTGCTGTTCATCGTCAGAACCCTGCTGCT GCACCCCGCCATCTTCGGACTGCACCACATCGGCATGCAGATGCGGATCGCCATGTTCA GCCTGATCTACAAGAAAACCCTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATC GGACAGCTGGTGAGCCTGCTGAGCAACAACCTGAACAAGTTCGACGAAGGCCTGGCCCT GGCCCACTTCGTGTGGATCGCCCCCCTGCAAGTGGCCCTGCTGATGGGCCTGATCTGGG AACTGCTGCAGGCCAGCGCCTTCTGCGGACTGGGATTCCTGATCGTGCTGGCCCTGTTC CAGGCCGGACTGGGGAGAATGATGATGAAGTACCGGGACCAGAGAGCCGGCAAGATCAG CGAGAGACTGGTCATCACCAGCGAGATGATCGAGAACATCCAGAGCGTGAAGGCCTACT GCTGGGAAGAGGCCATGGAAAAGATGATCGAGAACCTGCGGCAGACCGAGCTGAAGCTG ACAAGAAAGGCCGCCTACGTGCGCTACTTCAACAGCAGCGCCTTCTTCTTCAGCGGCTT CTTCGTGGTGTTCCTGAGCGTGCTGCCCTACGCCCTGATCAAGGGCATCATCCTGAGAA AGATCTTCACCACCATCAGCTTCTGCATCGTGCTGCGGATGGCCGTGACCAGACAGTTC CCCTGGGCCGTGCAGACCTGGTACGACAGCCTGGGCGCCATCAACAAGATCCAGGACTT CCTGCAGAAGCAAGAGTACAAGACCCTCGAGTACAACCTGACCACCACCGAGGTGGTCA TGGAAAACGTGACCGCCTTCTGGGAGGAAGGCTTCGGCGAGCTGTTCGAGAAGGCCAAG CAGAACAACAACAACCGCAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTT CAGCCTGCTGGGGACCCCCGTGCTGAAGGACATCAACTTCAAGATCGAGCGGGGACAGC TGCTGGCCGTGGCCGGAAGCACAGGCGCCGGAAAAACCAGCCTGCTCATGGTCATCATG GGCGAGCTGGAACCCAGCGAGGGCAAGATCAAGCACAGCGGCAGGATCAGCTTCTGCAG CCAGTTCAGCTGGATCATGCCCGGCACCATCAAAGAGAACATCATCTTCGGCGTGAGCT ACGACGAGTACAGATACCGCAGCGTGATCAAGGCCTGCCAGCTGGAAGAGGACATCAGC AAGTTCGCCGAGAAGGACAACATCGTGCTCGGCGAAGGCGGCATCACACTGAGCGGCGG ACAGAGGGCCAGAATCAGCCTGGCCAGAGCCGTGTACAAGGACGCCGACCTGTACCTGC TGGACAGCCCCTTCGGCTACCTGGACGTGCTGACCGAGAAAGAGATCTTCGAGAGCTGC GTGTGCAAGCTGATGGCCAACAAGACCCGGATCCTGGTCACCAGCAAGATGGAACACCT GAAGAAGGCCGACAAGATCCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCT TCAGCGAGCTGCAGAACCTGCAGCCCGACTTCAGCAGCAAACTGATGGGCTGCGACAGC TTCGACCAGTTCAGCGCCGAGCGGAGAAACAGCATCCTGACAGAGACACTGCACCGGTT CAGCCTGGAAGGCGACGCCCCCGTGAGCTGGACCGAGACAAAGAAGCAGAGCTTCAAGC AGACCGGCGAGTTCGGCGAGAAGCGGAAGAACAGCATCCTGAACCCCATCAACAGCATC CGGAAGTTCAGCATCGTCCAGAAAACCCCCCTGCAGATGAACGGCATCGAAGAGGACAG CGACGAGCCCCTGGAAAGACGGCTGAGCCTGGTGCCCGACAGCGAACAGGGCGAAGCCA TCCTGCCCCGGATCAGCGTGATCAGCACAGGCCCCACACTGCAGGCCCGGAGAAGGCAG AGCGTGCTGAACCTGATGACCCACAGCGTGAACCAGGGACAGAACATCCACAGAAAGAC CACCGCCAGCACACGGAAAGTGAGCCTGGCCCCCCAGGCCAACCTGACTGAGCTGGACA TCTACAGCAGACGGCTGAGCCAAGAGACAGGCCTGGAAATCAGCGAGGAAATCAACGAA GAGGACCTGAAAGAGTGCTTCTTCGACGACATGGAAAGCATCCCCGCCGTGACAACCTG GAACACCTACCTGCGGTACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGGT GCCTCGTGATCTTCCTGGCCGAAGTGGCCGCCAGCCTGGTGGTGCTGTGGCTGCTCGGA AACACCCCACTGCAGGACAAGGGCAACAGCACCCACAGCCGGAACAACAGCTACGCCGT GATCATCACCAGCACCAGCAGCTACTACGTGTTCTACATCTACGTGGGCGTCGCCGACA CTCTGCTCGCCATGGGCTTCTTCAGAGGACTGCCCCTGGTGCACACCCTGATCACCGTG	**3**

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	AGCAAGATCCTGCACCACAAGATGCTGCACAGCGTCCTGCAGGCCCCCATGAGCACACT GAACACCCTGAAAGCCGGCGGAATCCTGAACAGATTCAGCAAGGACATCGCCATCCTGG ACGACCTGCTGCCCCTGACCATCTTCGACTTCATCCAGCTGCTGCTGATCGTGATCGGC GCCATCGCCGTGGTGGCCGTGCTGCAGCCCTACATCTTCGTGGCCACCGTGCCCGTGAT CGTGGCCTTCATCATGCTGCGGGCCTACTTCCTGCAGACCAGCCAGCAGCTGAAGCAGC TCGAGAGCGAGGGCAGAAGCCCCATCTTCACCCACCTCGTGACCAGCCTGAAAGGCCTG TGGACCCTGAGAGCCTTCGGCAGACAGCCCTACTTCGAGACACTGTTCCACAAGGCCCT GAACCTGCACACCGCCAACTGGTTCCTGTACCTGAGCACCCTGCGGTGGTTCCAGATGA GGATCGAGATGATCTTCGTCATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTCACC ACTGGCGAAGGCGAGGGCAGAGTGGGAATCATCCTGACCCTGGCCATGAACATCATGAG CACACTCCAGTGGGCCGTGAACAGCAGCATCGACGTGGACAGCCTGATGCGGAGCGTGA GCCGGGTGTTCAAGTTCATCGACATGCCCACAGAGGGCAAGCCCACCAAGAGCACCAAG CCCTACAAGAACGGCCAGCTGAGCAAAGTCATGATCATCGAGAACAGCCACGTCAAGAA GGACGACATCTGGCCCAGCGGAGGCCAGATGACCGTGAAGGACCTGACCGCCAAGTACA CCGAAGGCGGAAACGCCATCCTGGAAAACATCAGCTTCAGCATCAGCCCCGGCCAGCGC GTGGGACTCCTGGGAAGAACCGGAAGCGGCAAGAGCACTCTGCTGAGCGCCTTCCTGAG ACTGCTGAACACCGAGGGCGAGATCCAGATCGACGGGGTGAGCTGGGACAGCATCACCC TGCAACAATGGCGGAAGGCCTTCGGCGTGATCCCCCAGAAGGTGTTCATCTTCAGCGGC ACGTTCCGGAAGAACCTGGACCCCTACGAGCAGTGGAGCGACCAAGAGATCTGGAAGGT GGCCGACGAAGTGGGACTGAGAAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGACTTCG TGCTGGTGGACGGCGGCTGCGTGCTGAGCCACGGACACAAGCAGCTGATGTGCCTGGCC AGAAGCGTGCTGAGCAAGGCCAAGATCCTGCTGCTCGACGAGCCCAGCGCCCACCTGGA CCCCGTGACCTACCAGATCATCCGGCGGACACTGAAGCAGGCCTTCGCCGACTGCACCG TGATCCTGTGCGAGCACAGAATCGAGGCCATGCTGGAATGCCAGCAGTTCCTGGTGATC GAAGAGAACAAAGTGCGGCAGTACGACAGCATCCAGAAGCTGCTGAACGAGCGGAGCCT GTTCAGACAGGCCATCAGCCCCAGCGACAGAGTGAAGCTGTTCCCCCACCGGAACAGCA GCAAGTGCAAGAGCAAGCCCCAGATCGCCGCCCTGAAAGAAGAAACCGAGGAAGAGGTG CAGGACACACGGCTGGCGGCCGTTTACCCATACGATGTTCCTGACTATGCGTGA	
CFTR-005	ATGCAGAGAAGCCCCCTGGAAAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGCTGGAC CCGGCCCATCCTGCGGAAGGGCTACAGACAGAGACTGGAACTGAGCGACATCTACCAGA TCCCCAGCGTGGACAGCGCCGACAACCTGAGCGAGAAGCTGGAAAGAGAGTGGGACAGA GAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGCCCTGCGGCGGTGCTTCTTCTG GCGGTTCATGTTCTACGGCATCTTCCTGTACCTGGGCGAAGTGACCAAAGCCGTGCAGC CCCTGCTGCTGGGCAGAATCATCGCCAGCTACGACCCCGACAACAAAGAGGAACGGAGC ATCGCCATCTACCTCGGCATCGGCCTGTGCCTGCTGTTCATCGTCAGAACCCTGCTGCT GCACCCCGCCATCTTCGGACTGCACCACATCGGCATGCAGATGCGGATCGCCATGTTCA GCCTGATCTACAAGAAAACCCTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATC GGACAGCTGGTGAGCCTGCTGAGCAACAACCTGAACAAGTTCGACGAAGGCCTGGCCCT GGCCCACTTCGTGTGGATCGCCCCCCTGCAAGTGGCCCTGCTGATGGGCCTGATCTGGG AACTGCTGCAGGCCAGCGCCTTCTGCGGACTGGGATTCCTGATCGTGCTGGCCCTGTTC CAGGCCGGACTGGGGAGAATGATGATGAAGTACCGGGACCAGAGAGCCGGCAAGATCAG CGAGAGACTGGTCATCACCAGCGAGATGATCGAGAACATCCAGAGCGTGAAGGCCTACT GCTGGGAAGAGGCCATGGAAAAGATGATCGAGAACCTGCGGCAGACCGAGCTGAAGCTG ACAAGAAAGGCCGCCTACGTGCGCTACTTCAACAGCAGCGCCTTCTTCTTCAGCGGCTT CTTCGTGGTGTTCCTGAGCGTGCTGCCCTACGCCCTGATCAAGGGCATCATCCTGAGAA AGATCTTCACCACCATCAGCTTCTGCATCGTGCTGCGGATGGCCGTGACCAGACAGTTC CCCTGGGCCGTGCAGACCTGGTACGACAGCCTGGGCGCCATCAACAAGATCCAGGACTT CCTGCAGAAGCAAGAGTACAAGACCCTCGAGTACAACCTGACCACCACCGAGGTGGTCA TGGAAAACGTGACCGCCTTCTGGGAGGAAGGCTTCGGCGAGCTGTTCGAGAAGGCCAAG CAGAACAACAACAACCGCAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTT CAGCCTGCTGGGGACCCCCGTGCTGAAGGACATCAACTTCAAGATCGAGCGGGGACAGC TGCTGGCCGTGGCCGGAAGCACAGGCGCCGGAAAAACCAGCCTGCTCATGGTCATCATG GGCGAGCTGGAACCCAGCGAGGGCAAGATCAAGCACAGCGGCAGGATCAGCTTCTGCAG CCAGTTCAGCTGGATCATGCCCGGCACCATCAAAGAGAACATCATCTTCGGCGTGAGCT ACGACGAGTACAGATACCGCAGCGTGATCAAGGCCTGCCAGCTGGAAGAGGACATCAGC AAGTTCGCCGAGAAGGACAACATCGTGCTCGGCGAAGGCGGCATCACACTGAGCGGCGG ACAGAGGGCCAGAATCAGCCTGGCCAGAGCCGTGTACAAGGACGCCGACCTGTACCTGC TGGACAGCCCCTTCGGCTACCTGGACGTGCTGACCGAGAAAGAGATCTTCGAGAGCTGC GTGTGCAAGCTGATGGCCAACAAGACCCGGATCCTGGTCACCAGCAAGATGGAACACCT	**4**

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	GAAGAAGGCCGACAAGATCCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCT TCAGCGAGCTGCAGAACCTGCAGCCCGACTTCAGCAGCAAACTGATGGGCTGCGACAGC TTCGACCAGTTCAGCGCCGAGCGGAGAAACAGCATCCTGACAGAGACACTGCACCGGTT CAGCCTGGAAGGCGACGCCCCCGTGAGCTGGACCGAGACAAAGAAGCAGAGCTTCAAGC AGACCGGCGAGTTCGGCGAGAAGCGGAAGAACAGCATCCTGAACCCCATCAACAGCATC CGGAAGTTCAGCATCGTCCAGAAAACCCCCCTGCAGATGAACGGCATCGAAGAGGACAG CGACGAGCCCCTGGAAAGACGGCTGAGCCTGGTGCCCGACAGCGAACAGGGCGAAGCCA TCCTGCCCCGGATCAGCGTGATCAGCACAGGCCCCACACTGCAGGCCCGGAGAAGGCAG AGCGTGCTGAACCTGATGACCCACAGCGTGAACCAGGGACAGAACATCCACAGAAAGAC CACCGCCAGCACACGGAAAGTGAGCCTGGCCCCCCAGGCCAACCTGACTGAGCTGGACA TCTACAGCAGACGGCTGAGCCAAGAGACAGGCCTGGAAATCAGCGAGGAAATCAACGAA GAGGACCTGAAAGAGTGCTTCTTCGACGACATGGAAAGCATCCCCGCCGTGACAACCTG GAACACCTACCTGCGGTACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGGT GCCTCGTGATCTTCCTGGCCGAAGTGGCCGCCAGCCTGGTGGTGCTGTGGCTGCTCGGA AACACCCCACTGCAGGACAAGGGCAACAGCACCCACAGCCGGAACAACAGCTACGCCGT GATCATCACCAGCACCAGCAGCTACTACGTGTTCTACATCTACGTGGGCGTCGCCGACA CTCTGCTCGCCATGGGCTTCTTCAGAGGACTGCCCCTGGTGCACACCCTGATCACCGTG AGCAAGATCCTGCACCACAAGATGCTGCACAGCGTCCTGCAGGCCCCCATGAGCACACT GAACACCCTGAAAGCCGGCGGAATCCTGAACAGATTCAGCAAGGACATCGCCATCCTGG ACGACCTGCTGCCCCTGACCATCTTCGACTTCATCCAGCTGCTGCTGATCGTGATCGGC GCCATCGCCGTGGTGGCCGTGCTGCAGCCCTACATCTTCGTGGCCACCGTGCCCGTGAT CGTGGCCTTCATCATGCTGCGGGCCTACTTCCTGCAGACCAGCCAGCAGCTGAAGCAGC TCGAGAGCGAGGGCAGAAGCCCCATCTTCACCCACCTCGTGACCAGCCTGAAAGGCCTG TGGACCCTGAGAGCCTTCGGCAGACAGCCCTACTTCGAGACACTGTTCCACAAGGCCCT GAACCTGCACACCGCCAACTGGTTCCTGTACCTGAGCACCCTGCGGTGGTTCCAGATGA GGATCGAGATGATCTTCGTCATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTCACC ACTGGCGAAGGCGAGGGCAGAGTGGGAATCATCCTGACCCTGGCCATGAACATCATGAG CACACTCCAGTGGGCCGTGAACAGCAGCATCGACGTGGACAGCCTGATGCGGAGCGTGA GCCGGGTGTTCAAGTTCATCGACATGCCCACAGAGGGCAAGCCCACCAAGAGCACCAAG CCCTACAAGAACGGCCAGCTGAGCAAAGTCATGATCATCGAGAACAGCCACGTCAAGAA GGACGACATCTGGCCCAGCGGAGGCCAGATGACCGTGAAGGACCTGACCGCCAAGTACA CCGAAGGCGGAAACGCCATCCTGGAAAACATCAGCTTCAGCATCAGCCCCGGCCAGCGC GTGGGACTCCTGGGAAGAACCGGAAGCGGCAAGAGCACTCTGCTGAGCGCCTTCCTGAG ACTGCTGAACACCGAGGGCGAGATCCAGATCGACGGGGTGAGCTGGGACAGCATCACCC TGCAACAATGGCGGAAGGCCTTCGGCGTGATCCCCCAGAAGGTGTTCATCTTCAGCGGC ACGTTCCGGAAGAACCTGGACCCCTACGAGCAGTGGAGCGACCAAGAGATCTGGAAGGT GGCCGACGAAGTGGGACTGAGAAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGACTTCG TGCTGGTGGACGGCGGCTGCGTGCTGAGCCACGGACACAAGCAGCTGATGTGCCTGGCC AGAAGCGTGCTGAGCAAGGCCAAGATCCTGCTGCTCGACGAGCCCAGCGCCCACCTGGA CCCCGTGACCTACCAGATCATCCGGCGGACACTGAAGCAGGCCTTCGCCGACTGCACCG TGATCCTGTGCGAGCACAGAATCGAGGCCATGCTGGAATGCCAGCAGTTCCTGGTGATC GAAGAGAACAAAGTGCGGCAGTACGACAGCATCCAGAAGCTGCTGAACGAGCGGAGCCT GTTCAGACAGGCCATCAGCCCCAGCGACAGAGTGAAGCTGTTCCCCCACCGGAACAGCA GCAAGTGCAAGAGCAAGCCCCAGATCGCCGCCCTGAAAGAAGAAACCGAGGAAGAGGTG CAGGACACACGGCTGTGA	
Wild Type CFTR	ATGCAGAGGTCGCCTCTGGAAAAGGCCAGCGTTGTCTCCAAACTTTTTTTCAGCTGGAC CAGACCAATTTTGAGGAAAGGATACAGACAGCGCCTGGAATTGTCAGACATATACCAAA TCCCTTCTGTTGATTCTGCTGACAATCTATCTGAAAAATTGGAAAGAGAATGGGATAGA GAGCTGGCTTCAAAGAAAAATCCTAAACTCATTAATGCCCTTCGGCGATGTTTTTTCTG GAGATTTATGTTCTATGGAATCTTTTTATATTTAGGGGAAGTCACCAAAGCAGTACAGC CTCTCTTACTGGGAAGAATCATAGCTTCCTATGACCCGGATAACAAGGAGGAACGCTCT ATCGCGATTTATCTAGGCATAGGCTTATGCCTTCTCTTTATTGTGAGGACACTGCTCCT ACACCCAGCCATTTTTGGCCTTCATCACATTGGAATGCAGATGAGAATAGCTATGTTTA GTTTGATTTATAAGAAGACTTTAAAGCTGTCAAGCCGTGTTCTAGATAAAATAAGTATT GGACAACTTGTTAGTCTCCTTTCCAACAACCTGAACAAATTTGATGAAGGACTTGCATT GGCACATTTCGTGTGGATCGCTCCTTTGCAAGTGGCACTCCTCATGGGGCTAATCTGGG AGTTGTTACAGGCGTCTGCCTTCTGTGGACTTGGTTTCCTGATAGTCCTTGCCCTTTTT CAGGCTGGGCTAGGGAGAATGATGATGAAGTACAGAGATCAGAGAGCTGGGAAGATCAG TGAAAGACTTGTGATTACCTCAGAAATGATTGAAAATATCCAATCTGTTAAGGCATACT GCTGGGAAGAAGCAATGGAAAAAATGATTGAAAACTTAAGACAAACAGAACTGAAACTG	**5**

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	ACTCGGAAGGCAGCCTATGTGAGATACTTCAATAGCTCAGCCTTCTTCTTCTCAGGGTT CTTTGTGGTGTTTTTATCTGTGCTTCCCTATGCACTAATCAAAGGAATCATCCTCCGGA AAATATTCACCACCATCTCATTCTGCATTGTTCTGCGCATGGCGGTCACTCGGCAATTT CCCTGGGCTGTACAAACATGGTATGACTCTCTTGGAGCAATAAACAAAATACAGGATTT CTTACAAAAGCAAGAATATAAGACATTGGAATATAACTTAACGACTACAGAAGTAGTGA TGGAGAATGTAACAGCCTTCTGGGAGGAGGGATTTGGGGAATTATTTGAGAAAGCAAAA CAAAACAATAACAATAGAAAAACTTCTAATGGTGATGACAGCCTCTTCTTCAGTAATTT CTCACTTCTTGGTACTCCTGTCCTGAAAGATATTAATTTCAAGATAGAAAGAGGACAGT TGTTGGCGGTTGCTGGATCCACTGGAGCAGGCAAGACTTCACTTCTAATGATGATTATG GGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTCATTCTGTTC TCAGTTTTCCTGGATTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTTTCCT ATGATGAATATAGATACAGAAGCGTCATCAAAGCATGCCAACTAGAAGAGGACATCTCC AAGTTTGCAGAGAAAGACAATATAGTTCTTGGAGAAGGTGGAATCACACTGAGTGGAGG TCAACGAGCAAGAATTTCTTTAGCAAGAGCAGTATACAAAGATGCTGATTTGTATTTAT TAGACTCTCCTTTTGGATACCTAGATGTTTTAACAGAAAAAGAAATATTTGAAAGCTGT GTCTGTAAACTGATGGCTAACAAAACTAGGATTTTGGTCACTTCTAAAATGGAACATTT AAAGAAAGCTGACAAAATATTAATTTTGAATGAAGGTAGCAGCTATTTTTATGGGACAT TTTCAGAACTCCAAAATCTACAGCCAGACTTTAGCTCAAAACTCATGGGATGTGATTCT TTCGACCAATTTAGTGCAGAAAGAAGAAATTCAATCCTAACTGAGACCTTACACCGTTT CTCATTAGAAGGAGATGCTCCTGTCTCCTGGACAGAAACAAAAAAACAATCTTTTAAAC AGACTGGAGAGTTTGGGGAAAAAAGGAAGAATTCTATTCTCAATCCAATCAACTCTATA CGAAAATTTTCCATTGTGCAAAAGACTCCCTTACAAATGAATGGCATCGAAGAGGATTC TGATGAGCCTTTAGAGAGAAGGCTGTCCTTAGTACCAGATTCTGAGCAGGGAGAGGCGA TACTGCCTCGCATCAGCGTGATCAGCACTGGCCCCACGCTTCAGGCACGAAGGAGGCAG TCTGTCCTGAACCTGATGACACACTCAGTTAACCAAGGTCAGAACATTCACCGAAAGAC AACAGCATCCACACGAAAAGTGTCACTGGCCCCTCAGGCAAACTTGACTGAACTGGATA TATATTCAAGAAGGTTATCTCAAGAAACTGGCTTGGAAATAAGTGAAGAAATTAACGAA GAAGACTTAAAGGAGTGCCTTTTTGATGATATGGAGAGCATACCAGCAGTGACTACATG GAACACATACCTTCGATATATTACTGTCCACAAGAGCTTAATTTTTGTGCTAATTTGGT GCTTAGTAATTTTTCTGGCAGAGGTGGCTGCTTCTTTGGTTGTGCTGTGGCTCCTTGGA AACACTCCTCTTCAAGACAAAGGGAATAGTACTCATAGTAGAAATAACAGCTATGCAGT GATTATCACCAGCACCAGTTCGTATTATGTGTTTTACATTTACGTGGGAGTAGCCGACA CTTTGCTTGCTATGGGATTCTTCAGAGGTCTACCACTGGTGCATACTCTAATCACAGTG TCGAAAATTTTACACCACAAAATGTTACATTCTGTTCTTCAAGCACCTATGTCAACCCT CAACACGTTGAAAGCAGGTGGGATTCTTAATAGATTCTCCAAAGATATAGCAATTTTGG ATGACCTTCTGCCTCTTACCATATTTGACTTCATCCAGTTGTTATTAATTGTGATTGGA GCTATAGCAGTTGTCGCAGTTTTACAACCCTACATCTTTGTTGCAACAGTGCCAGTGAT AGTGGCTTTTATTATGTTGAGAGCATATTTCCTCCAAACCTCACAGCAACTCAAACAAC TGGAATCTGAAGGCAGGAGTCCAATTTTCACTCATCTTGTTACAAGCTTAAAAGGACTA TGGACACTTCGTGCCTTCGGACGGCAGCCTTACTTTGAAACTCTGTTCCACAAAGCTCT GAATTTACATACTGCCAACTGGTTCTTGTACCTGTCAACACTGCGCTGGTTCCAAATGA GAATAGAAATGATTTTTGTCATCTTCTTCATTGCTGTTACCTTCATTTCCATTTTAACA ACAGGAGAAGGAGAAGGAAGAGTTGGTATTATCCTGACTTTAGCCATGAATATCATGAG TACATTGCAGTGGGCTGTAAACTCCAGCATAGATGTGGATAGCTTGATGCGATCTGTGA GCCGAGTCTTTAAGTTCATTGACATGCCAACAGAAGGTAAACCTACCAAGTCAACCAAA CCATACAAGAATGGCCAACTCTCGAAAGTTATGATTATTGAGAATTCACACGTGAAGAA AGATGACATCTGGCCCTCAGGGGGCCAAATGACTGTCAAAGATCTCACAGCAAAATACA CAGAAGGTGGAAATGCCATATTAGAGAACATTTCCTTCTCAATAAGTCCTGGCCAGAGG GTGGGCCTCTTGGGAAGAACTGGATCAGGGAAGAGTACTTTGTTATCAGCTTTTTTGAG ACTACTGAACACTGAAGGAGAAATCCAGATCGATGGTGTGTCTTGGGATTCAATAACTT TGCAACAGTGGAGGAAAGCCTTTGGAGTGATACCACAGAAAGTATTTATTTTTTCTGGA ACATTTAGAAAAAACTTGGATCCCTATGAACAGTGGAGTGATCAAGAAATATGGAAAGT TGCAGATGAGGTTGGGCTCAGATCTGTGATAGAACAGTTTCCTGGGAAGCTTGACTTTG TCCTTGTGGATGGGGGCTGTGTCCTAAGCCATGGCCACAAGCAGTTGATGTGCTTGGCT AGATCTGTTCTCAGTAAGGCGAAGATCTTGCTGCTTGATGAACCCAGTGCTCATTTGGA TCCAGTAACATACCAAATAATTAGAAGAACTCTAAAACAAGCATTTGCTGATTGCACAG TAATTCTCTGTGAACACAGGATAGAAGCAATGCTGGAATGCCAACAATTTTTGGTCATA GAAGAGAACAAAGTGCGGCAGTACGATTCCATCCAGAAACTGCTGAACGAGAGGAGCCT CTTCCGGCAAGCCATCAGCCCCTCCGACAGGGTGAAGCTCTTTCCCCACCGGAACTCAA	

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	GCAAGTGCAAGTCTAAGCCCCAGATTGCTGCTCTGAAAGAGGAGACAGAAGAAGAGGTG CAAGATACAAGGCTTTAG	
CFTR0006	ATGTACCCATACGATGTTCCTGACTATGCGGCGGCCGTTATGCAGAGAAGCCCCCTGGA AAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGCTGGACCCGGCCCATCCTGCGGAAGG GCTACAGACAGAGACTGGAACTGAGCGACATCTACCAGATCCCCAGCGTGGACAGCGCC GACAACCTGAGCGAGAAGCTGGAAAGAGAGTGGGACAGAGAGCTGGCCAGCAAGAAGAA CCCCAAGCTGATCAACGCCCTGCGGCGGTGCTTCTTCTGGCGGTTCATGTTCTACGGCA TCTTCCTGTACCTGGGCGAAGTGACCAAAGCCGTGCAGCCCCTGCTGCTGGGCAGAATC ATCGCCAGCTACGACCCCGACAACAAAGAGGAACGGAGCATCGCCATCTACCTCGGCAT CGGCCTGTGCCTGCTGTTCATCGTCAGAACCCTGCTGCTGCACCCCGCCATCTTCGGAC TGCACCACATCGGCATGCAGATGCGGATCGCCATGTTCAGCCTGATCTACAAGAAAACC CTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATCGGACAGCTGGTGAGCCTGCT GAGCAACAACCTGAACAAGTTCGACGAAGGCCTGGCCCTGGCCCACTTCGTGTGGATCG CCCCCCTGCAAGTGGCCCTGCTGATGGGCCTGATCTGGGAACTGCTGCAGGCCAGCGCC TTCTGCGGACTGGGATTCCTGATCGTGCTGGCCCTGTTCCAGGCCGGACTGGGGAGAAT GATGATGAAGTACCGGGACCAGAGAGCCGGCAAGATCAGCGAGAGACTGGTCATCACCA GCGAGATGATCGAGAACATCCAGAGCGTGAAGGCCTACTGCTGGGAAGAGGCCATGGAA AAGATGATCGAGAACCTGCGGCAGACCGAGCTGAAGCTGACAAGAAAGGCCGCCTACGT GCGCTACTTCAACAGCAGCGCCTTCTTCTTCAGCGGCTTCTTCGTGGTGTTCCTGAGCG TGCTGCCCTACGCCCTGATCAAGGGCATCATCCTGAGAAAGATCTTCACCACCATCAGC TTCTGCATCGTGCTGCGGATGGCCGTGACCAGACAGTTCCCCTGGGCCGTGCAGACCTG GTACGACAGCCTGGGCGCCATCAACAAGATCCAGGACTTCCTGCAGAAGCAAGAGTACA AGACCCTCGAGTACAACCTGACCACCACCGAGGTGGTCATGGAAAACGTGACCGCCTTC TGGGAGGAAGGCTTCGGCGAGCTGTTCGAGAAGGCCAAGCAGAACAACAACAACCGCAA GACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTTCAGCCTGCTGGGGACCCCCG TGCTGAAGGACATCAACTTCAAGATCGAGCGGGGACAGCTGCTGGCCGTGGCCGGAAGC ACAGGCGCCGGAAAAACCAGCCTGCTCATGGTCATCATGGGCGAGCTGGAACCCAGCGA GGGCAAGATCAAGCACAGCGGCAGGATCAGCTTCTGCAGCCAGTTCAGCTGGATCATGC CCGGCACCATCAAAGAGAACATCATCTTCGGCGTGAGCTACGACGAGTACAGATACCGC AGCGTGATCAAGGCCTGCCAGCTGGAAGAGGACATCAGCAAGTTCGCCGAGAAGGACAA CATCGTGCTCGGCGAAGGCGGCATCACACTGAGCGGCGGACAGAGGGCCAGAATCAGCC TGGCCAGAGCCGTGTACAAGGACGCCGACCTGTACCTGCTGGACAGCCCCTTCGGCTAC CTGGACGTGCTGACCGAGAAAGAGATCTTCGAGAGCTGCGTGTGCAAGCTGATGGCCAA CAAGACCCGGATCCTGGTCACCAGCAAGATGGAACACCTGAAGAAGGCCGACAAGATCC TGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCTTCAGCGAGCTGCAGAACCTG CAGCCCGACTTCAGCAGCAAACTGATGGGCTGCGACAGCTTCGACCAGTTCAGCGCCGA GCGGAGAAACAGCATCCTGACAGAGACACTGCACCGGTTCAGCCTGGAAGGCGACGCCC CCGTGAGCTGGACCGAGACAAAGAAGCAGAGCTTCAAGCAGACCGGCGAGTTCGGCGAG AAGCGGAAGAACAGCATCCTGAACCCCATCAACAGCATCCGGAAGTTCAGCATCGTCCA GAAAACCCCCCTGCAGATGAACGGCATCGAAGAGGACAGCGACGAGCCCCTGGAAAGAC GGCTGAGCCTGGTGCCCGACAGCGAACAGGGCGAAGCCATCCTGCCCCGGATCAGCGTG ATCAGCACAGGCCCCACACTGCAGGCCCGGAGAAGGCAGAGCGTGCTGAACCTGATGAC CCACAGCGTGAACCAGGGACAGAACATCCACAGAAAGACCACCGCCAGCACACGGAAAG TGAGCCTGGCCCCCCAGGCCAACCTGACTGAGCTGGACATCTACAGCAGACGGCTGAGC CAAGAGACAGGCCTGGAAATCAGCGAGGAAATCAACGAAGAGGACCTGAAAGAGTGCTT CTTCGACGACATGGAAAGCATCCCCGCCGTGACAACCTGGAACACCTACCTGCGGTACA TCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGGTGCCTCGTGATCTTCCTGGCC GAAGTGGCCGCCAGCCTGGTGGTGCTGTGGCTGCTCGGAAACACCCCACTGCAGGACAA GGGCAACAGCACCCACAGCCGGAACAACAGCTACGCCGTGATCATCACCAGCACCAGCA GCTACTACGTGTTCTACATCTACGTGGGCGTCGCCGACACTCTGCTCGCCATGGGCTTC TTCAGAGGACTGCCCCTGGTGCACACCCTGATCACCGTGAGCAAGATCCTGCACCACAA GATGCTGCACAGCGTCCTGCAGGCCCCCATGAGCACACTGAACACCCTGAAAGCCGGCG GAATCCTGAACAGATTCAGCAAGGACATCGCCATCCTGGACGACCTGCTGCCCCTGACC ATCTTCGACTTCATCCAGCTGCTGCTGATCGTGATCGGCGCCATCGCCGTGGTGGCCGT GCTGCAGCCCTACATCTTCGTGGCCACCGTGCCCGTGATCGTGGCCTTCATCATGCTGC GGGCCTACTTCCTGCAGACCAGCCAGCAGCTGAAGCAGCTCGAGAGCGAGGGCAGAAGC CCCATCTTCACCCACCTCGTGACCAGCCTGAAAGGCCTGTGGACCCTGAGAGCCTTCGG CAGACAGCCCTACTTCGAGACACTGTTCCACAAGGCCCTGAACCTGCACACCGCCAACT GGTTCCTGTACCTGAGCACCCTGCGGTGGTTCCAGATGAGGATCGAGATGATCTTCGTC ATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTCACCACTGGCGAAGGCGAGGGCAG	**23**

Construct	DNA sequence (from 5' to 3')	SEQ ID NO.
	AGTGGGAATCATCCTGACCCTGGCCATGAACATCATGAGCACACTCCAGTGGGCCGTGA ACAGCAGCATCGACGTGGACAGCCTGATGCGGAGCGTGAGCCGGGTGTTCAAGTTCATC GACATGCCCACAGAGGGCAAGCCCACCAAGAGCACCAAGCCCTACAAGAACGGCCAGCT GAGCAAAGTCATGATCATCGAGAACAGCCACGTCAAGAAGGACGACATCTGGCCCAGCG GAGGCCAGATGACCGTGAAGGACCTGACCGCCAAGTACACCGAAGGCGGAAACGCCATC CTGGAAAACATCAGCTTCAGCATCAGCCCCGGCCAGCGCGTGGGACTCCTGGGAAGAAC CGGAAGCGGCAAGAGCACTCTGCTGAGCGCCTTCCTGAGACTGCTGAACACCGAGGGCG AGATCCAGATCGACGGGGTGAGCTGGGACAGCATCACCCTGCAACAATGGCGGAAGGCC TTCGGCGTGATCCCCCAGAAGGTGTTCATCTTCAGCGGCACGTTCCGGAAGAACCTGGA CCCCTACGAGCAGTGGAGCGACCAAGAGATCTGGAAGGTGGCCGACGAAGTGGGACTGA GAAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGACTTCGTGCTGGTGGACGGCGGCTGC GTGCTGAGCCACGGACACAAGCAGCTGATGTGCCTGGCCAGAAGCGTGCTGAGCAAGGC CAAGATCCTGCTGCTCGACGAGCCCAGCGCCCACCTGGACCCCGTGACCTACCAGATCA TCCGGCGGACACTGAAGCAGGCCTTCGCCGACTGCACCGTGATCCTGTGCGAGCACAGA ATCGAGGCCATGCTGGAATGCCAGCAGTTCCTGGTGATCGAAGAGAACAAAGTGCGGCA GTACGACAGCATCCAGAAGCTGCTGAACGAGCGGAGCCTGTTCAGACAGGCCATCAGCC CCAGCGACAGAGTGAAGCTGTTCCCCCACCGGAACAGCAGCAAGTGCAAGAGCAAGCCC CAGATCGCCGCCCTGAAAGAAGAAACCGAGGAAGAGGTGCAGGACACACGGCTGTGA	

Untranslated Regions

[00123] In some embodiments, the polynucleotide of the present disclosure further comprises 3'- or 5'-untranslated regions or a 3'- or 5'-noncoding region. In some embodiments, said untranslated region or noncoding region improves a pharmacokinetic characteristic (e.g., a prolonged half-life) of said synthetic polynucleotide in a cell. In some embodiments, the polynucleotide of the present disclosure comprises a 5' untranslated region (UTR) or 3' UTR having at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to one set forth in SEQ ID NOs 6-22. In some embodiments, the polynucleotide comprises a 5' cap structure. In some embodiments, the 5' cap structure comprises a sequence having at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to SEQ ID 6. In some embodiments, the polynucleotide comprises a 3' poly adenosine tail. In some embodiments, the 3' poly adenosine tail comprises a sequence having at least 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity with SEQ ID NOs 7 or 8.

TABLE 2. Example untranslated region sequences

UTR	DNA sequence (from 5' to 3')	SEQ ID NO.
5'UTR	GGGAGACCCAAGCTGGCTAGCGTTTAAACTTCAGCTTGGCAATCCGGTACTGTTGGTA AAGCCACC	**6**
3' UTR-poly(A)-001	GAATTCTGCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAATTCG	**7**
3' UTR-poly(A)-002	GAATTCTGCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAATT	**8**
α-globin 5' UTR (HBA1)	GGGAGACATAAACCCTGGCGCGCTCGCGGCCCGGCACTCTTCTGGTCCCCACAGACTC AGAGAGAAGCCACC	**9**

UTR	DNA sequence (from 5' to 3')	SEQ ID NO.
α-globin 5' UTR (HBA2)	GGGAGACATAAACCCTGGCGCGCTCGCGGGCCGGCACTCTTCTGGTCCCCACAGACTC AGAGAGAAGCCACC	**10**
α-globin 5' UTR	GGGAGACTCTTCTGGTCCCCACAGACTCAGAGAGAACGCCACC	**11**
IRES of EMCV 5'-UTR	GTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCT GTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTC TGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTC TGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGC CAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTG TGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGG GCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGC ACATGCTTTACGTGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGG GGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAACC	**12**
IRES of TEV 5'-UTR	AAATAACAAATCTCAACACAACATATACAAAACAAACGAATCTCAAGCAATCAAGCAT TCTACTTCTATTGCAGCAATTTAAATCATTTCTTTTAAAGCAAAAGCAATTTTCTGAA AATTTTCACCATTTACGAACGATAGCA	**13**
ssRNA1 5'UTR	GGGAGACAAGAGAGAAAAGAAGAGCAAGAAGAAATATAAGAGCCACC	**14**
ssRNA2 5'UTR	GGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGCAATCCGGTACTGTTGGTA AAGCCACC	**15**
native 5' UTR	GTAGTAGGTCTTTGGCATTAGGAGCTTGAGCCCAGACGGCCCTAGCAGGGACCCCAGC GCCCGAGAGACC	**16**
TMV 3'-UTR	GGATTGTGTCCGTAATCACACGTGGTGCGTACGATAACGCATAGTGTTTTTCCCTCCA CTTAAATCGAAGGGTTGTGTCTTGGATCGCGCGGGTCAAATGTATATGGTTCATATAC ATCCGCAGGCACGTAATAAAGCGAGGGGTTCGAATCCCCCCGTTACCCCCGGTAGGGG CCCATTGTCTTC	**17**
MALAT1 3'-UTR	TCAGTAGGGTCATGAAGGTTTTTCTTTTCCTGAGAAAACAACACGTATTGTTTTCTCA GGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAAAAAAAAAGCAAAATTGTCTTC	**18**
NEAT2 3'-UTR	TCAGTAGGGTTGTAAAGGTTTTTCTTTTCCTGAGAAAACAACCTTTTGTTTTCTCAGG TTTTGCTTTTTGGCCTTTCCCTAGCTTTAAAAAAAAAAAAGCAAAATTGTCTTC	**19**
histone cluster 2, H3c 3'-UTR	GAAGTGGCGGTTCGGCCGGAGGTTCCATCGTATCCAAAAGGCTCTTTTCAGAGCCACC CATTGTCTTC	**20**
Native 3' UTR	AGAGCAGCATAAATGTTGACATGGGACATTTGCTCATGGAATTGGAGCTCGTGGGACA GTCACCTCATGGAATTGGAGCTCGTGGAACAGTTACCTCTGCCTCAGAAAACAAGGAT GAATTAAGTTTTTTTTTAAAAAAGAAACATTTGGTAAGGGGAATTGAGGACACTGATA TGGGTCTTGATAAATGGCTTCCTGGCAATAGTCAAATTGTGTGAAAGGTACTTCAAAT CCTTGAAGATTTACCACTTGTGTTTTGCAAGCCAGATTTTCCTGAAAACCCTTGCCAT GTGCTAGTAATTGGAAAGGCAGCTCTAAATGTCAATCAGCCTAGTTGATCAGCTTATT GTCTAGTGAAACTCGTTAATTTGTAGTGTTGGAGAAGAACTGAAATCATACTTCTTAG GGTTATGATTAAGTAATGATAACTGGAAACTTCAGCGGTTTATATAAGCTTGTATTCC TTTTTCTCTCCTCTCCCCATGATGTTTAGAAACACAACTATATTGTTTGCTAAGCATT CCAACTATCTCATTTCCAAGCAAGTATTAGAATACCACAGGAACCACAAGACTGCACA TCAAAATATGCCCCATTCAACATCTAGTGAGCAGTCAGGAAAGAGAACTTCCAGATCC TGGAAATCAGGGTTAGTATTGTCCAGGTCTACCAAAAATCTCAATATTTCAGATAATC ACAATACATCCCTTACCTGGGAAAGGGCTGTTATAATCTTTCACAGGGGACAGGATGG TTCCCTTGATGAAGAAGTTGATATGCCTTTTCCCAACTCCAGAAAGTGACAAGCTCAC AGACCTTTGAACTAGAGTTTAGCTGGAAAAGTATGTTAGTGCAAATTGTCACAGGACA GCCCTTCTTTCCACAGAAGCTCCAGGTAGAGGGTGTGTAAGTAGATAGGCCATGGGCA CTGTGGGTAGACACACATGAAGTCCAAGCATTTAGATGTATAGGTTGATGGTGGTATG TTTTCAGGCTAGATGTATGTACTTCATGCTGTCTACACTAAGAGAGAATGAGAGACAC ACTGAAGAAGCACCAATCATGAATTAGTTTTATATGCTTCTGTTTTATAATTTTGTGA AGCAAAATTTTTTCTCTAGGAAATATTTATTTTAATAATGTTTCAAACATATATAACA ATGCTGTATTTTAAAAGAATGATTATGAATTACATTTGTATAAAATAATTTTTATATT TGAAATATTGACTTTTTATGGCACTAGTATTTCTATGAAATATTATGTTAAAACTGGG ACAGGGGAGAACCTAGGGTGATATTAACCAGGGGCCATGAATCACCTTTTGGTCTGGA GGGAAGCCTTGGGGCTGATGCAGTTGTTGCCCACAGCTGTATGATTCCCAGCCAGCAC AGCCTCTTAGATGCAGTTCTGAAGAAGATGGTACCACCAGTCTGACTGTTTCCATCAA	**21**

UTR	DNA sequence (from 5' to 3')	SEQ ID NO.
	GGGTACACTGCCTTCTCAACTCCAAACTGACTCTTAAGAAGACTGCATTATATTTATT ACTGTAAGAAAATATCACTTGTCAATAAAATCCATACATTTGTGTGAAA	
ssRNA2 5'UTR (A32C)	GGGAGACCCAAGCTGGCTAGCGTTTAAACTTCAGCTTGGCAATCCGGTACTGTTGGTA AAGCCACC	**22**

Nucleotide analogues

[00124] In some embodiments of the synthetic polynucleotide, the polynucleotide may comprise one or more nucleotide analogues. In some embodiments, the nucleotide analogues replace uridines in a sequence. For example, a sequence using standard nucleotides (A, C, U, T, G) may comprises a uridine at a particular position in a sequence. A sequence may instead have a nucleotide analogue in place of the uridine. The nucleotide analogue may have structure that may still be recognized by the cellular translation machinery such that the polynucleotide comprising a nucleotide analogue may still be translated. The nucleotide analogue may be recognized as synonymous with a standard nucleotide. For example, the nucleotide analogue may be recognized as synonymous with uridine and the resulting translation product is generated as if the nucleotide analogue is a uridine. In some embodiments, at least about 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of nucleotides replacing uridine within said polynucleotide are nucleotide analogues. In some embodiments, fewer than about 15% of nucleotides within said polynucleotide are nucleotide analogues In some fewer than about 30% of the nucleotides are nucleotide analogues. In other cases, fewer than about 27.5%, fewer than about 25%, fewer than about 22.5%, fewer than about 20%, fewer than about 17.5%, fewer than about 15%, fewer than about 12.5%, fewer than about 10%, fewer than about 7.5%, fewer than about 5%, or fewer than about 2.5% of the nucleotides are nucleotide analogues.

[00125] A polyribonucleotide can have the same or a mixture of different nucleotide analogues or modified nucleotides. The nucleotide analogues or modified nucleotides can have structural changes that are naturally or not naturally occurring in messenger RNA. A mixture of various analogues or modified nucleotides can be used. For example, one or more analogues within a polynucleotide can have natural modifications, while another part has modifications that are not naturally found in mRNA. Additionally, some analogues or modified ribonucleotides can have a base modification, while other modified ribonucleotides have a sugar modification. In the same way, it is possible that all modifications are base modifications, or all modifications are sugar modifications or any suitable mixture thereof.

[00126] A nucleotide analogue or modified nucleotide can be selected from the group comprising pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 3-methyluridine, 5-carboxymethyl- uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethy1-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-l-methyl-pseudouridine, 2-thio-l-methyl-pseudouridine, 1-methyl-l-deaza-pseudouridine, 2-thio-1-methyl-l-deaza-pseudouridine, dihydrouridine, dihydropseudouridine, 2-thio-

dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-l-methyl-pseudoisocytidine, 4-thio-l-methyl-1-deaza-pseudoisocytidine, 1-methyl-l-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, 4-methoxy-l-methyl-pseudoisocytidine, 2-aminopurine, 2, 6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6- diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, 2-methoxy-adenine, inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

[00127] In some embodiments of the synthetic polynucleotide, the nucleotide analogue is a purine or pyrimidine analogue. In some cases, a polyribonucleotide of the disclosure comprises a modified pyrimidine, such as a modified uridine. A nucleotide analogue may be a pseudouridine (Ψ). A nucleotide analogue may be a methylpseudouridine. A nucleotide analogue may be a 1-methylpseudouridine ($m^1\Psi$). In some embodiments, the polynucleotide comprises a 1-methylpseudouridine. In some cases a uridine analogue is selected from pseudouridine 1-methylpseudouridine, 2-thiouridine (s^2U), 5-methyluridine (m^5U), 5-methoxyuridine (mo^5U), 4-thiouridine (s^4U), 5-bromouridine (Br^5U), 2´O-methyluridine (U2´m), 2´-amino-2´-deoxyuridine (U2´NH_2), 2´-azido-2´-deoxyuridine (U2´N_3), and 2´-fluoro-2´-deoxyuridine (U2´F).

[00128] In some embodiments, the synthetic polynucleotide comprises (e.g., one or more) 1-methylpseudouridine. In some embodiments, at least about 80% of nucleotides replacing uridine within said polynucleotide are 1-methylpseudouridine. In some embodiments, at least (about) 5%, 10%, 15%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of nucleotides replacing uridine within said polynucleotide are 1-methylpseudouridine. In some embodiments, 100% of nucleotides replacing uridine within said polynucleotide are 1-methylpseudouridine.

[00129] Nucleic Acid Constructs, Vectors, and Engineered Polyribonucleotides

[00130] In some embodiments, the present disclosure provides nucleic acid molecules, such as polynucleotides, which encode one or more polypeptides of interest. The term nucleic acid includes any

compound and/or substance that comprise a polymer of nucleotides. Nucleotide polymers that contain greater than 50% of ribose bases or ribonucleotide analogues are referred to as polyribonucleotides. Nucleotide polymers may use altered nucleotide usage that encode a protein or functional fragment thereof, such as CFTR. The sequence of the engineered polynucleotides can be derived from, for example, DNA, RNA, mRNA transcripts, genomic DNA, mitochondrial DNA, mitochondrial RNA, or another suitable nucleic acid that comprises the genetic information of a gene of interest. The nucleic acid constructs, vectors, engineered polyribonucleotides, or compositions can be derived from nucleic acids carrying mutated genes and polymorphisms.

[00131] In addition to the four canonical ribonucleotides, namely, adenosine, guanosine, cytidine and uridine, several cellular RNAs also contain a number of structurally diverse ribonucleotides. About a hundred structurally different nucleotides or nucleotide analogues have been identified in transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), messenger RNAs (mRNAs) and small nuclear RNAs (snRNAs). In tRNAs, some nucleotides can be important determinants of the specificity and efficiency of aminoacylation and codon recognition. Such structurally diverse ribonucleotides can be a modified ribonucleotide or a nucleotide analogue. In some cases, a polynucleotide of the disclosure is engineered to comprise a ribonucleotide analogue.

[00132] In some cases, a nucleic acid construct, a vector, or a polynucleotide is engineered to contain the four classical ribonucleotides and can be modified post-transcriptionally, after being administered to a subject. For instance, in some cases the disclosure provides a composition, vector, or a nucleic acid construct comprising a nucleic acid construct encoding CFTR, wherein fewer than 30% of the nucleic acids encoding CFTR are nucleotide analogues. In other cases, fewer than 27.5%, fewer than 25%, fewer than 22.5%, fewer than 20%, fewer than 17.5%, fewer than 15%, fewer than 12.5%, fewer than 10%, fewer than 7.5%, fewer than 5%, or fewer than 2.5% of the nucleotides encoding CFTR are nucleotide analogues.

[00133] Example nucleic acids that can form a polynucleotide of the disclosure include, but are not limited to, ribonucleic acids (RNAs), deoxyribonucleic acids (DNAs), or hybrids thereof. Example modified nucleotides that can form at least a fraction of a polynucleotide of the disclosure include, but are not limited to, pseudouridine (T) and 1-methylpseudouridine (m1Ψ).

[00134] A chemical modification can be located on one or more nucleoside(s) or the backbone of the nucleic acid molecule. They can be located on both a nucleoside and a backbone linkage. A modification can be engineered into a polynucleotide in vitro. Modified ribonucleotides and nucleic acid analogues can also be introduced post-transcriptionally by covalent modification of the classical ribonucleotides.

[00135] A nucleic acid construct, a vector, or an engineered polyribonucleotide of the disclosure can comprise purine and pyrimidine analogues. In some cases, a polyribonucleotide of the disclosure comprises a modified pyrimidine, such as a modified uridine. In some cases, a uridine analogue is selected from pseudouridine (Ψ), 1-methylpseudouridine (m1Ψ), 2-thiouridine (s2U), 5-methyluridine (m5U), 5-methoxyuridine (mo5U), 4-thiouridine (s4U), 5-bromouridine (Br5U), 2′O-methyluridine (U2′m), 2′-amino-2′-deoxyuridine (U2′NH2), 2′-azido-2′-deoxyuridine (U2′N3), and 2′-fluoro-2′-deoxyuridine (U2F).

[00136] In some instances, the nucleic acid construct(s), vector(s), engineered polyribonucleotide(s), or composition(s) encodes CFTR or a variant thereof at a level that is increased by a factor of at least about 1.5 as compared to levels within cells exposed to a composition comprising a nucleic acid construct that does not include the codons encoding CFTR or a variant thereof. In some cases, the factor is at least about 1.1, at least about 1.2, at least about 1.3, at least about 1.4, at least about 1.5, at least about 2, at least about 3, at least about 4, at least about 5, at least about 10, at least about 20, at least about 30, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 90, or at least about 100.

[00137] A polyribonucleotide can have the same or a mixture of different nucleotide analogues or modified nucleotides. The nucleotide analogues or modified nucleotides can have structural changes that are naturally or not naturally occurring in messenger RNA. A mixture of various analogues or modified nucleotides can be used. For example, one or more analogues within a polynucleotide can have natural modifications, while another part has modifications that are not naturally found in mRNA. Additionally, some analogues or modified ribonucleotides can have a base modification, while other modified ribonucleotides have a sugar modification. In the same way, it is possible that all modifications are base modifications or all modifications are sugar modifications or any suitable mixture thereof.

[00138] A nucleotide analogue or modified nucleotide can be selected from the group comprising pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine, dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, 4-methoxy-1-methyl-pseudoisocytidine, 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, 2-methoxy-adenine, inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-

thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

[00139] In some cases, at least about 5% of the nucleic acid construct(s), a vector(s), engineered polyribonucleotide(s), or compositions includes non-naturally occurring (e.g., modified, analogues, or engineered) uridine, adenosine, guanine, or cytosine, such as the nucleotides described herein. In some cases, 100% of the modified nucleotides in the composition are either 1-methylpseudouridine or pseudouridine. In some cases, at least about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% of the nucleic acid construct(s), a vector(s), engineered polyribonucleotide(s), or compositions includes non-naturally occurring uracil, adenine, guanine, or cytosine. In some cases, at most about 99%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 5%, 1%, of the nucleic acid construct(s), a vector(s), engineered polyribonucleotide(s), or compositions includes non-naturally occurring uracil, adenine, guanine, or cytosine.

[00140] A nucleic acid construct(s), a vector(s), or an engineered polyribonucleotide(s) of the disclosure can comprise one or more promoter sequences and any associated regulatory sequences. A promoter sequence and/or an associated regulatory sequence can comprise any number of modified or unmodified nucleotides, and any number of nucleic acid analogues. Promoter sequences and/or any associated regulatory sequences can comprise, for example, at least 2 bases or base pairs, 3 bases or base pairs, 4 bases or base pairs, 5 bases or base pairs, 6 bases or base pairs, 7 bases or base pairs, 8 bases or base pairs, 9 bases or base pairs, 10 bases or base pairs, 11 bases or base pairs, 12 bases or base pairs, 13 bases or base pairs, 14 bases or base pairs, 15 bases or base pairs, 16 bases or base pairs, 17 bases or base pairs, 18 bases or base pairs, 19 bases or base pairs, 20 bases or base pairs, 21 bases or base pairs, 22 bases or base pairs, 23 bases or base pairs, 24 bases or base pairs, 25 bases or base pairs, 26 bases or base pairs, 27 bases or base pairs, 28 bases or base pairs, 29 bases or base pairs, 30 bases or base pairs, 35 bases or base pairs, 40 bases or base pairs, 50 bases or base pairs, 75 bases or base pairs, 100 bases or base pairs, 150 bases or base pairs, 200 bases or base pairs, 300 bases or base pairs, 400 bases or base pairs, 500 bases or base pairs, 600 bases or base pairs, 700 bases or base pairs, 800 bases or base pairs, 900 bases or base pairs, 1000 bases or base pairs, 2000 bases or base pairs, 3000 bases or base pairs, 4000 bases or base pairs, 5000 bases or base pairs, at least 10000 bases or base pairs or more. A promoter sequence and/or an associated regulatory sequence can comprise any number of modified or unmodified nucleotides, for example, at most 10000 bases or base pairs, 5000 bases or base pairs, 4000 bases or base pairs, 3000 bases or base pairs, 2000 bases or base pairs, 1000 bases or base pairs, 900 bases or base pairs, 800 bases or base pairs, 700 bases or base pairs, 600 bases or base pairs, 500 bases or base pairs, 400 bases or base pairs, 300 bases or base pairs, 200 bases or base pairs, 100 bases or base pairs, 75 bases or base pairs, 50 bases or base pairs, 40 bases or base pairs, 35 bases or base pairs, 30 bases or base pairs, 29 bases or base pairs, 28 bases or base pairs, 27 bases

or base pairs, 26 bases or base pairs, 25 bases or base pairs, 24 bases or base pairs, 23 bases or base pairs, 22 bases or base pairs, 21 bases or base pairs, 20 bases or base pairs, 19 bases or base pairs, 18 bases or base pairs, 17 bases or base pairs, 16 bases or base pairs, 15 bases or base pairs, 14 bases or base pairs, 13 bases or base pairs, 12 bases or base pairs, 11 bases or base pairs, 10 bases or base pairs, 9 bases or base pairs, 8 bases or base pairs, 7 bases or base pairs, 6 bases or base pairs, 5 bases or base pairs, 4 bases or base pairs, 3 bases or base pairs or 2 bases or base pairs.

[00141] In some cases, less than all of the nucleotides in the promoter sequence or associated regulatory region are nucleotide analogues or modified nucleotides. For instance, in some cases, less than or equal to 99%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5% of the nucleotides in a promoter or associated regulatory region. In some cases, all of the nucleotides in a promoter or associated regulatory region are nucleic acid analogues or modified nucleotides.

[00142] A nucleic acid construct(s), a vector(s), an engineered polyribonucleotide(s), or compositions of the disclosure can comprise an engineered 5′ cap structure, or a 5′-cap can be added to a polyribonucleotide intracellularly. The 5′cap structure of an mRNA can be involved in binding to the mRNA Cap Binding Protein (CBP), which is responsible for mRNA stability in the cell and translation competency through the association of CBP with poly(A) binding protein to form the mature pseudo-circular mRNA species. The 5′cap structure can also be involved in nuclear export, increases in mRNA stability, and in assisting the removal of 5′ proximal introns during mRNA splicing.

[00143] A nucleic acid construct(s), a vector(s), or an engineered polyribonucleotide(s) can be 5′-end capped generating a 5′-GpppN-3′-triphosphate linkage between a terminal guanosine cap residue and the 5′-terminal transcribed sense nucleotide of the mRNA molecule. The cap-structure can comprise a modified or unmodified 7-methylguanosine linked to the first nucleotide via a 5′-5′ triphosphate bridge. This 5′-guanylate cap can then be methylated to generate an N7-methyl-guanylate residue (Cap-0 structure). The ribose sugars of the terminal and/or anteterminal transcribed nucleotides of the 5′ end of the mRNA may optionally also be 2′-O-methylated (Cap-1 structure). 5′-decapping through hydrolysis and cleavage of the guanylate cap structure may target a nucleic acid molecule, such as an mRNA molecule, for degradation.

[00144] In some cases, a cap can comprise further modifications, including the methylation of the 2′ hydroxy-groups of the first 2 ribose sugars of the 5′ end of the mRNA. For instance, an eukaryotic cap-1 has a methylated 2′-hydroxy group on the first ribose sugar, while a cap-2 has methylated 2′-hydroxy groups on the first two ribose sugars. The 5′ cap can be chemically similar to the 3′ end of an RNA molecule (the 5′ carbon of the cap ribose is bonded, and the free 3′-hydroxyls on both 5′- and 3′-ends of the capped transcripts. Such double modification can provide significant resistance to 5′ exonucleases. Non-limiting examples of 5′ cap structures that can be used with an engineered polyribonucleotide include, but are not limited to, m7G(5′)ppp(5′)N(Cap-0), m7G(5′)ppp(5′)N1mpNp (Cap-1), and m7G(5′)-ppp(5)N1mpN2mp (Cap-2).

[00145] Modifications to the modified mRNA of the present disclosure may generate a non-hydrolyzable cap structure preventing decapping and thus increasing mRNA half-life while facilitating efficient

translation. Because cap structure hydrolysis requires cleavage of 5′-ppp-5′triphosphate linkages, modified nucleotides may be used during the capping reaction. For example, a Vaccinia Capping Enzyme from New England Biolabs (Ipswich, Mass.) may be used with guanosine α-thiophosphate nucleotides according to the manufacturer's instructions to create a phosphorothioate linkage in the 5′-ppp-5′ cap. Additional modified guanosine nucleotides may be used such as α-methyl-phosphonate and seleno-phosphate nucleotides. Additional modifications include, but are not limited to, 2′-O-methylation of the ribose sugars of 5′-terminal and/or 5′-anteterminal nucleotides of the mRNA on the 2′-hydroxyl group of the sugar ring. Multiple distinct 5′-cap structures can be used to generate the 5′-cap of a polyribonucleotide.

[00146] The modified mRNA may be capped post-transcriptionally. According to the present disclosure, 5′ terminal caps may include endogenous caps or cap analogues. According to the present disclosure, a 5′ terminal cap may comprise a guanine analogue. Useful guanine analogues include, but are not limited to, inosine, N1-methyl-guanosine, 2′fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

[00147] Further, a nucleic acid construct(s), a vector(s), or an engineered polyribonucleotide(s) can contain one or more internal ribosome entry site(s) (IRES). IRES sequences can initiate protein synthesis in absence of the 5′ cap structure. An IRES sequence can also be the sole ribosome binding site, or it can serve as one of multiple ribosome binding sites of an mRNA. Engineered polyribonucleotides containing more than one functional ribosome binding site can encode several peptides or polypeptides that are translated by the ribosomes (“polycistronic or multicistronic polynucleotides”). An engineered polynucleotide described here can comprise at least 1 IRES sequence, two IRES sequences, three IRES sequences, four IRES sequences, five IRES sequences, six IRES sequences, seven IRES sequences, eight IRES sequences, nine IRES sequences, ten IRES sequences, or another suitable number are present in an engineered polyribonucleotide. Examples of IRES sequences that can be used according to the present disclosure include without limitation, those from tobacco etch virus (TEV), picornaviruses (e.g., FMDV), pest viruses (CFFV), polio viruses (PV), encephalomyocarditis viruses (EMCV), foot-and-mouth disease viruses (FMDV), hepatitis C viruses (HCV), classical swine fever viruses (CSFV), murine leukemia virus (MLV), simian immune deficiency viruses (SIV) or cricket paralysis viruses (CrPV). An IRES sequence can be derived, for example, from commercially available vectors such as the IRES sequences available from Clontech™, GeneCopoeia™, or Sigma-Aldrich™. IRES sequences can be, for example, at least 150 bases or base pairs, 200 bases or base pairs, 300 bases or base pairs, 400 bases or base pairs, 500 bases or base pairs, 600 bases or base pairs, 700 bases or base pairs, 800 bases or base pairs, 900 bases or base pairs, 1000 bases or base pairs, 2000 bases or base pairs, 3000 bases or base pairs, 4000 bases or base pairs, 5000 bases or base pairs, or 10000 bases or base pairs. IRES sequences can at most 10000 bases or base pairs, 5000 bases or base pairs, 4000 bases or base pairs, 3000 bases or base pairs, 2000 bases or base pairs, 1000 bases or base pairs, 900 bases or base pairs, 800 bases or base pairs, 700 bases or base pairs, 600 bases or base pairs, 500 bases or base pairs, 400 bases or base pairs, 300 bases or base pairs, 200 bases or base pairs, 100 bases or base pairs, 50 bases or base pairs, or 10 bases or base pairs.

[00148] A nucleic acid construct(s), a vector(s), or an engineered polyribonucleotide(s) of the disclosure can comprise one or more untranslated regions. An untranslated region can comprise any number of modified or unmodified nucleotides. Untranslated regions (UTRs) of a gene are transcribed but not translated into a polypeptide. In some cases, an untranslated sequence can increase the stability of the nucleic acid molecule and the efficiency of translation. The regulatory features of a UTR can be incorporated into the modified mRNA molecules of the present disclosure, for instance, to increase the stability of the molecule. The specific features can also be incorporated to ensure controlled down-regulation of the transcript in case they are misdirected to undesired organs sites. Some 5′ UTRs play roles in translation initiation. A 5′ UTR can comprise a Kozak sequence which is involved in the process by which the ribosome initiates translation of many genes. Kozak sequences can have the consensus GCC(R)CCAUGG, where R is a purine (adenine or guanine) that is located three bases upstream of the start codon (AUG). A Kozak sequence may comprise GCCACC. 5′ UTRs may form secondary structures which are involved in binding of translation elongation factor. In some cases, one can increase the stability and protein production of the engineered polynucleotide molecules of the disclosure, by engineering the features typically found in abundantly expressed genes of specific target organs. For example, introduction of 5′UTR of liver-expressed mRNA, such as albumin, serum amyloid A, Apolipoprotein AB/E, transferrin, alpha fetoprotein, erythropoietin, or Factor VIII, can be used to increase expression of an engineered polynucleotide in a liver. Likewise, use of 5′ UTR from muscle proteins (MyoD, Myosin, Myoglobin, Myogenin, Herculin), for endothelial cells (Tie-1, CD36), for myeloid cells (C/EBP, AML1, G-CSF, GM-CSF, CD1 lb, 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) can be used to increase expression of an engineered polynucleotide in a desired cell or tissue.

[00149] Other non-UTR sequences can be incorporated into the 5′ (or 3′ UTR) UTRs of the polyribonucleotides of the present disclosure. The 5′ and/or 3′ UTRs can provide stability and/or translation efficiency of polyribonucleotides. For example, introns or portions of intron sequences can be incorporated into the flanking regions of an engineered polyribonucleotide. Incorporation of intronic sequences can also increase the rate of translation of the polyribonucleotide.

[00150] 3′ UTRs may have stretches of Adenosines and Uridines embedded therein. These AU rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU rich elements (AREs) can be separated into classes: Class I AREs contain several dispersed copies of an AUUUA motif within U-rich regions. C-Myc and MyoD contain class I AREs. Class II AREs possess two or more overlapping UUAUUUA(U/A)(U/A) nonamers. Molecules containing this type of AREs include GM-CSF and TNF-α. Class III ARES are less well defined. These U rich regions do not contain an AUUUA motif c-Jun and Myogenin are two well-studied examples of this class. Proteins binding to the AREs may destabilize the messenger, whereas members of the ELAV family, such as HuR, may increase the stability of mRNA. HuR may bind to AREs of all the three classes.

Engineering the HuR specific binding sites into the 3′ UTR of nucleic acid molecules can lead to HuR binding and thus, stabilization of the message in vivo.

[00151] Engineering of 3′ UTR AU rich elements (AREs) can be used to modulate the stability of an engineered polyribonucleotide. One or more copies of an ARE can be engineered into a polyribonucleotide to modulate the stability of a polyribonucleotide. AREs can be identified, removed or mutated to increase the intracellular stability and thus increase translation and production of the resultant protein. Transfection experiments can be conducted in relevant cell lines, using engineered polyribonucleotides and protein production can be assayed at various time points post-transfection. For example, cells can be transfected with different ARE-engineering molecules and by using an ELISA kit to the relevant protein and assaying protein produced at 6 hours, 12 hours, 24 hours, 48 hours, and 7 days post-transfection.

[00152] An untranslated region can comprise any number of nucleotides. An untranslated region can comprise a length of about 1 to about 10 bases or base pairs, about 10 to about 20 bases or base pairs, about 20 to about 50 bases or base pairs, about 50 to about 100 bases or base pairs, about 100 to about 500 bases or base pairs, about 500 to about 1000 bases or base pairs, about 1000 to about 2000 bases or base pairs, about 2000 to about 3000 bases or base pairs, about 3000 to about 4000 bases or base pairs, about 4000 to about 5000 bases or base pairs, about 5000 to about 6000 bases or base pairs, about 6000 to about 7000 bases or base pairs, about 7000 to about 8000 bases or base pairs, about 8000 to about 9000 bases or base pairs, or about 9000 to about 10000 bases or base pairs in length. An untranslated region can comprise a length of for example, at least 1 base or base pair, 2 bases or base pairs, 3 bases or base pairs, 4 bases or base pairs, 5 bases or base pairs, 6 bases or base pairs, 7 bases or base pairs, 8 bases or base pairs, 9 bases or base pairs, 10 bases or base pairs, 20 bases or base pairs, 30 bases or base pairs, 40 bases or base pairs, 50 bases or base pairs, 60 bases or base pairs, 70 bases or base pairs, 80 bases or base pairs, 90 bases or base pairs, 100 bases or base pairs, 200 bases or base pairs, 300 bases or base pairs, 400 bases or base pairs, 500 bases or base pairs, 600 bases or base pairs, 700 bases or base pairs, 800 bases or base pairs, 900 bases or base pairs, 1000 bases or base pairs, 2000 bases or base pairs, 3000 bases or base pairs, 4000 bases or base pairs, 5000 bases or base pairs, 6000 bases or base pairs, 7000 bases or base pairs, 8000 bases or base pairs, 9000 bases or base pairs, or 10000 bases or base pairs in length.

[00153] An engineered polyribonucleotide of the disclosure can comprise one or more introns. An intron can comprise any number of modified or unmodified nucleotides. An intron can comprise, for example, at least 1 base or base pair, 50 bases or base pairs, 100 bases or base pairs, 150 bases or base pairs, 200 bases or base pairs, 300 bases or base pairs, 400 bases or base pairs, 500 bases or base pairs, 600 bases or base pairs, 700 bases or base pairs, 800 bases or base pairs, 900 bases or base pairs, 1000 bases or base pairs, 2000 bases or base pairs, 3000 bases or base pairs, 4000 bases or base pairs, or 5000 bases or base pairs. In some cases, an intron can comprise, for example, at most 10000 bases or base pairs, 5000 bases or base pairs, 4000 bases or base pairs, 3000 bases or base pairs, 2000 bases or base pairs, 1000 bases or base pairs, 900 bases or base pairs, 800 bases or base pairs, 700 bases or base pairs, 600 bases or base pairs, 500 bases

or base pairs, 400 bases or base pairs, 300 bases or base pairs, 200 bases or base pairs, or 100 bases or base pairs.

[00154] In some cases, a percentage of the nucleotides in an intron are modified. For instance, in some cases, fewer than 99%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 5% or 1% of the nucleotides in an intron are modified. In some cases, all of the nucleotides in an intron are modified.

[00155] An engineered polyribonucleotide of the disclosure can comprise a polyA sequence. A polyA sequence (e.g., polyA tail) can comprise any number of nucleotides. A polyA sequence can comprise a length of about 1 to about 10 bases or base pairs, about 10 to about 20 bases or base pairs, about 20 to about 50 bases or base pairs, about 50 to about 100 bases or base pairs, about 100 to about 500 bases or base pairs, about 500 to about 1000 bases or base pairs, about 1000 to about 2000 bases or base pairs, about 2000 to about 3000 bases or base pairs, about 3000 to about 4000 bases or base pairs, about 4000 to about 5000 bases or base pairs, about 5000 to about 6000 bases or base pairs, about 6000 to about 7000 bases or base pairs, about 7000 to about 8000 bases or base pairs, about 8000 to about 9000 bases or base pairs, or about 9000 to about 10000 bases or base pairs in length. In some examples, a polyA sequence is at least about 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, or 200 nucleotides in length. A polyA sequence can comprise a length of for example, at least 1 base or base pair, 2 bases or base pairs, 3 bases or base pairs, 4 bases or base pairs, 5 bases or base pairs, 6 bases or base pairs, 7 bases or base pairs, 8 bases or base pairs, 9 bases or base pairs, 10 bases or base pairs, 20 bases or base pairs, 30 bases or base pairs, 40 bases or base pairs, 50 bases or base pairs, 60 bases or base pairs, 70 bases or base pairs, 80 bases or base pairs, 90 bases or base pairs, 100 bases or base pairs, 200 bases or base pairs, 300 bases or base pairs, 400 bases or base pairs, 500 bases or base pairs, 600 bases or base pairs, 700 bases or base pairs, 800 bases or base pairs, 900 bases or base pairs, 1000 bases or base pairs, 2000 bases or base pairs, 3000 bases or base pairs, 4000 bases or base pairs, 5000 bases or base pairs, 6000 bases or base pairs, 7000 bases or base pairs, 8000 bases or base pairs, 9000 bases or base pairs, or 10000 bases or base pairs in length. A polyA sequence can comprise a length of at most 100 bases or base pairs, 90 bases or base pairs, 80 bases or base pairs, 70 bases or base pairs, 60 bases or base pairs, 50 bases or base pairs, 40 bases or base pairs, 30 bases or base pairs, 20 bases or base pairs, 10 bases or base pairs, or 5 bases or base pairs.

[00156] In some cases, a percentage of the nucleotides in a poly-A sequence are modified. For instance, in some cases, fewer than 99%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 5% or 1% of the nucleotides in a poly-A sequence are modified. In some cases, all of the nucleotides in a poly-A are modified.

[00157] A linker sequence can comprise any number of nucleotides. A linker can be attached to the modified nucleobase at an N-3 or C-5 position. The linker attached to the nucleobase can be diethylene glycol, dipropylene glycol, triethylene glycol, tripropylene glycol, tetraethylene glycol, tetraethylene glycol, divalent alkyl, alkenyl, alkynyl moiety, ester, amide, or an ether moiety. A linker sequence can comprise a length of about 1 to about 10 bases or base pairs, about 10 to about 20 bases or base pairs, about 20 to about

50 bases or base pairs, about 50 to about 100 bases or base pairs, about 100 to about 500 bases or base pairs, about 500 to about 1000 bases or base pairs, about 1000 to about 2000 bases or base pairs, about 2000 to about 3000 bases or base pairs, about 3000 to about 4000 bases or base pairs, about 4000 to about 5000 bases or base pairs, about 5000 to about 6000 bases or base pairs, about 6000 to about 7000 bases or base pairs, about 7000 to about 8000 bases or base pairs, about 8000 to about 9000 bases or base pairs, or about 9000 to about 10000 bases or base pairs in length. A linker sequence can comprise a length of for example, at least 1 base or base pair, 2 bases or base pairs, 3 bases or base pairs, 4 bases or base pairs, 5 bases or base pairs, 6 bases or base pairs, 7 bases or base pairs, 8 bases or base pairs, 9 bases or base pairs, 10 bases or base pairs, 20 bases or base pairs, 30 bases or base pairs, 40 bases or base pairs, 50 bases or base pairs, 60 bases or base pairs, 70 bases or base pairs, 80 bases or base pairs, 90 bases or base pairs, 100 bases or base pairs, 200 bases or base pairs, 300 bases or base pairs, 400 bases or base pairs, 500 bases or base pairs, 600 bases or base pairs, 700 bases or base pairs, 800 bases or base pairs, 900 bases or base pairs, 1000 bases or base pairs, 2000 bases or base pairs, 3000 bases or base pairs, 4000 bases or base pairs, 5000 bases or base pairs, 6000 bases or base pairs, 7000 bases or base pairs, 8000 bases or base pairs, 9000 bases or base pairs, or at least 10000 bases or base pairs in length. A linker at most 10000 bases or base pairs, 5000 bases or base pairs, 4000 bases or base pairs, 3000 bases or base pairs, 2000 bases or base pairs, 1000 bases or base pairs, 900 bases or base pairs, 800 bases or base pairs, 700 bases or base pairs, 600 bases or base pairs, 500 bases or base pairs, 400 bases or base pairs, 300 bases or base pairs, 200 bases or base pairs, or 100 bases or base pairs in length.

[00158] In some cases, a percentage of the nucleotides in a linker sequence are modified. For instance, in some cases, fewer than 99%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 5% or 1% of the nucleotides in a linker sequence are modified. In some cases, all of the nucleotides in a linker sequence are modified.

[00159] In some cases, a nucleic acid construct(s), a vector(s), or an engineered polyribonucleotide(s) can include at least one stop codon before the 3′untranslated region (UTR). In some cases, a nucleic acid construct(s), a vector(s), or an engineered polyribonucleotide(s) includes multiple stop codons. The stop codon can be selected from TGA, TAA and TAG. The stop codon may be modified or unmodified. In some cases, the nucleic acid construct(s), vector(s), or engineered polyribonucleotide(s) includes the stop codon TGA and one additional stop codon. In some cases, the nucleic acid construct(s), vector(s), or engineered polyribonucleotide(s) includes the addition of the TAA stop codon.

LIPID COMPOSITIONS

[00160] In some embodiments of the (e.g., pharmaceutical) composition, the lipid composition comprises: (1) an ionizable cationic lipid; and (2) a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid. In some embodiments, the (e.g., pharmaceutical) composition further comprises a zwitterionic lipid (e.g. a phospholipid).

Ionizable Cationic Lipids

[00161] In some embodiments of the lipid composition of the present application, the lipid composition comprises an ionizable cationic lipid. In some embodiments, the cationic ionizable lipids contain one or more groups which is protonated at physiological pH but may deprotonated and has no charge at a pH above 8, 9, 10, 11, or 12. The ionizable cationic group may contain one or more protonatable amines which are able to form a cationic group at physiological pH. The cationic ionizable lipid compound may also further comprise one or more lipid components such as two or more fatty acids with C_6-C_{24} alkyl or alkenyl carbon groups. These lipid groups may be attached through an ester linkage or may be further added through a Michael addition to a sulfur atom. In some embodiments, these compounds may be a dendrimer, a dendron, a polymer, or a combination thereof.

[00162] In some embodiments of the lipid composition of the present application, the ionizable cationic lipids refer to lipid and lipid-like molecules with nitrogen atoms that can acquire charge (pKa). These lipids may be known in the literature as cationic lipids. These molecules with amino groups typically have between 2 and 6 hydrophobic chains, often alkyl or alkenyl such as C_6-C_{24} alkyl or alkenyl groups, but may have at least 1 or more that 6 tails. In some embodiments, these cationic ionizable lipids are dendrimers, which are a polymer exhibiting regular dendritic branching, formed by the sequential or generational addition of branched layers to or from a core and are characterized by a core, at least one interior branched layer, and a surface branched layer. (See *Petar R. Dvornic and Donald A. Tomalia in Chem. in Britain, 641-645, August 1994*.) In other embodiments, the term "dendrimer" as used herein is intended to include, but is not limited to, a molecular architecture with an interior core, interior layers (or "generations") of repeating units regularly attached to this initiator core, and an exterior surface of terminal groups attached to the outermost generation. A "dendron" is a species of dendrimer having branches emanating from a focal point which is or can be joined to a core, either directly or through a linking moiety to form a larger dendrimer. In some embodiments, the dendrimer structures have radiating repeating groups from a central core which doubles with each repeating unit for each branch. In some embodiments, the dendrimers described herein may be described as a small molecule, medium-sized molecules, lipids, or lipid-like material. These terms may be used to described compounds described herein which have a dendron like appearance (*e.g.* molecules which radiate from a single focal point).

[00163] While dendrimers are polymers, dendrimers may be preferable to traditional polymers because they have a controllable structure, a single molecular weight, numerous and controllable surface functionalities, and traditionally adopt a globular conformation after reaching a specific generation. Dendrimers can be prepared by sequentially reactions of each repeating unit to produce monodisperse, tree-like and/or generational structure polymeric structures. Individual dendrimers consist of a central core molecule, with a dendritic wedge attached to one or more functional sites on that central core. The dendrimeric surface layer can have a variety of functional groups disposed thereon including anionic, cationic, hydrophilic, or lipophilic groups, according to the assembly monomers used during the preparation.

[00164] Modifying the functional groups and/or the chemical properties of the core, repeating units, and the surface or terminating groups, their physical properties can be modulated. Some properties which can be varied include, but are not limited to, solubility, toxicity, immunogenicity and bioattachment capability. Dendrimers are often described by their generation or number of repeating units in the branches. A dendrimer consisting of only the core molecule is referred to as Generation 0, while each consecutive repeating unit along all branches is Generation 1, Generation 2, and so on until the terminating or surface group. In some embodiments, half generations are possible resulting from only the first condensation reaction with the amine and not the second condensation reaction with the thiol.

[00165] Preparation of dendrimers requires a level of synthetic control achieved through series of stepwise reactions comprising building the dendrimer by each consecutive group. Dendrimer synthesis can be of the convergent or divergent type. During divergent dendrimer synthesis, the molecule is assembled from the core to the periphery in a stepwise process involving attaching one generation to the previous and then changing functional groups for the next stage of reaction. Functional group transformation is necessary to prevent uncontrolled polymerization. Such polymerization would lead to a highly branched molecule that is not monodisperse and is otherwise known as a hyperbranched polymer. Due to steric effects, continuing to react dendrimer repeat units leads to a sphere shaped or globular molecule, until steric overcrowding prevents complete reaction at a specific generation and destroys the molecule's monodispersity. Thus, in some embodiments, the dendrimers of G1-G10 generation are specifically contemplated. In some embodiments, the dendrimers comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 repeating units, or any range derivable therein. In some embodiments, the dendrimers used herein are G0, G1, G2, or G3. However, the number of possible generations (such as 11, 12, 13, 14, 15, 20, or 25) may be increased by reducing the spacing units in the branching polymer.

[00166] Additionally, dendrimers have two major chemical environments: the environment created by the specific surface groups on the termination generation and the interior of the dendritic structure which due to the higher order structure can be shielded from the bulk media and the surface groups. Because of these different chemical environments, dendrimers have found numerous different potential uses including in therapeutic applications.

[00167] In some embodiments of the lipid composition of the present application, the dendrimers are assembled using the differential reactivity of the acrylate and methacrylate groups with amines and thiols. The dendrimers may include secondary or tertiary amines and thioethers formed by the reaction of an acrylate group with a primary or secondary amine and a methacrylate with a mercapto group. Additionally, the repeating units of the dendrimers may contain groups which are degradable under physiological conditions. In some embodiments, these repeating units may contain one or more germinal diethers, esters, amides, or disulfides groups. In some embodiments, the core molecule is a monoamine which allows dendritic polymerization in only one direction. In other embodiments, the core molecule is a polyamine with multiple different dendritic branches which each may comprise one or more repeating units. The dendrimer may be formed by removing one or more hydrogen atoms from this core. In some embodiments,

these hydrogen atoms are on a heteroatom such as a nitrogen atom. In some embodiments, the terminating group is a lipophilic groups such as a long chain alkyl or alkenyl group. In other embodiments, the terminating group is a long chain haloalkyl or haloalkenyl group. In other embodiments, the terminating group is an aliphatic or aromatic group containing an ionizable group such as an amine ($-NH_2$) or a carboxylic acid ($-CO_2H$). In still other embodiments, the terminating group is an aliphatic or aromatic group containing one or more hydrogen bond donors such as a hydroxide group, an amide group, or an ester.

[00168] The cationic ionizable lipids of the present application may contain one or more asymmetrically-substituted carbon or nitrogen atoms, and may be isolated in optically active or racemic form. Thus, all chiral, diastereomeric, racemic form, epimeric form, and all geometric isomeric forms of a chemical formula are intended, unless the specific stereochemistry or isomeric form is specifically indicated. Cationic ionizable lipids may occur as racemates and racemic mixtures, single enantiomers, diastereomeric mixtures and individual diastereomers. In some embodiments, a single diastereomer is obtained. The chiral centers of the cationic ionizable lipids of the present application can have the *S* or the *R* configuration. Furthermore, it is contemplated that one or more of the cationic ionizable lipids may be present as constitutional isomers. In some embodiments, the compounds have the same formula but different connectivity to the nitrogen atoms of the core. Without wishing to be bound by any theory, it is believed that such cationic ionizable lipids exist because the starting monomers react first with the primary amines and then statistically with any secondary amines present. Thus, the constitutional isomers may present the fully reacted primary amines and then a mixture of reacted secondary amines.

[00169] Chemical formulas used to represent cationic ionizable lipids of the present application will typically only show one of possibly several different tautomers. For example, many types of ketone groups are known to exist in equilibrium with corresponding enol groups. Similarly, many types of imine groups exist in equilibrium with enamine groups. Regardless of which tautomer is depicted for a given formula, and regardless of which one is most prevalent, all tautomers of a given chemical formula are intended.

[00170] The cationic ionizable lipids of the present application may also have the advantage that they may be more efficacious than, be less toxic than, be longer acting than, be more potent than, produce fewer side effects than, be more easily absorbed than, and/or have a better pharmacokinetic profile (*e.g.*, higher oral bioavailability and/or lower clearance) than, and/or have other useful pharmacological, physical, or chemical properties over, compounds known in the prior art, whether for use in the indications stated herein or otherwise.

[00171] In addition, atoms making up the cationic ionizable lipids of the present application are intended to include all isotopic forms of such atoms. Isotopes, as used herein, include those atoms having the same atomic number but different mass numbers. By way of general example and without limitation, isotopes of hydrogen include tritium and deuterium, and isotopes of carbon include ^{13}C and ^{14}C.

[00172] It should be recognized that the particular anion or cation forming a part of any salt form of a cationic ionizable lipids provided herein is not critical, so long as the salt, as a whole, is pharmacologically

acceptable. Additional examples of pharmaceutically acceptable salts and their methods of preparation and use are presented in *Handbook of Pharmaceutical Salts: Properties, and Use* (2002), which is incorporated herein by reference.

[00173] In some embodiments of the lipid composition of the present application, the ionizable cationic lipid is a dendrimer or dendron. In some embodiments, the ionizable cationic lipid comprises an ammonium group which is positively charged at physiological pH and contains at least two hydrophobic groups. In some embodiments, the ammonium group is positively charged at a pH from about 6 to about 8. In some embodiments, the ionizable cationic lipid is a dendrimer or dendron. In some embodiments, the ionizable cationic lipid comprises at least two C_6-C_{24} alkyl or alkenyl groups.

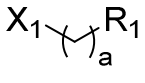
Dendrimers of Formula (I)

[00174] In some embodiments of the lipid composition, the ionizable cationic lipid comprises at least two C_8-C_{24} alkyl groups. In some embodiments, the ionizable cationic lipid is a dendrimer further defined by the formula:

Core-Repeating Unit-Terminating Group (D-I)

wherein the core is linked to the repeating unit by removing one or more hydrogen atoms from the core and replacing the atom with the repeating unit and wherein:

the core has the formula:


(D-II)

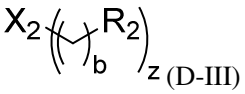
wherein:

X_1 is amino or alkylamino$_{(C\leq12)}$, dialkylamino$_{(C\leq12)}$, heterocycloalkyl$_{(C\leq12)}$, heteroaryl$_{(C\leq12)}$, or a substituted version thereof;

R_1 is amino, hydroxy, or mercapto, or alkylamino$_{(C\leq12)}$, dialkylamino$_{(C\leq12)}$, or a substituted version of either of these groups; and

a is 1, 2, 3, 4, 5, or 6; or

the core has the formula:


(D-III)

wherein:

X_2 is $N(R_5)_y$;

R_5 is hydrogen, alkyl$_{(C\leq18)}$, or substituted alkyl$_{(C\leq18)}$; and

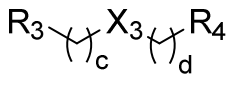
y is 0, 1, or 2, provided that the sum of y and z is 3;

R_2 is amino, hydroxy, or mercapto, or alkylamino$_{(C\leq12)}$, dialkylamino$_{(C\leq12)}$, or a substituted version of either of these groups;

b is 1, 2, 3, 4, 5, or 6; and

z is 1, 2, 3; provided that the sum of z and y is 3; or

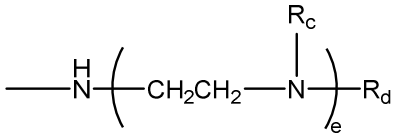
the core has the formula:

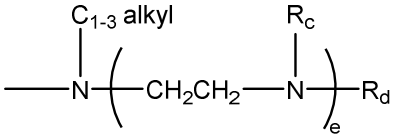
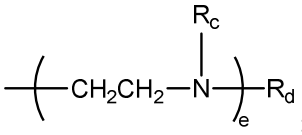


(D-IV)

wherein:

X_3 is $-NR_6-$, wherein R_6 is hydrogen, $alkyl_{(C\leq 8)}$, or substituted $alkyl_{(C\leq 8)}$, $-O-$, or $alkylaminodiyl_{(C\leq 8)}$, $alkoxydiyl_{(C\leq 8)}$, $arenediyl_{(C\leq 8)}$, $heteroarenediyl_{(C\leq 8)}$, $heterocycloalkanediyl_{(C\leq 8)}$, or a substituted version of any of these groups;

R_3 and R_4 are each independently amino, hydroxy, or mercapto, or $alkylamino_{(C\leq 12)}$, $dialkylamino_{(C\leq 12)}$, or a substituted version of either of these groups; or a group of the formula: $-N(R_f)_f(CH_2CH_2N(R_c))_eR_d$, ,

, or ;

wherein:

e and f are each independently 1, 2, or 3; provided that the sum of e and f is 3;

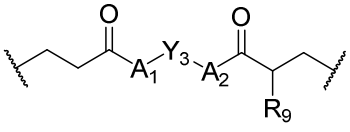
R_c, R_d, and R_f are each independently hydrogen, $alkyl_{(C\leq 6)}$, or substituted $alkyl_{(C\leq 6)}$;

c and d are each independently 1, 2, 3, 4, 5, or 6; or

the core is $alkylamine_{(C\leq 18)}$, $dialkylamine_{(C\leq 36)}$, $heterocycloalkane_{(C\leq 12)}$, or a substituted version of any of these groups;

wherein the repeating unit comprises a degradable diacyl and a linker;

the degradable diacyl group has the formula:



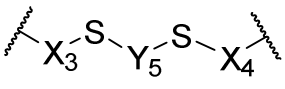
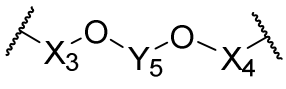
(D-VII)

wherein:

A_1 and A_2 are each independently $-O-$, -S-, or $-NR_a-$, wherein:

R_a is hydrogen, $alkyl_{(C\leq 6)}$, or substituted $alkyl_{(C\leq 6)}$;

Y_3 is $alkanediyl_{(C\leq 12)}$, $alkenediyl_{(C\leq 12)}$, $arenediyl_{(C\leq 12)}$, or a substituted version of any of these groups; or a group of the formula:

 or 

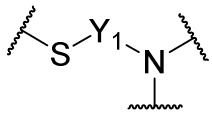
wherein:

X_3 and X_4 are alkanediyl$_{(C\leq12)}$, alkenediyl$_{(C\leq12)}$, arenediyl$_{(C\leq12)}$, or a substituted version of any of these groups;

Y_5 is a covalent bond, alkanediyl$_{(C\leq12)}$, alkenediyl$_{(C\leq12)}$, arenediyl$_{(C\leq12)}$, or a substituted version of any of these groups; and

R_9 is alkyl$_{(C\leq8)}$ or substituted alkyl$_{(C\leq8)}$;

the linker group has the formula:



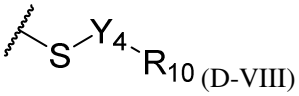
(D-VI)

wherein:

Y_1 is alkanediyl$_{(C\leq12)}$, alkenediyl$_{(C\leq12)}$, arenediyl$_{(C\leq12)}$, or a substituted version of any of these groups; and

wherein when the repeating unit comprises a linker group, then the linker group comprises an independent degradable diacyl group attached to both the nitrogen and the sulfur atoms of the linker group if n is greater than 1, wherein the first group in the repeating unit is a degradable diacyl group, wherein for each linker group, the next repeating unit comprises two degradable diacyl groups attached to the nitrogen atom of the linker group; and wherein n is the number of linker groups present in the repeating unit; and

the terminating group has the formula:



(D-VIII)

wherein:

Y_4 is alkanediyl$_{(C\leq18)}$ or an alkanediyl$_{(C\leq18)}$ wherein one or more of the hydrogen atoms on the alkanediyl$_{(C\leq18)}$ has been replaced with –OH, –F, –Cl, –Br, –I, –SH, –OCH_3, –OCH_2CH_3, –SCH_3, or –$OC(O)CH_3$;

R_{10} is hydrogen, carboxy, hydroxy, or

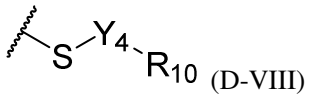
aryl$_{(C\leq12)}$, alkylamino$_{(C\leq12)}$, dialkylamino$_{(C\leq12)}$, *N*-heterocycloalkyl$_{(C\leq12)}$, –C(O)N(R_{11})–alkanediyl$_{(C\leq6)}$–heterocycloalkyl$_{(C\leq12)}$, –C(O)–alkylamino$_{(C\leq12)}$, –C(O)–dialkylamino$_{(C\leq12)}$, –C(O)–*N*-heterocycloalkyl$_{(C\leq12)}$, wherein:

R_{11} is hydrogen, alkyl$_{(C\leq6)}$, or substituted alkyl$_{(C\leq6)}$;

wherein the final degradable diacyl in the chain is attached to a terminating group;

n is 0, 1, 2, 3, 4, 5, or 6;

or a pharmaceutically acceptable salt thereof. In some embodiments, the terminating group is further defined by the formula:



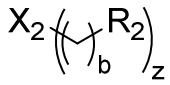
(D-VIII)

wherein:

Y_4 is alkanediyl$_{(C\leq 18)}$; and

R_{10} is hydrogen. In some embodiments, A_1 and A_2 are each independently –O– or –NR_a–.

[00175] In some embodiments of the dendrimer of formula (D-I), the core is further defined by the formula:



(D-III)

wherein:

X_2 is $N(R_5)_y$;

R_5 is hydrogen or alkyl$_{(C\leq 8)}$, or substituted alkyl$_{(C\leq 18)}$; and

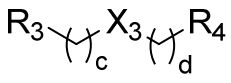
y is 0, 1, or 2, provided that the sum of y and z is 3;

R_2 is amino, hydroxy, or mercapto, or alkylamino$_{(C\leq 12)}$, dialkylamino$_{(C\leq 12)}$, or a substituted version of either of these groups;

b is 1, 2, 3, 4, 5, or 6; and

z is 1, 2, 3; provided that the sum of z and y is 3.

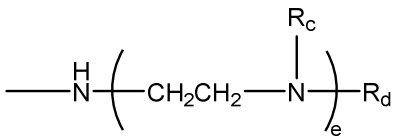
[00176] In some embodiments of the dendrimer of formula (D-I), the core is further defined by the formula:

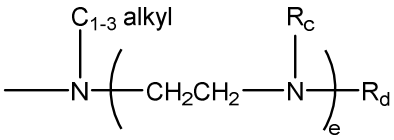
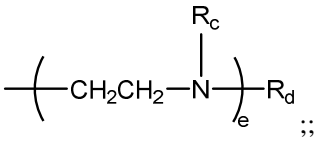


(D-IV)

wherein:

X_3 is –NR_6–, wherein R_6 is hydrogen, alkyl$_{(C\leq 8)}$, or substituted alkyl$_{(C\leq 8)}$, –O–, or alkylaminodiyl$_{(C\leq 8)}$, alkoxydiyl$_{(C\leq 8)}$, arenediyl$_{(C\leq 8)}$, heteroarenediyl$_{(C\leq 8)}$, heterocycloalkanediyl$_{(C\leq 8)}$, or a substituted version of any of these groups;

R_3 and R_4 are each independently amino, hydroxy, or mercapto, or alkylamino$_{(C\leq 12)}$, dialkylamino$_{(C\leq 12)}$, or a substituted version of either of these groups; or a group of the formula: –$N(R_f)_f(CH_2CH_2N(R_c))_eR_d$, ,

, or ;;

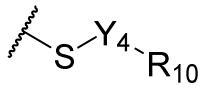
wherein:

e and f are each independently 1, 2, or 3; provided that the sum of e and f is 3;

R_c, R_d, and R_f are each independently hydrogen, alkyl$_{(C\leq 6)}$, or substituted alkyl$_{(C\leq 6)}$;

c and d are each independently 1, 2, 3, 4, 5, or 6.

[00177] In some embodiments of the dendrimer of formula (I), the the terminating group is represented by the formula:



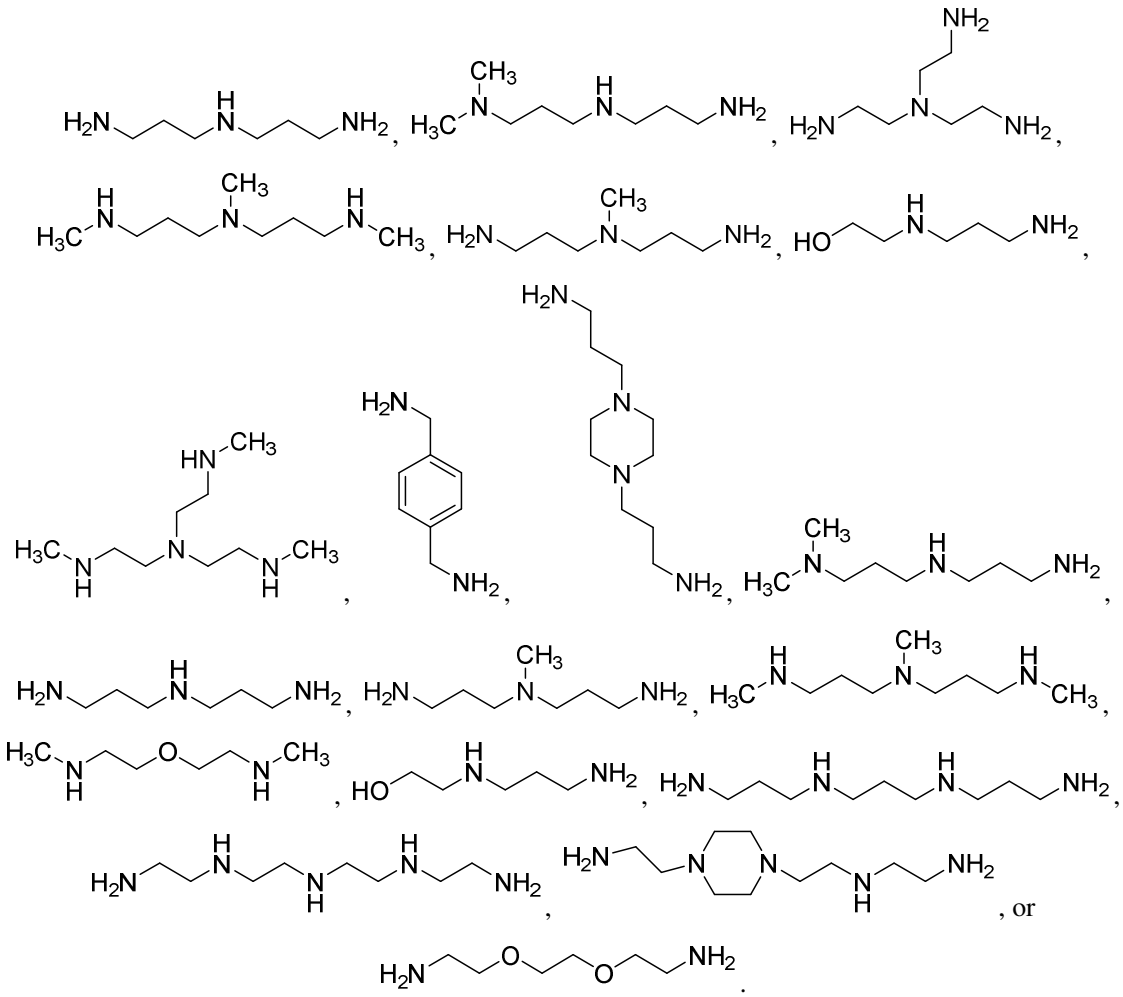
(D-VIII),

wherein:

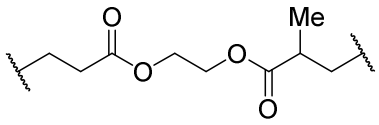
Y_4 is alkanediyl$_{(C\leq 18)}$; and

R_{10} is hydrogen.

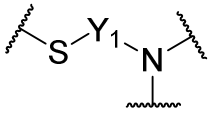
[00178] In some embodiments of the dendrimer of formula (D-I), the core is further defined as:



[00179] In some embodiments of the dendrimer of formula (D-I), the degradable diacyl is further defined as:



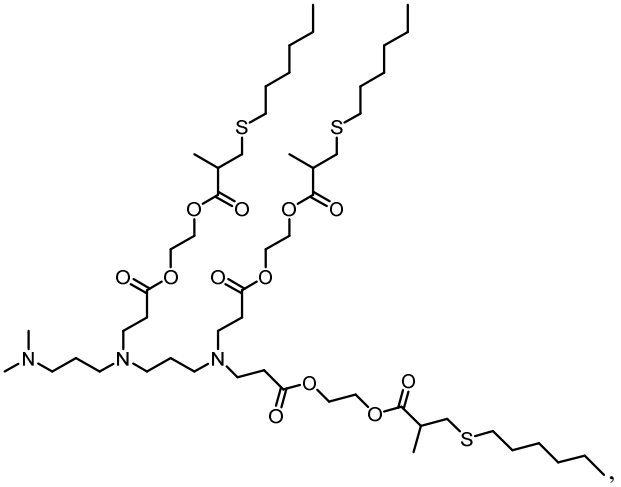
[00180] In some embodiments of the dendrimer of formula (D-I), the linker is further defined as



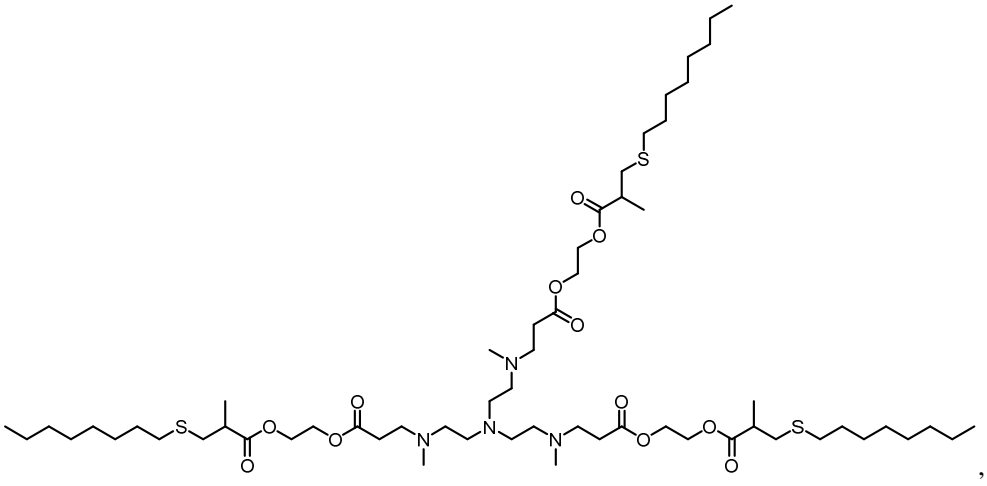
(D-VI),

wherein Y_1 is alkanediyl$_{(C\leq 8)}$ or substituted alkanediyl$_{(C\leq 8)}$.

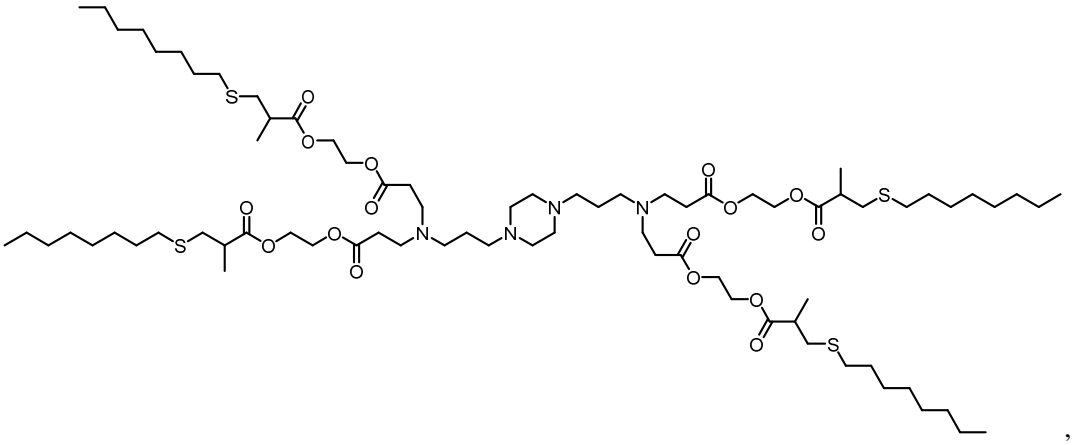
[00181] In some embodiments of the dendrimer of formula (D-I), the dendrimer is selected from the group consisting of:



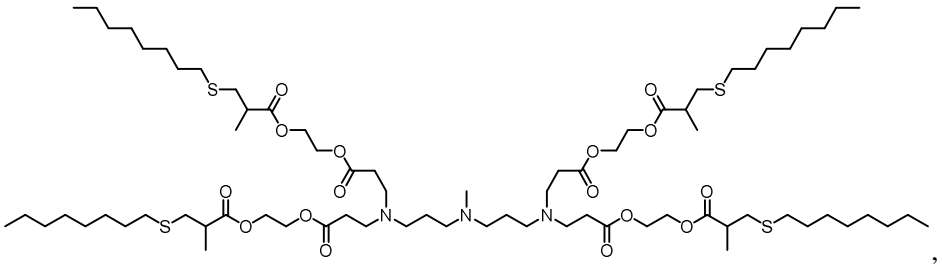
,



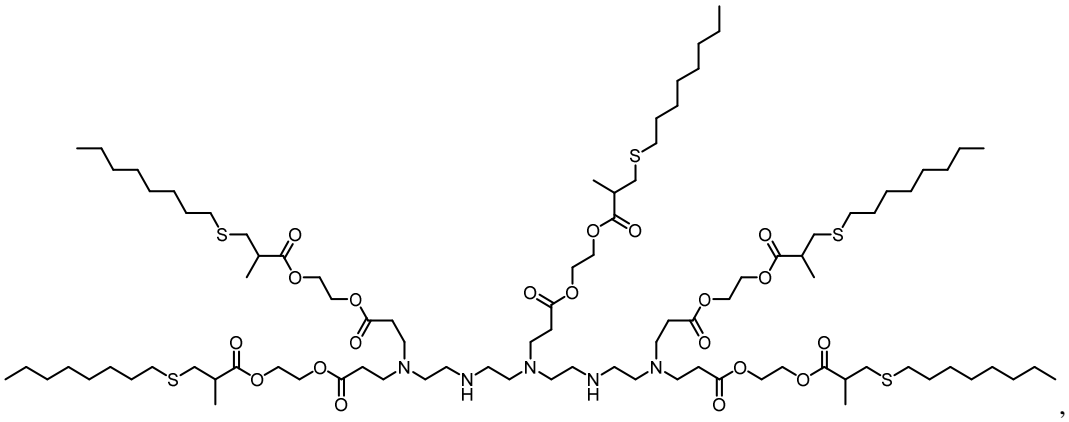
,



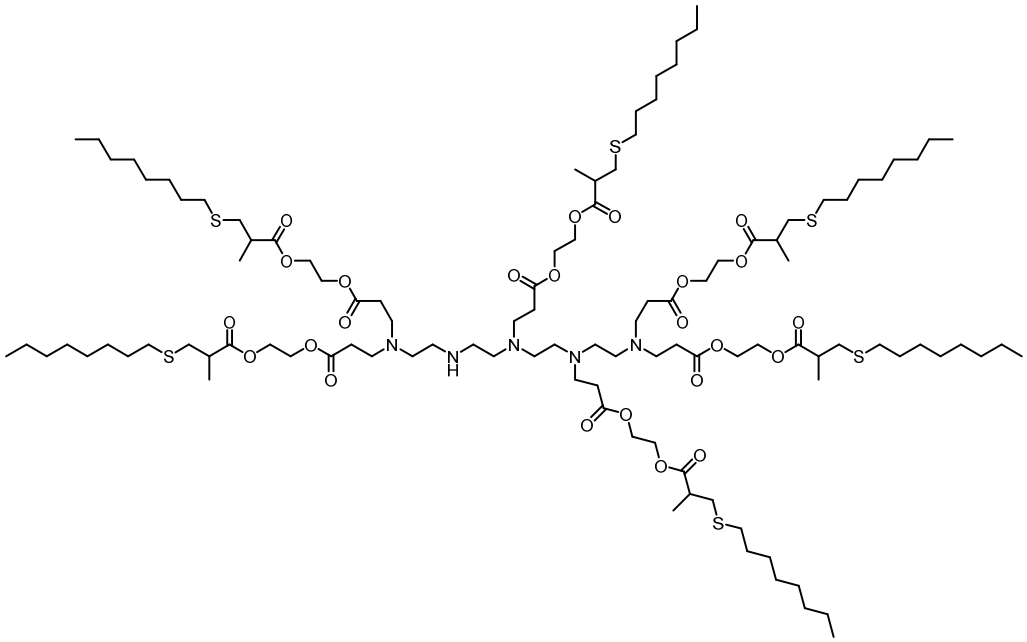
,



,



,

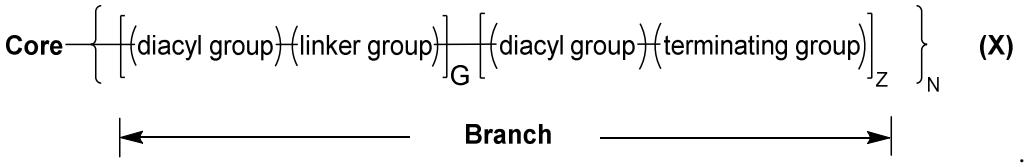


or ;

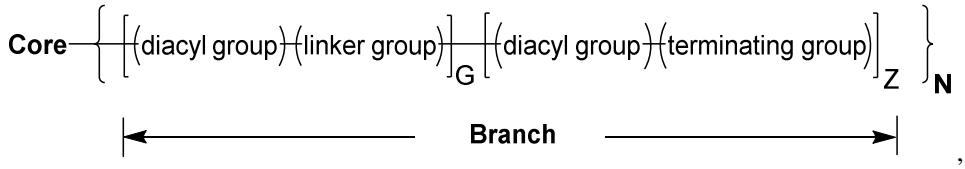
and pharmaceutically acceptable salts thereof.

Dendrimers of Formula (X)

[00182] In some embodiments of the lipid composition, the ionizable cationic lipid is a dendrimer of the formula $\mathbf{Core}\text{-}\left(\mathbf{Branch}\right)_{\mathbf{N}}$. In some embodiments, the ionizable cationic lipid is a dendrimer of the formula

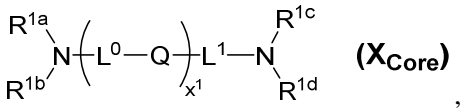


[00183] In some embodiments of the lipid composition, the ionizable cationic lipid is a dendrimer of a generation (g) having a structural formula:



or a pharmaceutically acceptable salt thereof, wherein:

(a) the core comprises a structural formula (X_{Core}):



wherein:

Q is independently at each occurrence a covalent bond, -O-, -S-, -NR^2-, or -$CR^{3a}R^{3b}$-;

R^2 is independently at each occurrence R^{1g} or -L^2-$NR^{1e}R^{1f}$;

R^{3a} and R^{3b} are each independently at each occurrence hydrogen or an optionally substituted (e.g., C_1-C_6, such as C_1-C_3) alkyl;

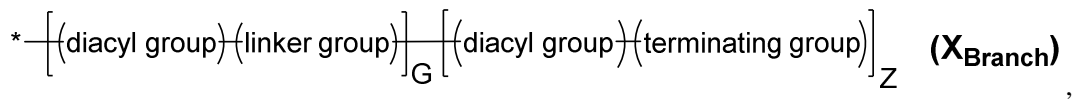
R^{1a}, R^{1b}, R^{1c}, R^{1d}, R^{1e}, R^{1f}, and R^{1g} (if present) are each independently at each occurrence a point of connection to a branch, hydrogen, or an optionally substituted (e.g., C_1-C_{12}) alkyl;

L^0, L^1, and L^2 are each independently at each occurrence selected from a covalent bond, alkylene, heteroalkylene, [alkylene]-[heterocycloalkyl]-[alkylene], [alkylene]-(arylene)-[alkylene], heterocycloalkyl, and arylene; or,

alternatively, part of L^1 form a (e.g., C_4-C_6) heterocycloalkyl (e.g., containing one or two nitrogen atoms and, optionally, an additional heteroatom selected from oxygen and sulfur) with one of R^{1c} and R^{1d}; and

x^1 is 0, 1, 2, 3, 4, 5, or 6; and

(b) each branch of the plurality (N) of branches independently comprises a structural formula (X_{Branch}):



wherein:

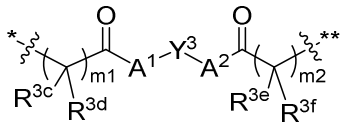
* indicates a point of attachment of the branch to the core;

g is 1, 2, 3, or 4;

$Z = 2^{(g-1)}$;

G=0, when g=1; or G = $\sum_{i=0}^{i=g-2} 2^i$, when g≠1;

(c) each diacyl group independently comprises a structural formula



, wherein:

* indicates a point of attachment of the diacyl group at the proximal end thereof;

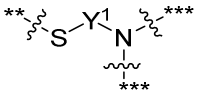
** indicates a point of attachment of the diacyl group at the distal end thereof;

Y^3 is independently at each occurrence an optionally substituted (e.g., C_1-C_{12}); alkylene, an optionally substituted (e.g., C_1-C_{12}) alkenylene, or an optionally substituted (e.g., C_1-C_{12}) arenylene;

A^1 and A^2 are each independently at each occurrence -O-, -S-, or -NR^4-, wherein: R^4 is hydrogen or optionally substituted (e.g., C_1-C_6) alkyl;

m^1 and m^2 are each independently at each occurrence 1, 2, or 3; and

R^{3c}, R^{3d}, R^{3e}, and R^{3f} are each independently at each occurrence hydrogen or an optionally substituted (e.g., C_1-C_8) alkyl; and

(d) each linker group independently comprises a structural formula 

, wherein:

** indicates a point of attachment of the linker to a proximal diacyl group;

*** indicates a point of attachment of the linker to a distal diacyl group; and

Y_1 is independently at each occurrence an optionally substituted (e.g., C_1-C_{12}) alkylene, an optionally substituted (e.g., C_1-C_{12}) alkenylene, or an optionally substituted (e.g., C_1-C_{12}) arenylene; and

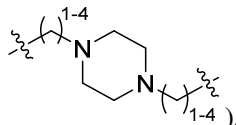
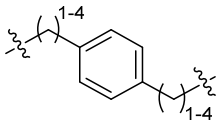
(e) each terminating group is independently selected from optionally substituted (e.g., C_1-C_{18}, such as C_4-C_{18}) alkylthiol, and optionally substituted (e.g., C_1-C_{18}, such as C_4-C_{18}) alkenylthiol.

[00184] In some embodiments of X_{Core}, Q is independently at each occurrence a covalent bond, -O-, -S-, -NR^2-, or -$CR^{3a}R^{3b}$. In some embodiments of X_{Core} Q is independently at each occurrence a covalent bond. In some embodiments of X_{Core} Q is independently at each occurrence an -O-. In some embodiments of X_{Core} Q is independently at each occurrence a -S-. In some embodiments of X_{Core} Q is independently at each occurrence a -NR^2 and R^2 is independently at each occurrence R^{1g} or -L^2-$NR^{1e}R^{1f}$. In some embodiments of X_{Core} Q is independently at each occurrence a -$CR^{3a}R^{3b}$ R^{3a}, and R^{3a} and R^{3b} are each independently at each occurrence hydrogen or an optionally substituted alkyl (e.g., C_1-C_6, such as C_1-C_3).

[00185] In some embodiments of X_{Core}, R^{1a}, R^{1b}, R^{1c}, R^{1d}, R^{1e}, R^{1f}, and R^{1g} (if present) are each independently at each occurrence a point of connection to a branch, hydrogen, or an optionally substituted alkyl. In some embodiments of X_{Core}, R^{1a}, R^{1b}, R^{1c}, R^{1d}, R^{1e}, R^{1f}, and R^{1g} (if present) are each independently

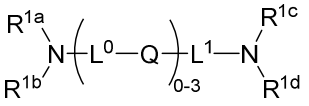
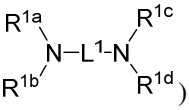
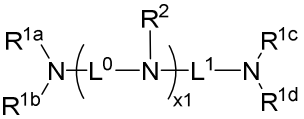
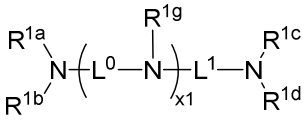
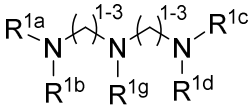
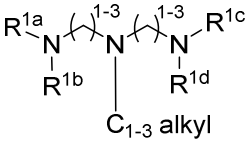
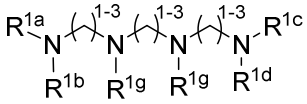
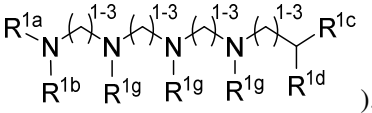
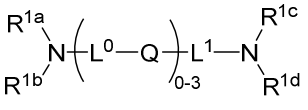
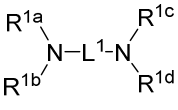
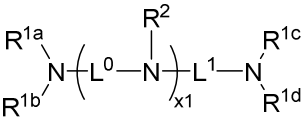
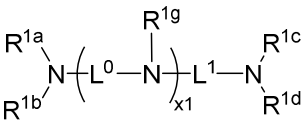
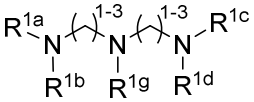
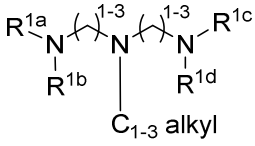
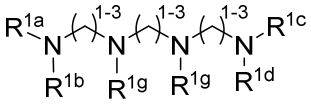
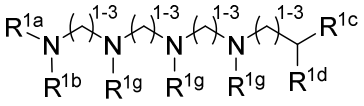
at each occurrence a point of connection to a branch, hydrogen. In some embodiments of X_{Core}, R^{1a}, R^{1b}, R^{1c}, R^{1d}, R^{1e}, R^{1f}, and R^{1g} (if present) are each independently at each occurrence a point of connection to a branch an optionally substituted alkyl (e.g., C_1-C_{12}).

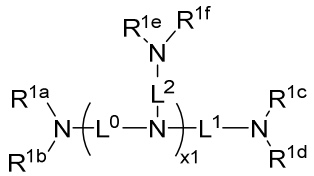
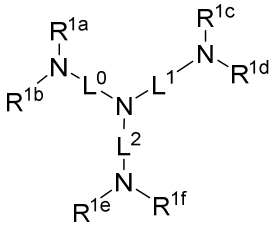
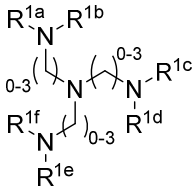
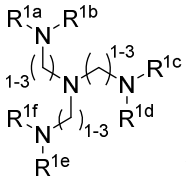
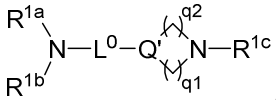
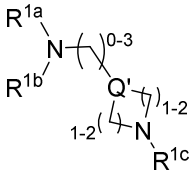
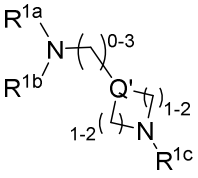
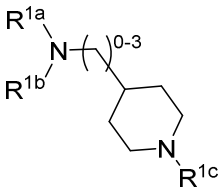
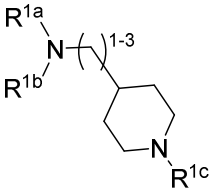
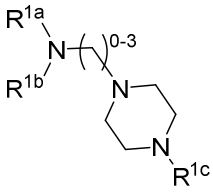
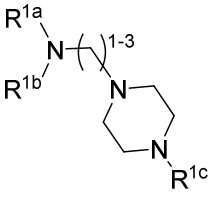
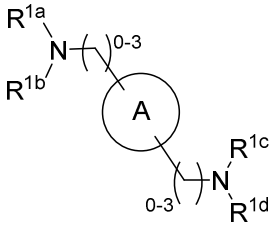
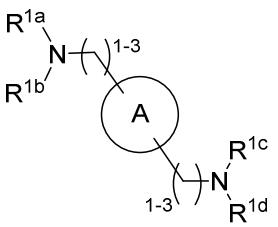
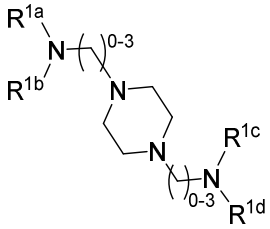
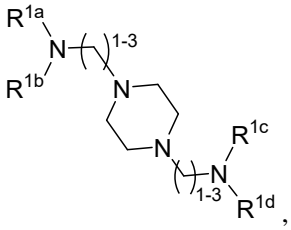
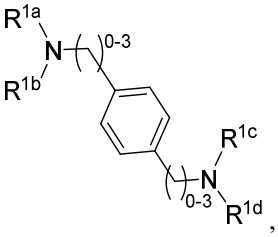
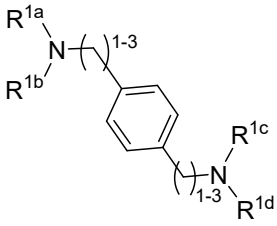
[00186] In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence selected from a covalent bond, alkylene, heteroalkylene, [alkylene]-[heterocycloalkyl]-[alkylene], [alkylene]-(arylene)-[alkylene], heterocycloalkyl, and arylene; or, alternatively, part of L^1 form a heterocycloalkyl (e.g., C_4-C_6 and containing one or two nitrogen atoms and, optionally, an additional heteroatom selected from oxygen and sulfur) with one of R^{1c} and R^{1d}. In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a covalent bond. In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a hydrogen. In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be an alkylene (e.g., C_1-C_{12}, such as C_1-C_6 or C_1-C_3). In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a heteroalkylene (e.g., C_1-C_{12}, such as C_1-C_8 or C_1-C_6). In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a heteroalkylene (e.g., C_2-C_8 alkyleneoxide, such as oligo(ethyleneoxide)). In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a [alkylene]-[heterocycloalkyl]-[alkylene] [(e.g., C_1-C_6) alkylene]-[(e.g., C_4-C_6) heterocycloalkyl]-[(e.g., C_1-C_6) alkylene]. In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a [alkylene]-(arylene)-[alkylene] [(e.g., C_1-C_6) alkylene]-(arylene)-[(e.g., C_1-C_6) alkylene]. In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a [alkylene]-(arylene)-[alkylene] (e.g., [(e.g., C_1-C_6) alkylene]-phenylene-[(e.g., C_1-C_6) alkylene]). In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be a heterocycloalkyl (e.g., C_4-C_6heterocycloalkyl). In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence can be an arylene (e.g., phenylene). In some embodiments of X_{Core}, part of L^1 form a heterocycloalkyl with one of R^{1c} and R^{1d}. In some embodiments of X_{Core}, part of L^1 form a heterocycloalkyl (e.g., C_4-C_6 heterocycloalkyl) with one of R^{1c} and R^{1d} and the heterocycloalkyl can contain one or two nitrogen atoms and, optionally, an additional heteroatom selected from oxygen and sulfur.

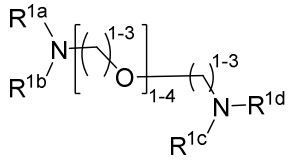
[00187] In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence selected from a covalent bond, C_1-C_6 alkylene (e.g., C_1-C_3 alkylene), C_2-C_{12} (e.g., C_2-C_8) alkyleneoxide (e.g., oligo(ethyleneoxide), such as -$(CH_2CH_2O)_{1\text{-}4}$-(CH_2CH_2)-), [(C_1-C_4) alkylene]-[(C_4-C_6) heterocycloalkyl]-[(C_1-C_4) alkylene] (e.g., ), and [(C_1-C_4) alkylene]-phenylene-[(C_1-C_4) alkylene] (e.g., ). In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence selected from C_1-C_6 alkylene (e.g., C_1-C_3 alkylene), -(C_1-C_3 alkylene-O)$_{1\text{-}4}$-(C_1-C_3 alkylene), -(C_1-C_3 alkylene)-phenylene-(C_1-C_3 alkylene)-, and -(C_1-C_3 alkylene)-piperazinyl-(C_1-C_3 alkylene)-. In

some embodiments of $X_{Core,}$ L^0, L^1, and L^2 are each independently at each occurrence C_1-C_6 alkylene (e.g., C_1-C_3 alkylene). In some embodiments, L^0, L^1, and L^2 are each independently at each occurrence C_2-C_{12} (e.g., C_2-C_8) alkyleneoxide (e.g., -(C_1-C_3 alkylene-O)$_{1-4}$-(C_1-C_3 alkylene)). In some embodiments of X_{Core}, L^0, L^1, and L^2 are each independently at each occurrence selected from [(C_1-C_4) alkylene]-[(C_4-C_6) heterocycloalkyl]-[(C_1-C_4) alkylene] (e.g., -(C_1-C_3 alkylene)-phenylene-(C_1-C_3 alkylene)-) and [(C_1-C_4) alkylene]-[(C_4-C_6) heterocycloalkyl]-[(C_1-C_4) alkylene] (e.g., -(C_1-C_3 alkylene)-piperazinyl-(C_1-C_3 alkylene)-).

[00188] In some embodiments of $X_{Core,}$ x^1 is 0, 1, 2, 3, 4, 5, or 6. In some embodiments of $X_{Core,}$ x^1 is 0. In some embodiments of X_{Core}, x^1 is 1. In some embodiments of X_{Core}, x^1 is 2. In some embodiments of $X_{Core,}$ x^1 is 0, 3. In some embodiments of X_{Core} x^1 is 4. In some embodiments of X_{Core} x^1 is 5. In some embodiments of $X_{Core,}$ x^1 is 6.

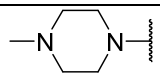
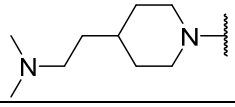
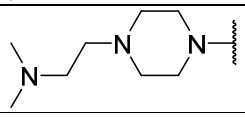
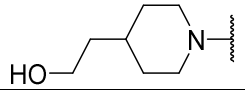
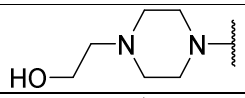
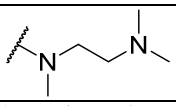
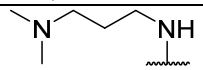
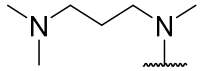
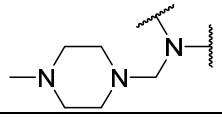
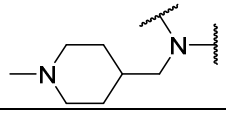
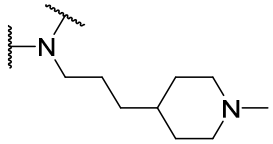
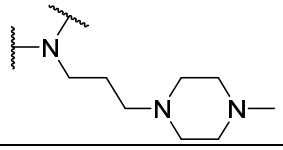
[00189] In some embodiments of X_{Core}, the core comprises a structural formula:  (e.g., ). In some embodiments of X_{Core}, the core comprises a structural formula: . In some embodiments of X_{Core}, the core comprises a structural formula:  (e.g., , , , or ). In some embodiments of X_{Core}, the core comprises a structural formula:  (e.g., ). In some embodiments of X_{Core}, the core comprises a structural formula: . In some embodiments of X_{Core}, the core comprises a structural formula:  (e.g., , , , or ). In some

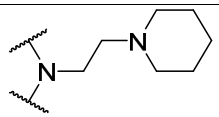
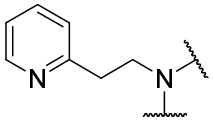
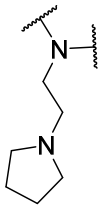
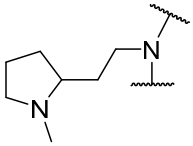
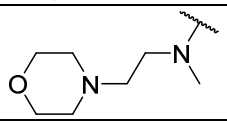
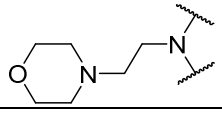
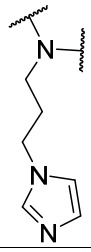
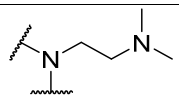
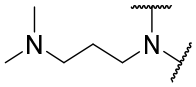
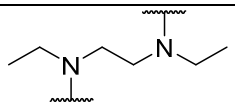
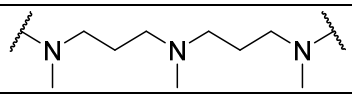
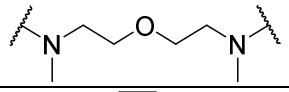
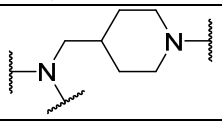
embodiments of X_{Core}, the core comprises a structural formula:  (e.g.,  , such as  or ). In some embodiments of X_{Core}, the core comprises a structural formula:  , wherein Q' is $-NR^2-$ or $-CR^{3a}R^{3b}-$; q^1 and q^2 are each independently 1 or 2. In some embodiments of X_{Core}, the core comprises a structural formula:  or  (e.g.,  ,  ,  , or ). In some embodiments of X_{Core}, the core comprises a structural formula  or  (e.g.,  ,  ,  , or ), wherein ring A is an

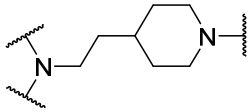
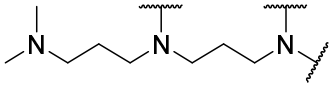
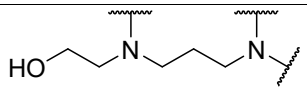
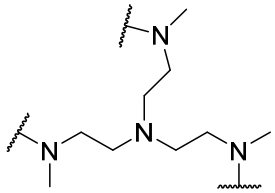
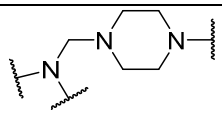
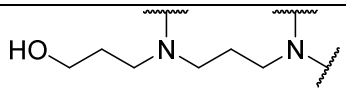
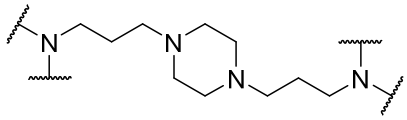
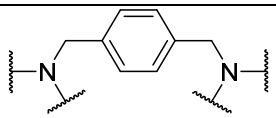
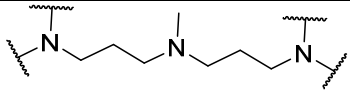
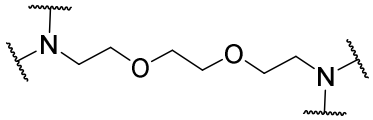
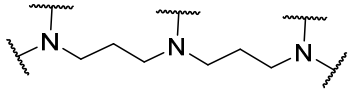
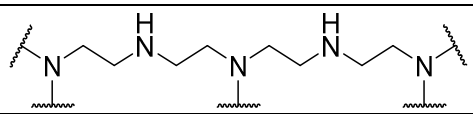
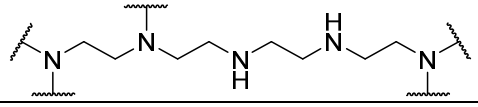
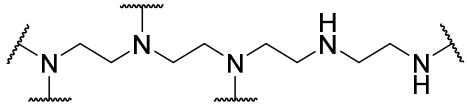
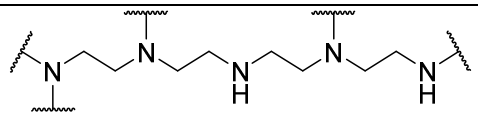
optionally substituted aryl or an optionally substituted (e.g., C_3-C_{12}, such as C_3-C_5) heteroaryl. In some embodiments of $X_{Core,}$ the core comprises has a structural formula .

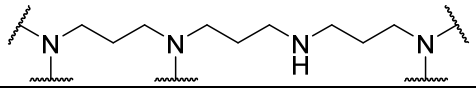
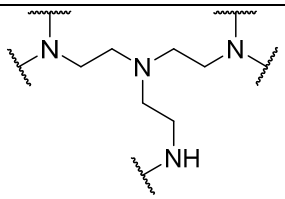
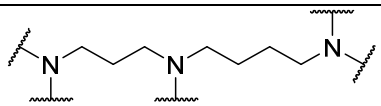
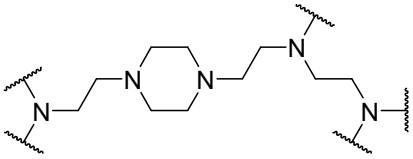
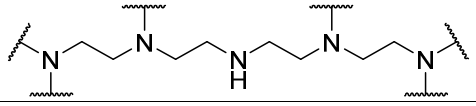
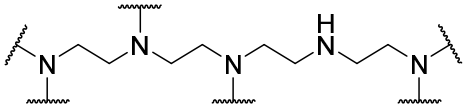
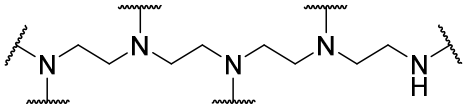
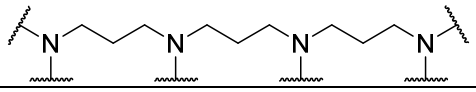
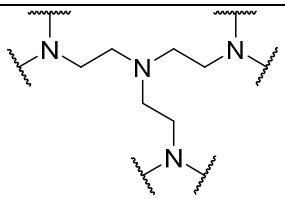
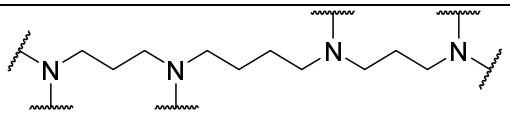
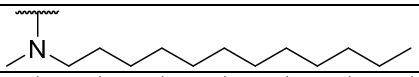
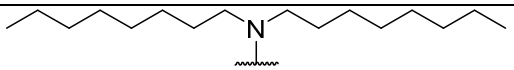
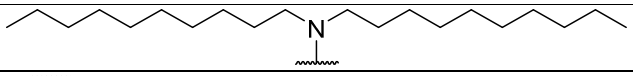
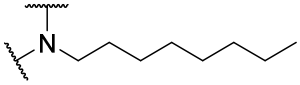
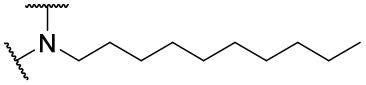
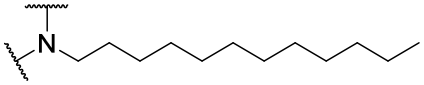
[00190] In some embodiments of X_{Core}, the core comprises a structural formula set forth in **Table. 3** and pharmaceutically acceptable salts thereof, wherein * indicates a point of attachment of the core to a branch of the plurality of branches. In some embodiments, the example cores of **Table. 3** are not limiting of the stereoisomers (i.e. enantiomers, diastereomers) listed.

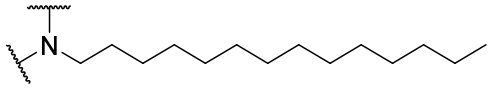
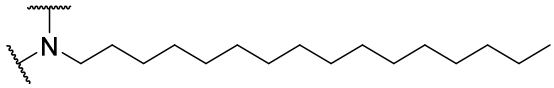
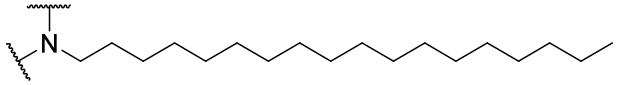
Table 3. Example core structures

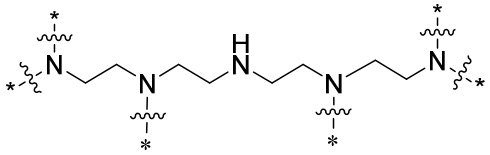
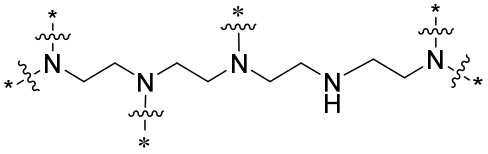
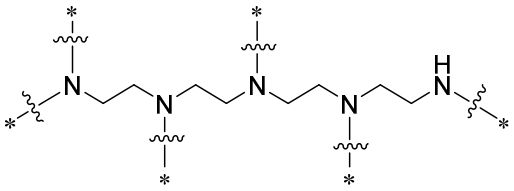
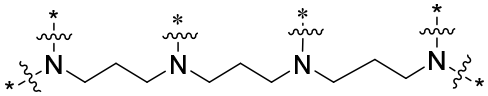
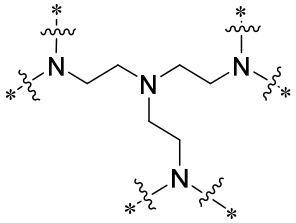
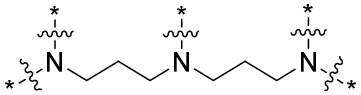
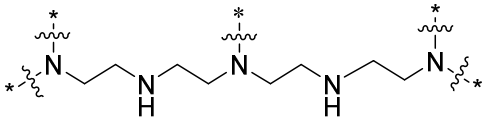
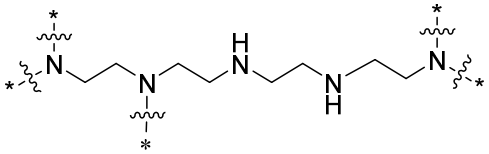
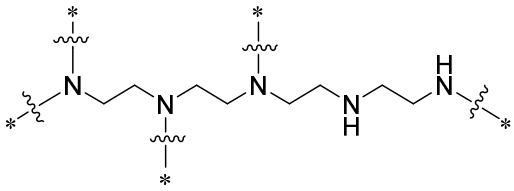
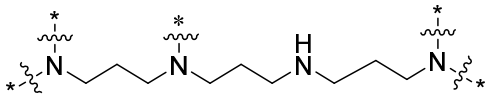
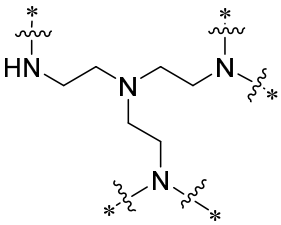
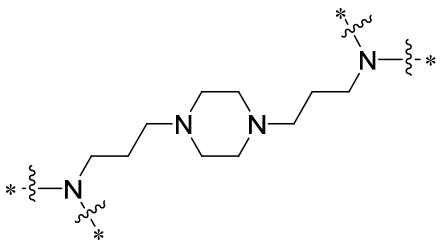
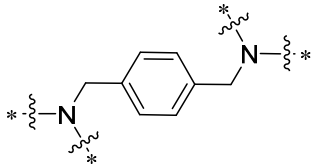
ID #	Structure
1A1	
1A2-1	
1A2-2	
1A3-1	
1A3-2	
1A4	
1A5-1	
1A5-2	
2A1-1	
2A1-2	
2A2-1	
2A2-2	

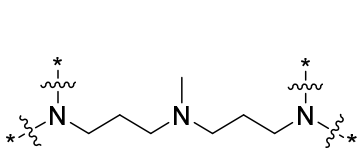
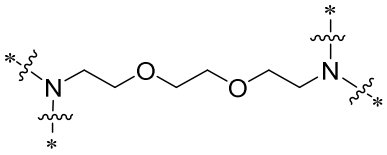
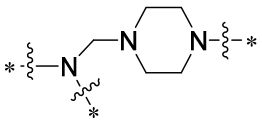
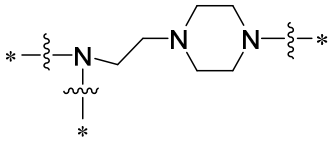
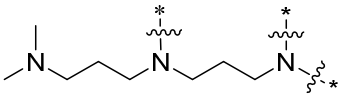
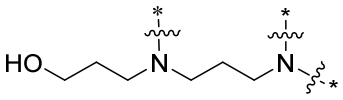
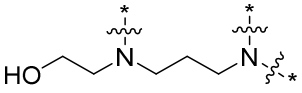
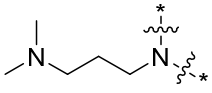
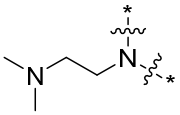
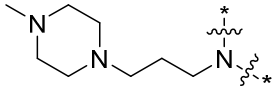
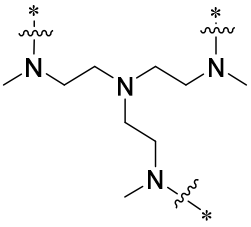
ID #	Structure
2A3	
2A4	
2A5	
2A6	
2A7-1	
2A7-2	
2A8	
2A9	
2A9V	
2A10	
2A11	
2A12	
3A1	

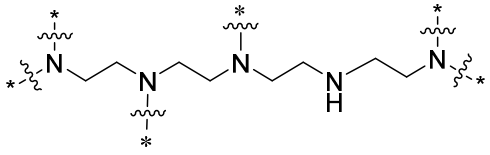
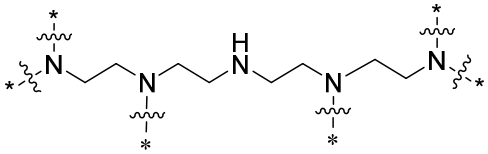
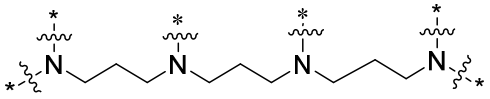
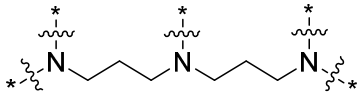
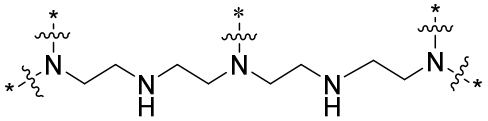
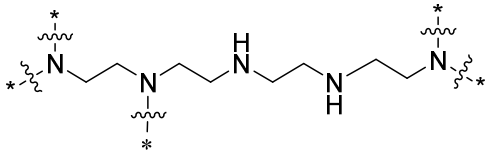
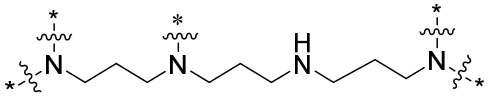
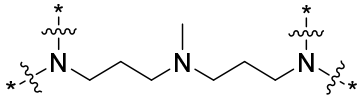
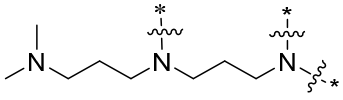
ID #	Structure
3A2	
3A3	
3A4	
3A5	
3A6	
3A7	
4A1	
4A2	
4A3	
4A4	
5A1	
5A2-1 (5-arm)	
5A2-2 (5-arm)	
5A2-3 (5-arm)	
5A2-4 (5-arm)	

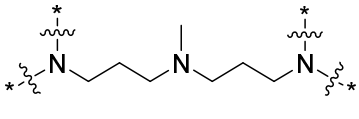
ID #	Structure
5A3-1 (5-arm)	
5A4-1 (5-arm)	
5A5	
5A6	
5A2-4 (6 arm)	
5A2-5 (6 arm)	
5A2-6 (6 arm)	
5A3-2 (6 arm)	
5A4-2 (6 arm)	
6A4	
1H1	
1H2	
1H3	
2H1	
2H2	
2H3	

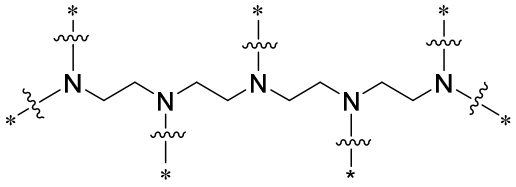
ID #	Structure
2H4	
2H5	
2H6	

[00191] In some embodiments of X_{Core}, the core comprises a structural formula selected from the group consisting of: , , , , , , , , , , , , ,


,

,

,

,

,

,

,

,

,

,

, and pharmaceutically acceptable salts thereof, wherein * indicates a point of attachment of the core to a branch of the plurality of branches or H. In some embodiments, wherein * indicates a point of attachment of the core to a branch of the plurality of branches.

[00192] In some embodiments of X_{Core}, the core comprises a structural formula selected from the group consisting of:

,

,

,

,

,

,

,

,

, and pharmaceutically acceptable salts thereof, wherein * indicates a point of attachment of the core to a branch of the plurality of branches.

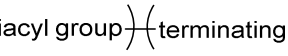
[00193] In some embodiments of X_{Core}, the core has the structure , wherein * indicates a point of attachment of the core to a branch of the plurality of branches or H. In some embodiments, at least 2 branches are attached to the core. In some embodiments, at least 3 branches are attached to the core. In some embodiments, at least 4 branches are attached to the core.

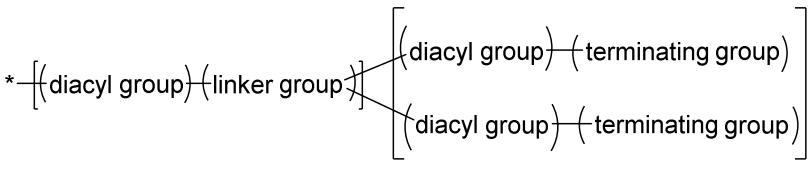
[00194] In some embodiments of X_{Core}, the core has the structure , wherein * indicates a point of attachment of the core to a branch of the plurality of branches or H. In some embodiments, at least 4 branches are attached to the core. In some embodiments, at least 5 branches are attached to the core. In some embodiments, at least 6 branches are attached to the core.

[00195] In some embodiments, the plurality (N) of branches comprises at least 3 branches, at least 4 branches, at least 5 branches. In some embodiments, the plurality (N) of branches comprises at least 3 branches. In some embodiments, the plurality (N) of branches comprises at least 4 branches. In some embodiments, the plurality (N) of branches comprises at least 5 branches.

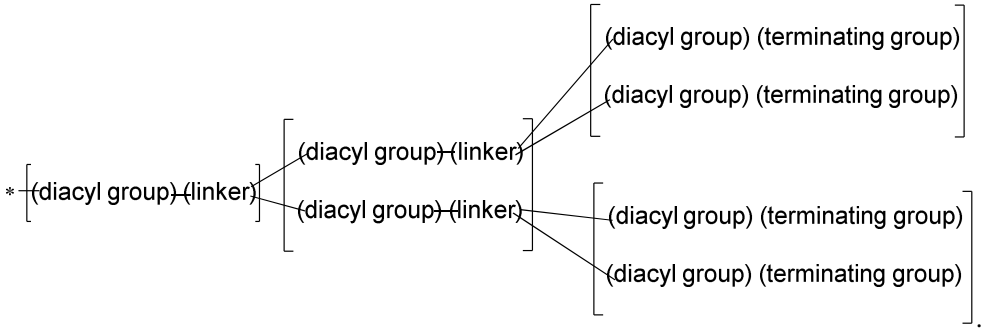
[00196] In some embodiments of X_{Branch}, g is 1, 2, 3, or 4. In some embodiments of X_{Branch}, g is 1. In some embodiments of X_{Branch}, g is 2. In some embodiments of X_{Branch}, g is 3. In some embodiments of X_{Branch}, g is 4.

[00197] In some embodiments of X_{Branch}, $Z = 2^{(g-1)}$ and when g=1, G=0. In some embodiments of X_{Branch}, $Z = 2^{(g-1)}$ and $G = \sum_{i=0}^{i=g-2} 2^i$, when g≠1.

[00198] In some embodiments of X_{Branch}, g=1, G=0, Z=1, and each branch of the plurality of branches comprises a structural formula each branch of the plurality of branches comprises a structural formula .

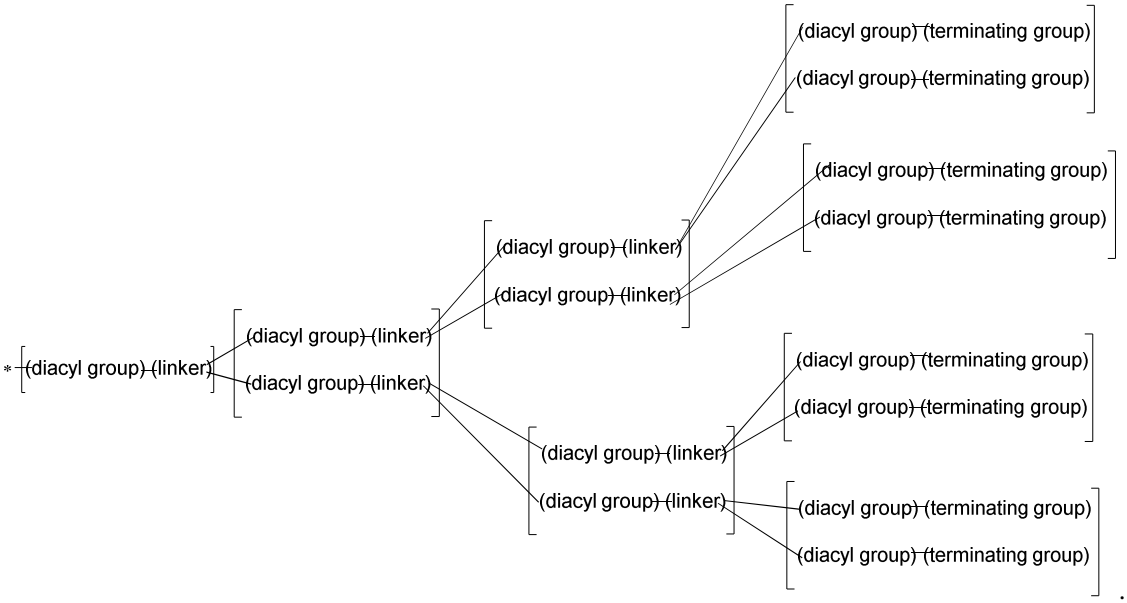
[00199] In some embodiments of X_{Branch}, g=2, G=1, Z=2, and each branch of the plurality of branches comprises a structural formula .

[00200] In some embodiments of X_{Branch}, g=3, G=3, Z=4, and each branch of the plurality of branches comprises a structural formula



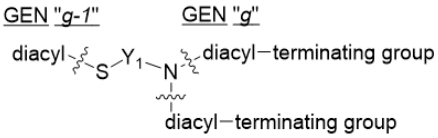
.

[00201] In some embodiments of X_{Branch}, g=4, G=7, Z=8, and each branch of the plurality of branches comprises a structural formula



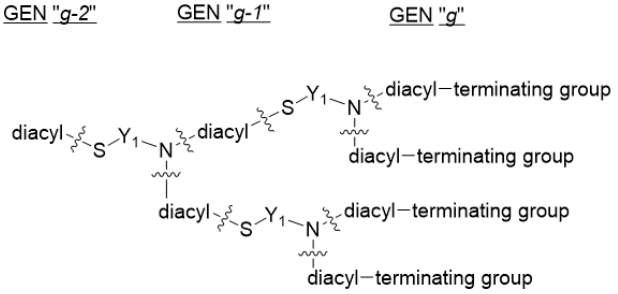
.

[00202] In some embodiments, the dendrimers described herein with a generation (g) = 1 has the structure:



.

[00203] In some embodiments, the dendrimers described herein with a generation (g) = 1 has the structure:

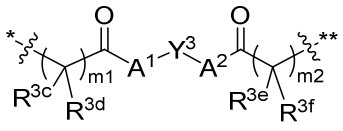


.

[00204] The example formulation of the dendrimers described herein for generations 1-4 is shown in **Table 4**. The number of diacyl groups, linker groups, and terminating groups can be calculated based on g.

Table 4. Formulation of Dendrimer Groups Based on Generation (g)

	$g = 1$	$g = 2$	$g = 3$	$g = 4$	
# of diacyl grp	1	1+2=3	$1+2+2^2=7$	$1+2+2^2+2^3=15$	$1+2+\ldots+2^{g-1}$
# of linker grp	0	1	1+2	$1+2+2^2$	$1+2+\ldots+2^{g-2}$
# of terminating grp	1	2	2^2	2^3	$2^{(g-1)}$

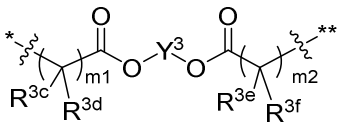
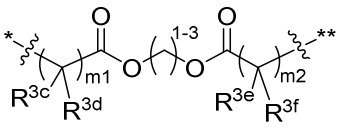
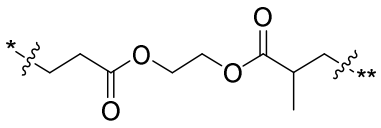
[00205] In some embodiments, the diacyl group independently comprises a structural formula 

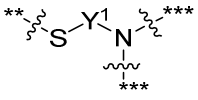
, * indicates a point of attachment of the diacyl group at the proximal end thereof, and ** indicates a point of attachment of the diacyl group at the distal end thereof.

[00206] In some embodiments of the diacyl group of X_{Branch}, Y^3 is independently at each occurrence an optionally substituted; alkylene, an optionally substituted alkenylene, or an optionally substituted arenylene. In some embodiments of the diacyl group of X_{Branch}, Y^3 is independently at each occurrence an optionally substituted alkylene (e.g., C_1-C_{12}). In some embodiments of the diacyl group of X_{Branch}, Y^3 is independently at each occurrence an optionally substituted alkenylene (e.g., C_1-C_{12}). In some embodiments of the diacyl group of X_{Branch}, Y^3 is independently at each occurrence an optionally substituted arenylene (e.g., C_1-C_{12}).

[00207] In some embodiments of the diacyl group of X_{Branch}, A^1 and A^2 are each independently at each occurrence -O-, -S-, or -NR^4-. In some embodiments of the diacyl group of X_{Branch}, A^1 and A^2 are each independently at each occurrence -O-. In some embodiments of the diacyl group of X_{Branch}, A^1 and A^2 are each independently at each occurrence -S-. In some embodiments of the diacyl group of X_{Branch}, A^1 and A^2 are each independently at each occurrence -NR^4- and R^4 is hydrogen or optionally substituted alkyl (e.g., C_1-C_6). In some embodiments of the diacyl group of X_{Branch}, m^1 and m^2 are each independently at each occurrence 1, 2, or 3. In some embodiments of the diacyl group of X_{Branch}, m^1 and m^2 are each independently at each occurrence 1. In some embodiments of the diacyl group of X_{Branch}, m^1 and m^2 are each independently at each occurrence 2. In some embodiments of the diacyl group of X_{Branch}, m^1 and m^2 are each independently at each occurrence 3. In some embodiments of the diacyl group of X_{Branch}, R^{3c}, R^{3d}, R^{3e}, and R^{3f} are each independently at each occurrence hydrogen or an optionally substituted alkyl. In some embodiments of the diacyl group of X_{Branch}, R^{3c}, R^{3d}, R^{3e}, and R^{3f} are each independently at each occurrence hydrogen. In some embodiments of the diacyl group of X_{Branch}, R^{3c}, R^{3d}, R^{3e}, and R^{3f} are each independently at each occurrence an optionally substituted (e.g., C_1-C_8) alkyl.

[00208] In some embodiments of the diacyl group, A^1 is -O- or -NH-. In some embodiments of the diacyl group, A^1 is -O-. In some embodiments of the diacyl group, A^2 is -O- or -NH-. In some embodiments of the diacyl group, A^2 is -O-. In some embodiments of the diacyl group, Y^3 is C_1-C_{12} (e.g., C_1-C_6, such as C_1-C_3) alkylene.

[00209] In some embodiments of the diacyl group, the diacyl group independently at each occurrence comprises a structural formula  (e.g., , such as ), and optionally R^{3c}, R^{3d}, R^{3e}, and R^{3f} are each independently at each occurrence hydrogen or C_1-C_3 alkyl.

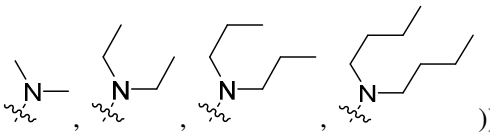
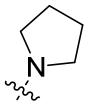
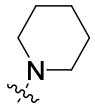
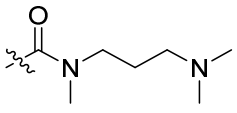
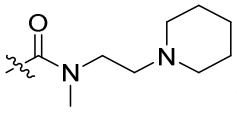
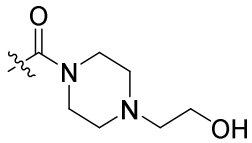
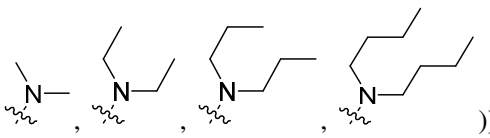
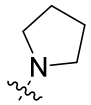
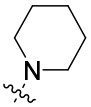
[00210] In some embodiments, linker group independently comprises a structural formula , ** indicates a point of attachment of the linker to a proximal diacyl group, and *** indicates a point of attachment of the linker to a distal diacyl group.

[00211] In some embodiments of the linker group of X_{Branch} if present, Y_1 is independently at each occurrence an optionally substituted alkylene, an optionally substituted alkenylene, or an optionally substituted arenylene. In some embodiments of the linker group of X_{Branch} if present, Y_1 is independently at each occurrence an optionally substituted alkylene (e.g., C_1-C_{12}). In some embodiments of the linker group of X_{Branch} if present, Y_1 is independently at each occurrence an optionally substituted alkenylene (e.g., C_1-C_{12}). In some embodiments of the linker group of X_{Branch} if present, Y_1 is independently at each occurrence an optionally substituted arenylene (e.g., C_1-C_{12}).

[00212] In some embodiments of the terminating group of X_{Branch}, each terminating group is independently selected from optionally substituted alkylthiol and optionally substituted alkenylthiol. In some embodiments of the terminating group of X_{Branch}, each terminating group is an optionally substituted alkylthiol (e.g., C_1-C_{18}, such as C_4-C_{18}). In some embodiments of the terminating group of X_{Branch}, each terminating group is optionally substituted alkenylthiol (e.g., C_1-C_{18}, such as C_4-C_{18}).

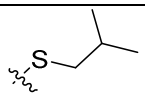
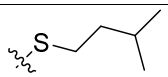
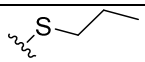
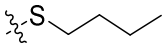
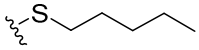
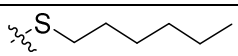
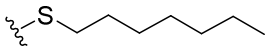
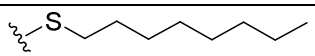
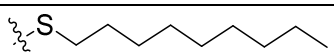
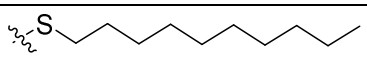
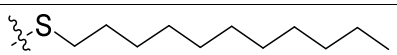
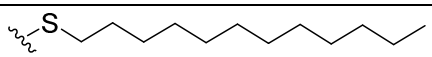
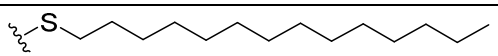
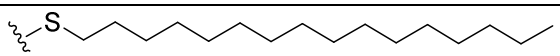
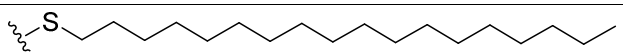
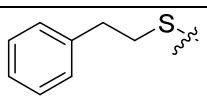
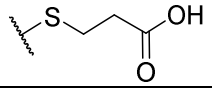
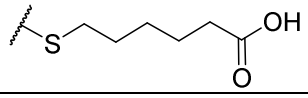
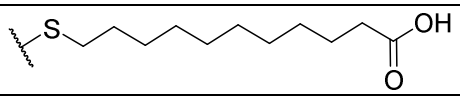
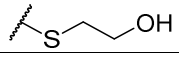
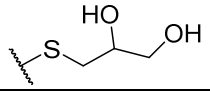
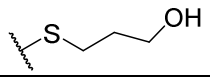
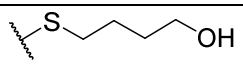
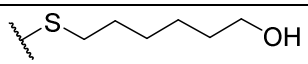
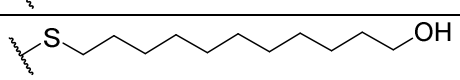
[00213] In some embodiments of the terminating group of X_{Branch}, each terminating group is independently C_1-C_{18} alkenylthiol or C_1-C_{18} alkylthiol, and the alkyl or alkenyl moiety is optionally substituted with one or more substituents each independently selected from halogen, C_6-C_{12} aryl, C_1-C_{12} alkylamino, C_4-C_6 *N*-heterocycloalkyl , -OH, -C(O)OH, –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_1-C_{12} alkylamino), –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_4-C_6 *N*-heterocycloalkyl), –C(O)–(C_1-C_{12} alkylamino), and –C(O)–(C_4-C_6 *N*-heterocycloalkyl), and the C_4-C_6 *N*-heterocycloalkyl moiety of any of the preceding substituents is optionally substituted with C_1-C_3 alkyl or C_1-C_3 hydroxyalkyl.

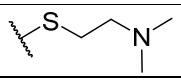
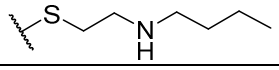
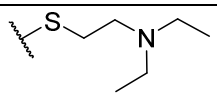
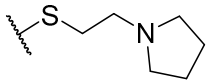
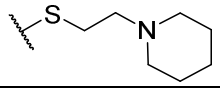
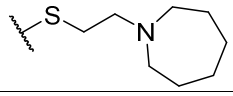
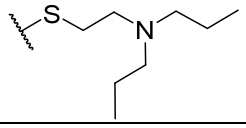
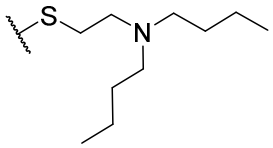
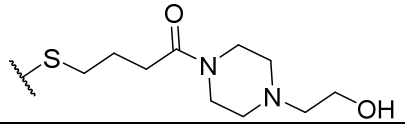
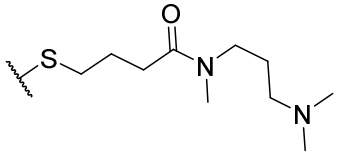
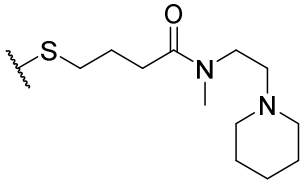
[00214] In some embodiments of the terminating group of X_{Branch}, each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkenylthiol or C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol, wherein the alkyl or alkenyl moiety is optionally substituted with one or more substituents each independently selected from halogen, C_6-C_{12} aryl (e.g., phenyl), C_1-C_{12} (e.g., C_1-C_8) alkylamino (e.g., C_1-C_6 mono-alkylamino (such as -$NHCH_2CH_2CH_2CH_3$)

or C_1-C_8 di-alkylamino (such as )), C_4-C_6 *N*-heterocycloalkyl (e.g., *N*-pyrrolidinyl (), *N*-piperidinyl (), *N*-azepanyl ()), -OH, -C(O)OH, –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_1-C_{12} alkylamino (e.g., mono- or di-alkylamino)) (e.g., ), –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_4-C_6 *N*-heterocycloalkyl) (e.g., ), –C(O)–(C_1-C_{12} alkylamino (e.g., mono- or di-alkylamino)), and –C(O)–(C_4-C_6 *N*-heterocycloalkyl) (e.g., ), wherein the C_4-C_6 *N*-heterocycloalkyl moiety of any of the preceding substituents is optionally substituted with C_1-C_3 alkyl or C_1-C_3 hydroxyalkyl. In some embodiments of the terminating group of X_{Branch}, each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol, wherein the alkyl moiety is optionally substituted with one substituent **-**OH. In some embodiments of the terminating group of X_{Branch}, each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol, wherein the alkyl moiety is optionally substituted with one substituent selected from C_1-C_{12} (e.g., C_1-C_8) alkylamino (e.g., C_1-C_6 mono-alkylamino (such as -$NHCH_2CH_2CH_2CH_3$) or C_1-C_8 di-alkylamino (such as )) and C_4-C_6 *N*-heterocycloalkyl (e.g., *N*-pyrrolidinyl (), *N*-piperidinyl (), *N*-azepanyl ()). In some embodiments of the terminating group of X_{Branch}, each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkenylthiol or C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol. In some embodiments of the terminating group of X_{Branch}, each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol.

[00215] In some embodiments of the terminating group of X_{Branch}, each terminating group is independently a structural set forth in **Table 5.** In some embodiments, the dendrimers described herein can comprise a terminating group or pharmaceutically acceptable salt, or thereof selected in **Table 5**. In some embodiments, the example terminating group of **Table 5** are not limiting of the stereoisomers (i.e. enantiomers, diastereomers) listed.

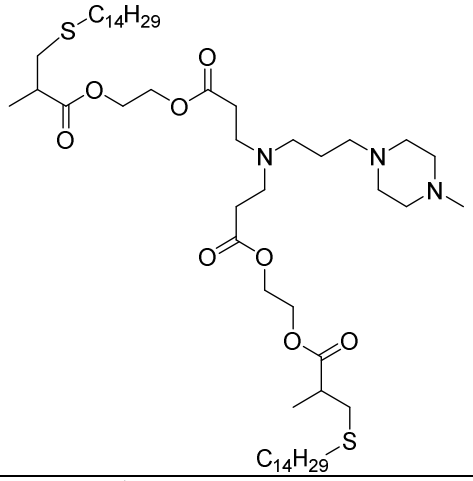
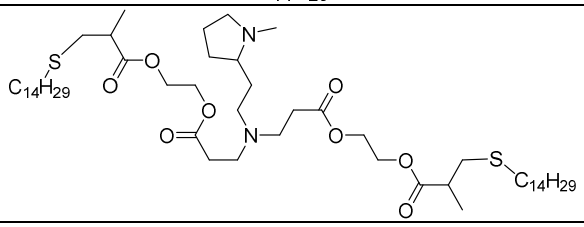
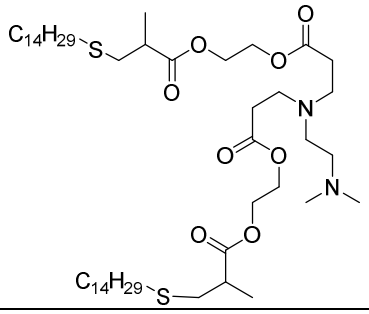
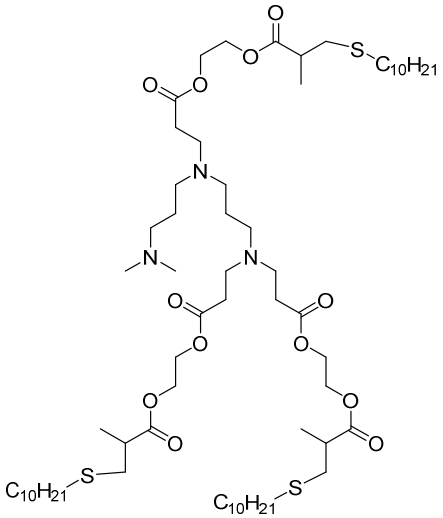
Table 5. Example terminating group / peripheries structures

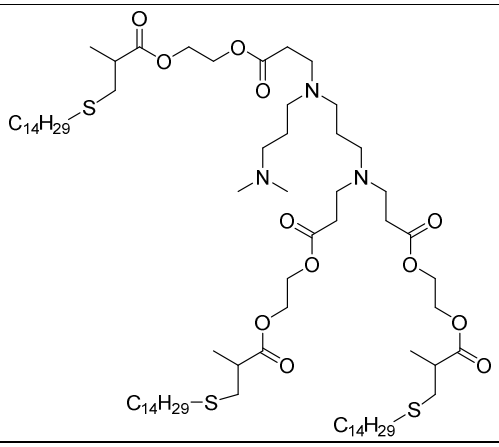
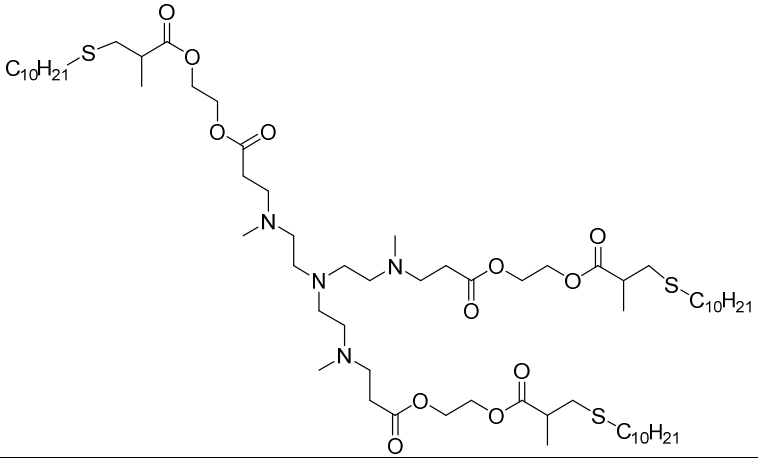
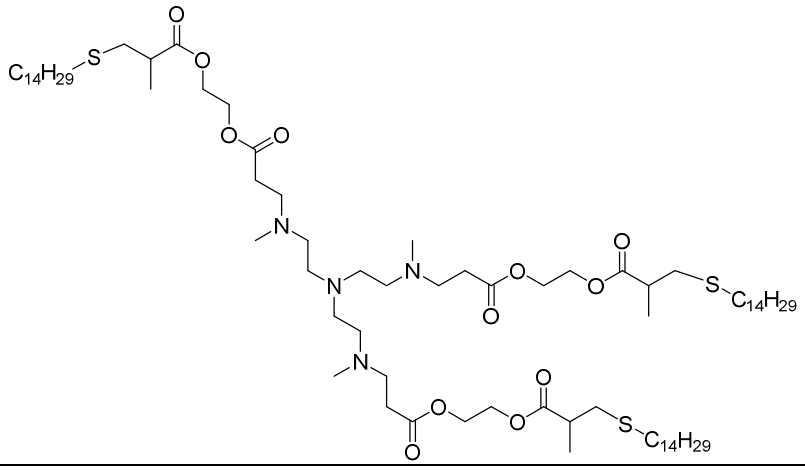
ID #	Structure
SC1	
SC2	
SC3	
SC4	
SC5	
SC6	
SC7	
SC8	
SC9	
SC10	
SC11	
SC12	
SC14	
SC16	
SC18	
SC19	
SO1	
SO2	
SO3	
SO4	
SO5	
SO6	
SO7	
SO8	
SO9	

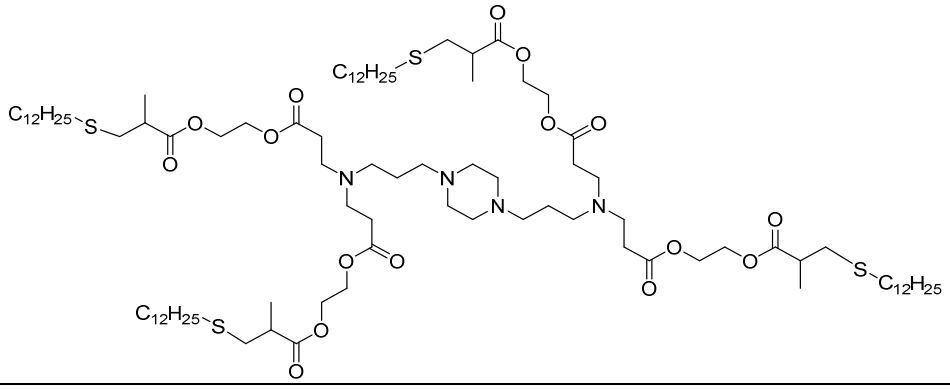
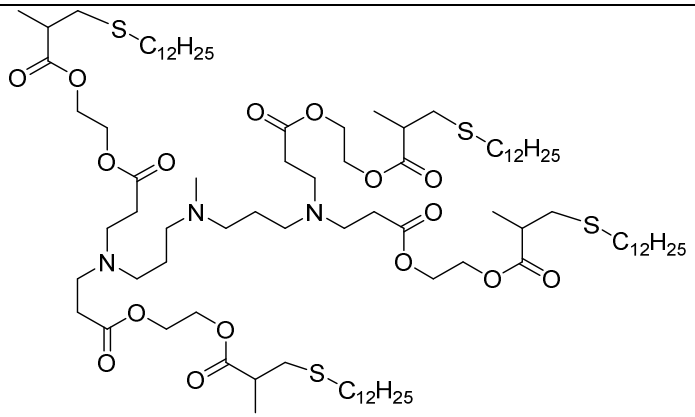
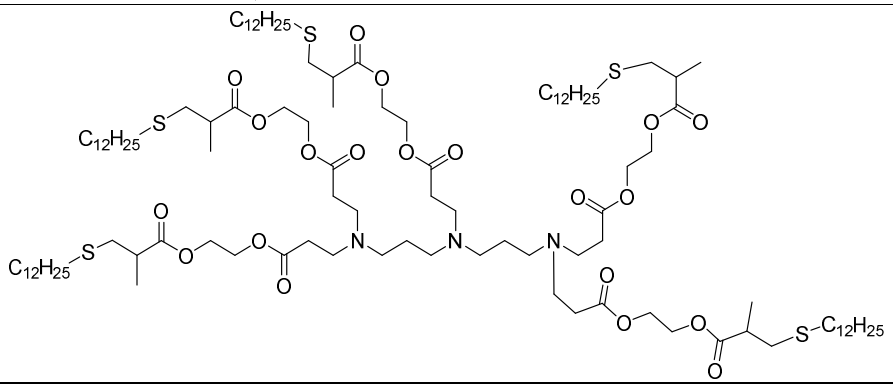
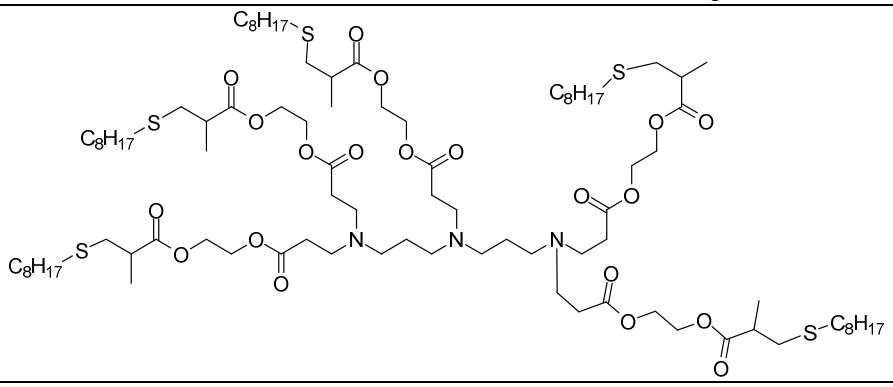
ID #	Structure
SN1	
SN2	
SN3	
SN4	
SN5	
SN6	
SN7	
SN8	
SN9	
SN10	
SN11	

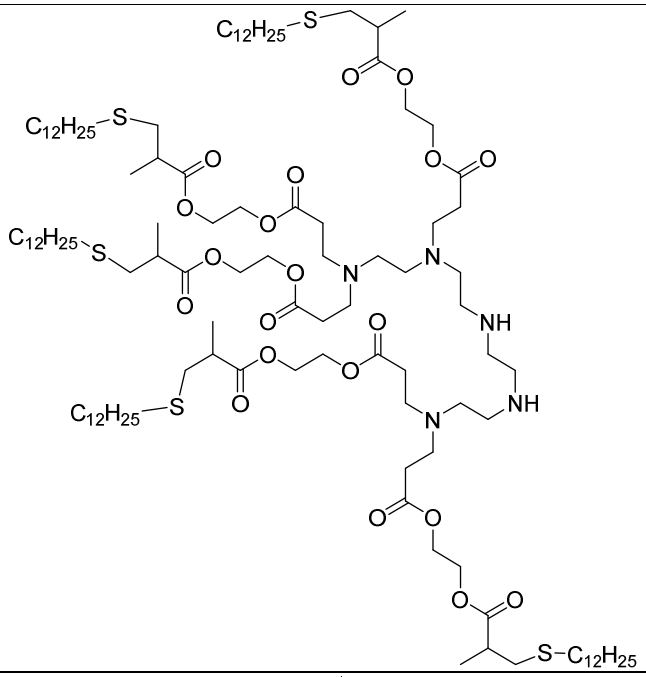
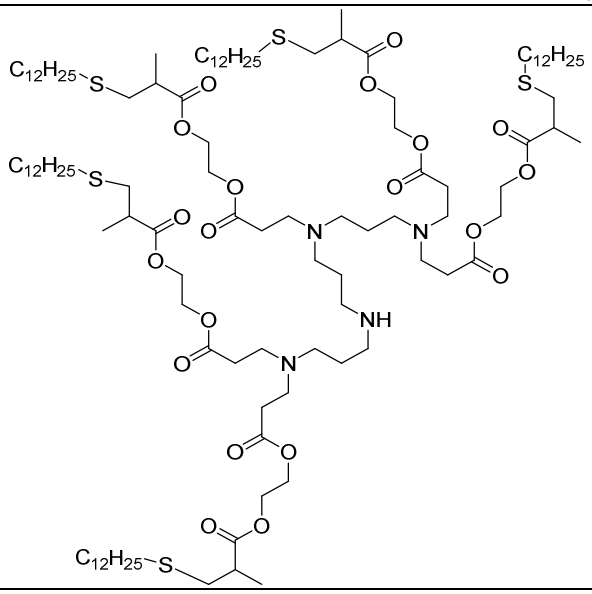
[00216] In some embodiments, the dendrimer of Formula (X) is selected from those set forth in **Table 6** and pharmaceutically acceptable salts thereof.

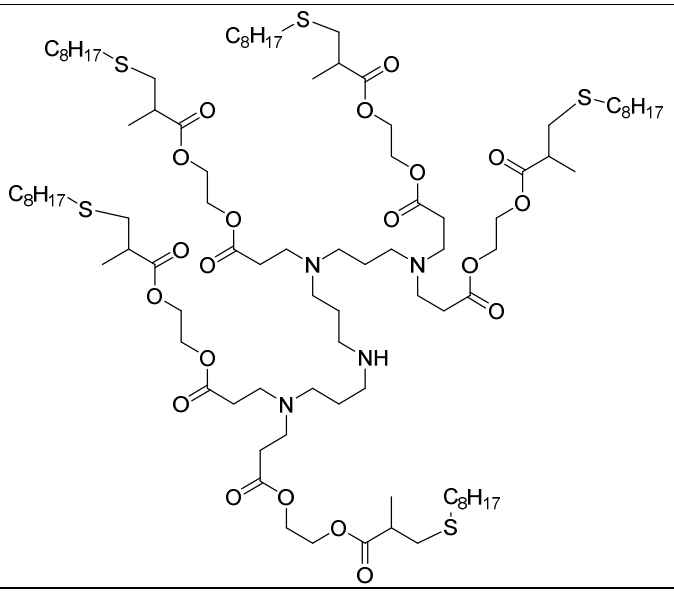
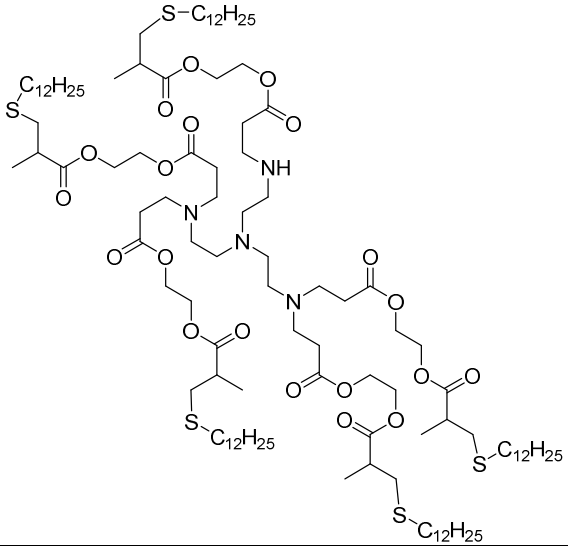
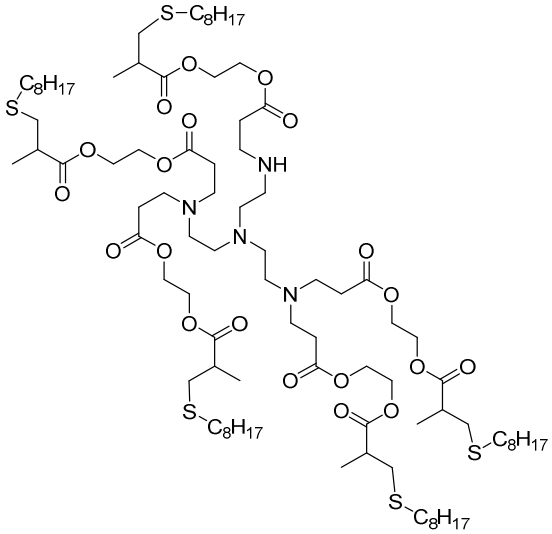
Table 6. Example ionizable cationic lipo-dendrimers

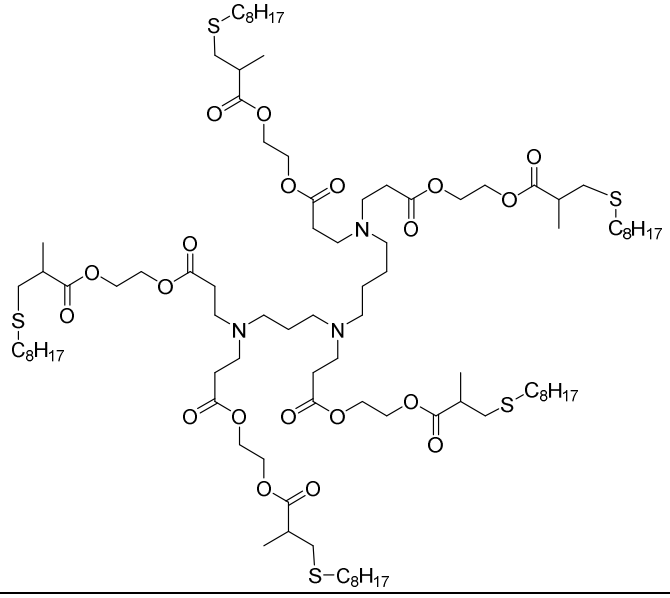
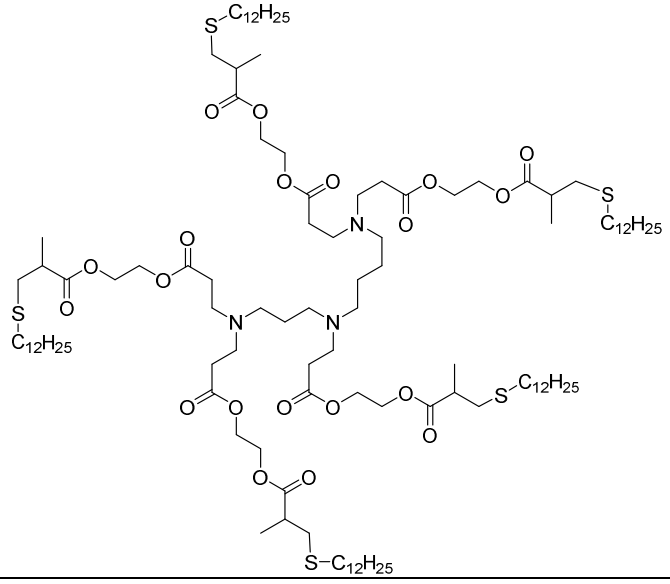
ID #	Structure
2A2-SC14	
2A6-SC14	
2A9-SC14	
3A3-SC10	

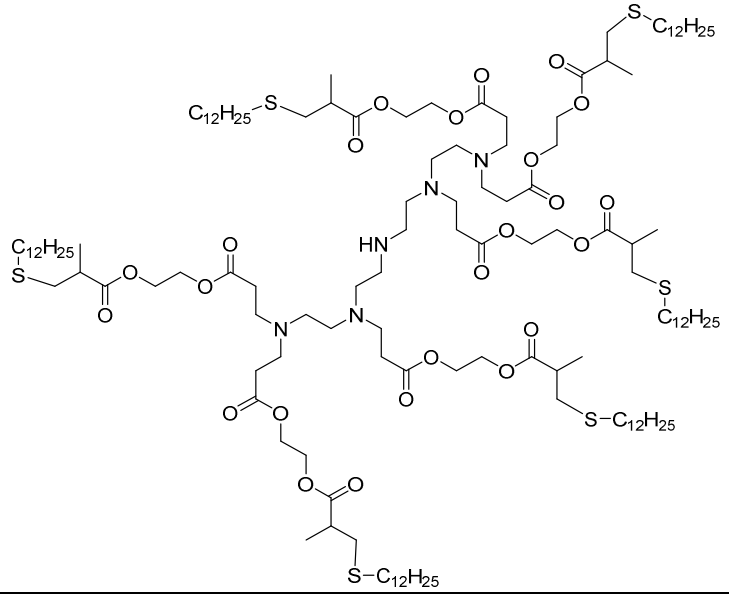
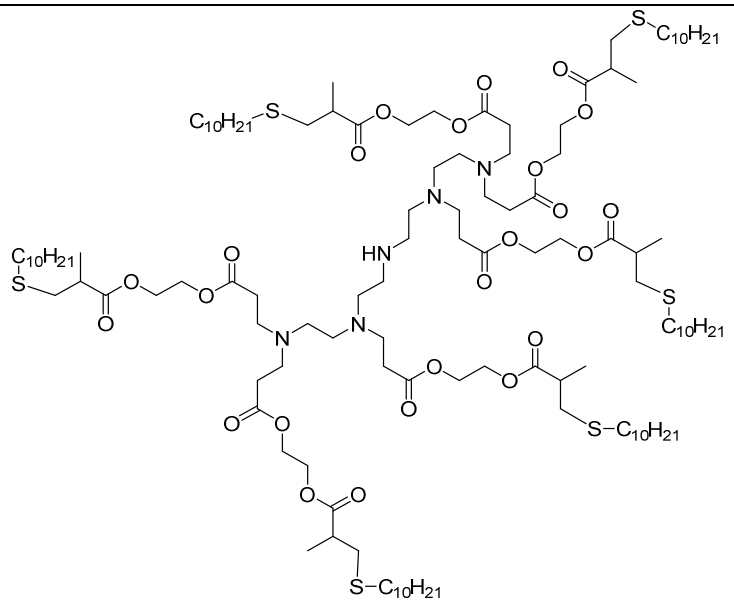
ID #	Structure
3A3-SC14	
3A5-SC10	
3A5-SC14	

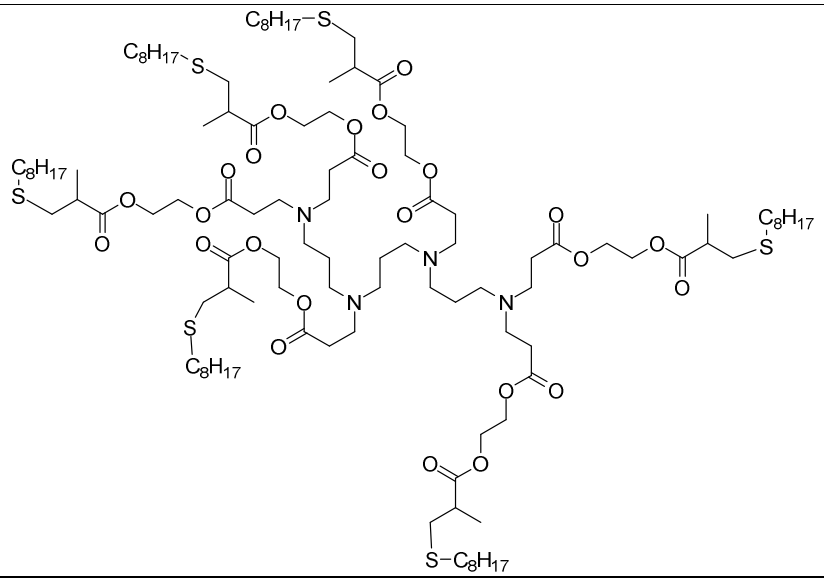
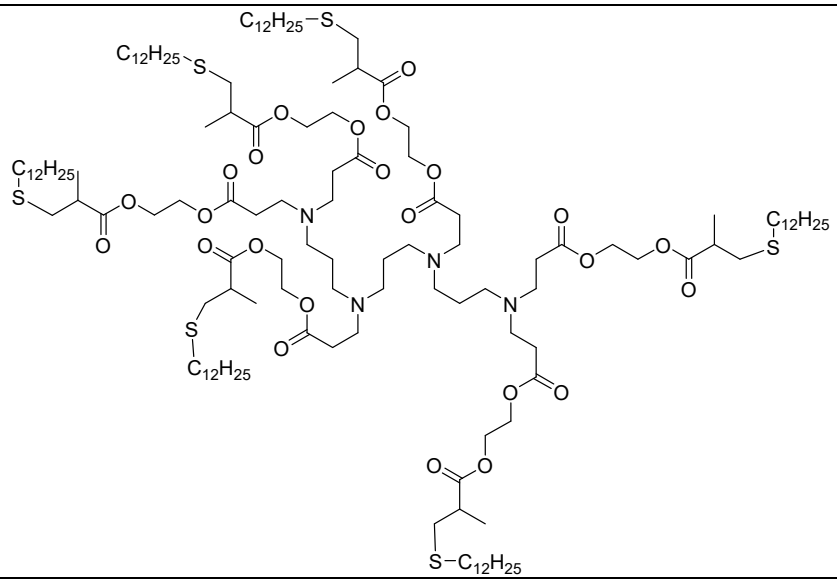
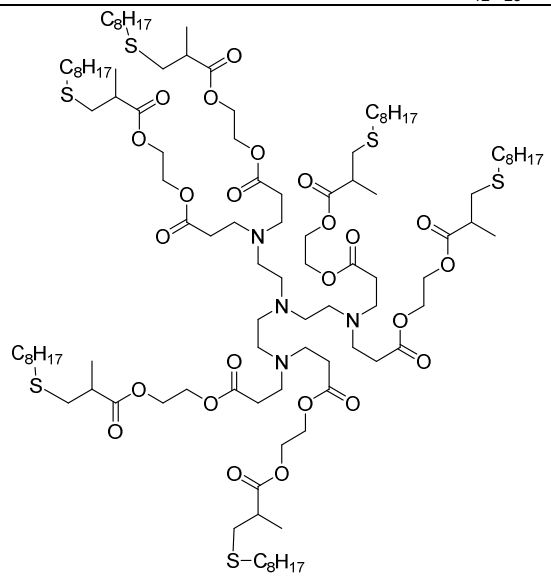
ID #	Structure
4A1-SC12	
4A3-SC12	
5A1-SC12	
5A1-SC8	

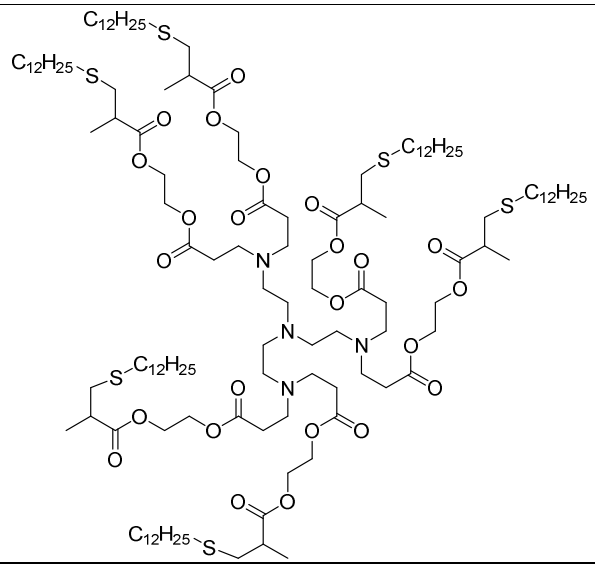
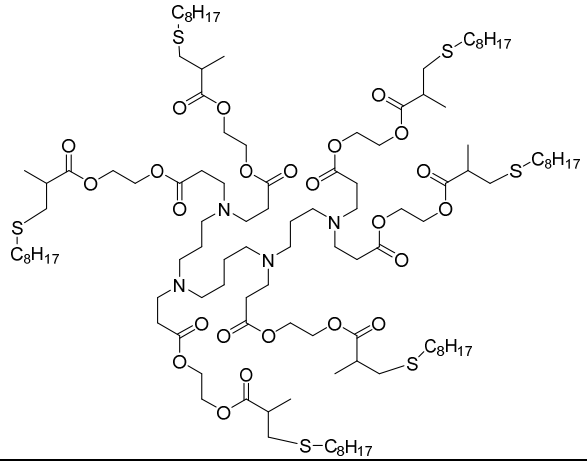
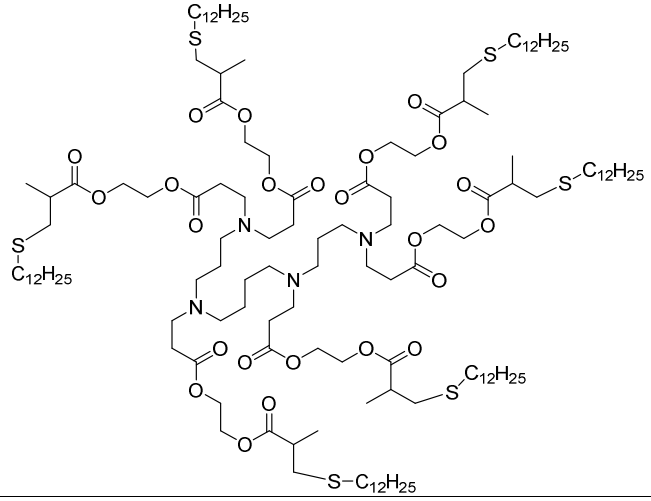
ID #	Structure
5A2-2-SC12 (5-arm)	
5A3-1-SC12 (5 arm)	

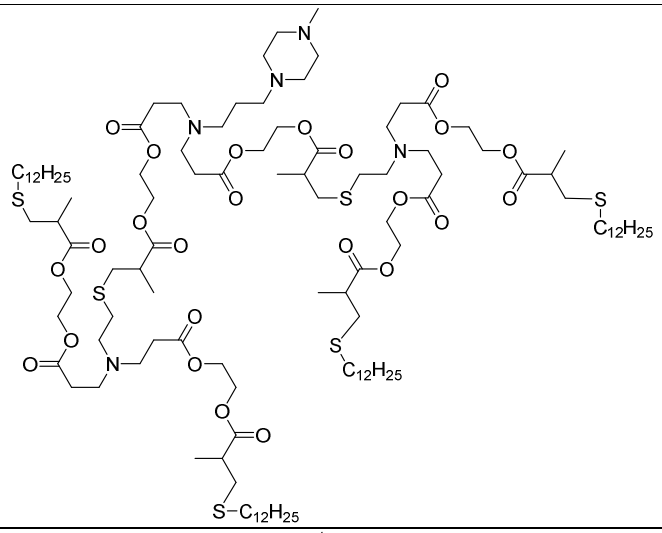
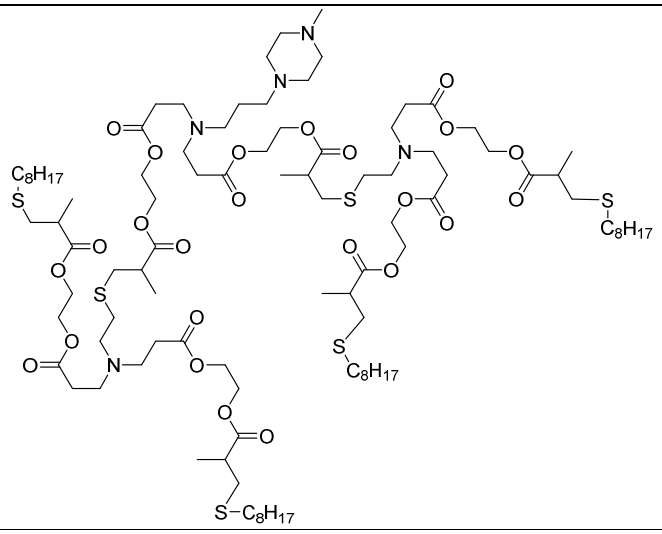
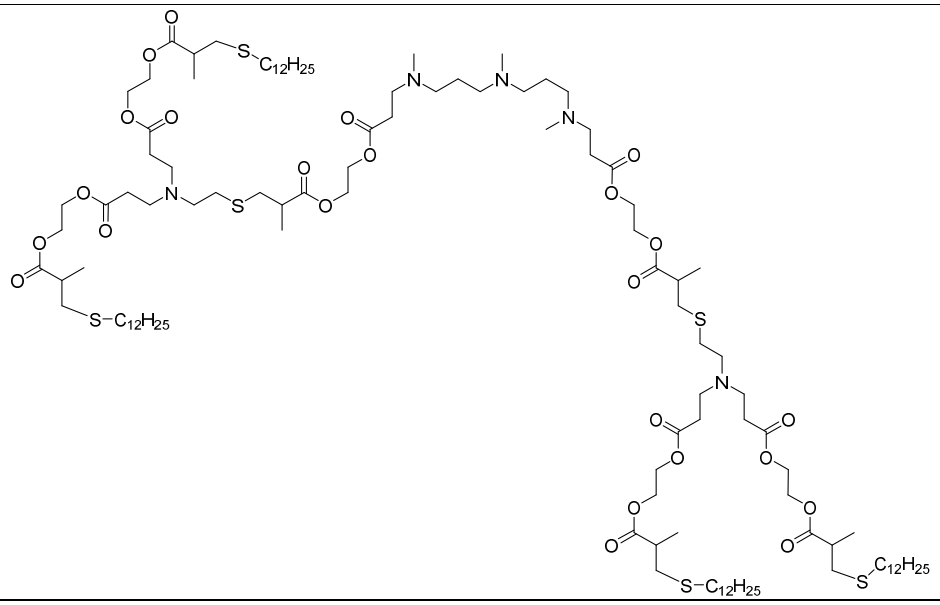
ID #	Structure
5A3-1-SC8 (5-arm)	
5A4-1-SC12 (5-arm)	
5A4-1-SC8 (5-arm)	

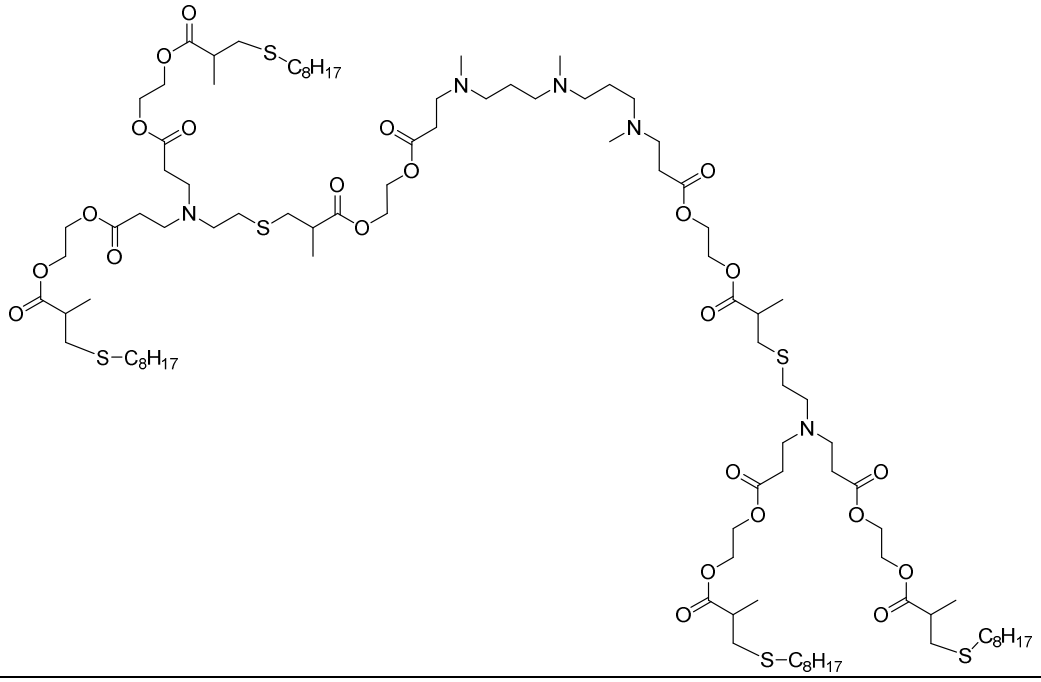
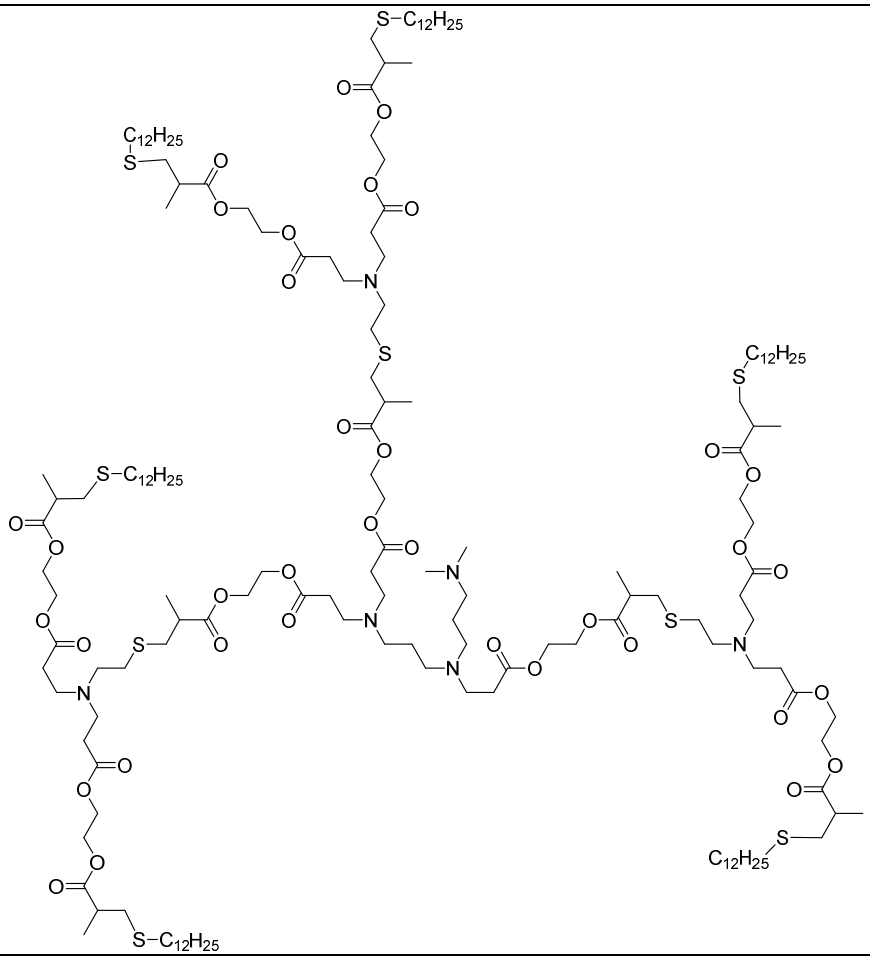
ID #	Structure
5A5-SC8	
5A5-SC12	

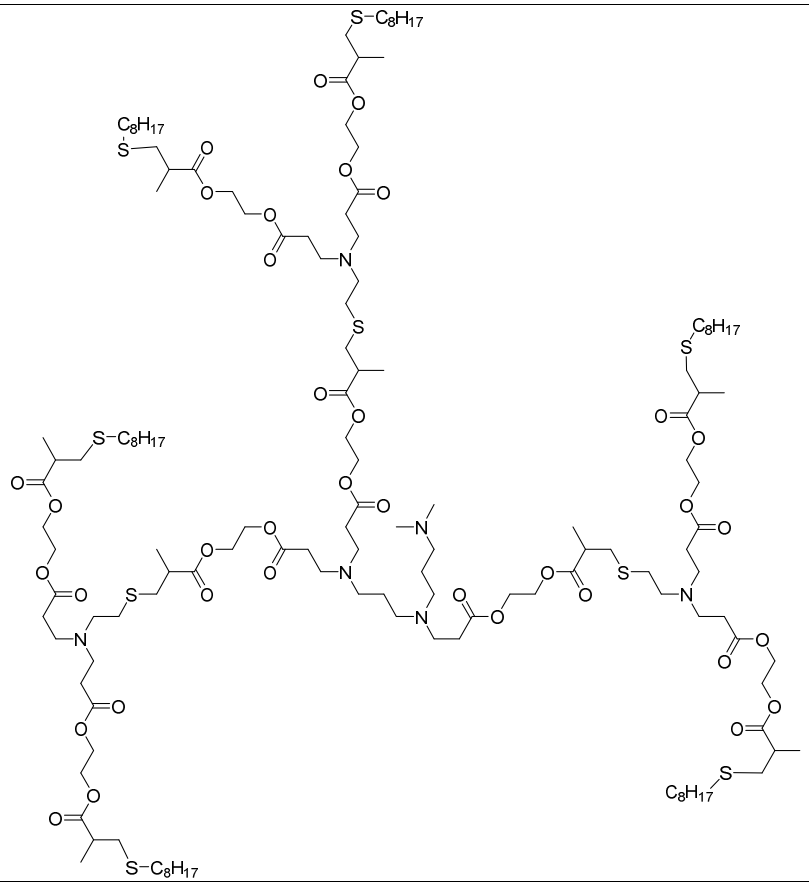
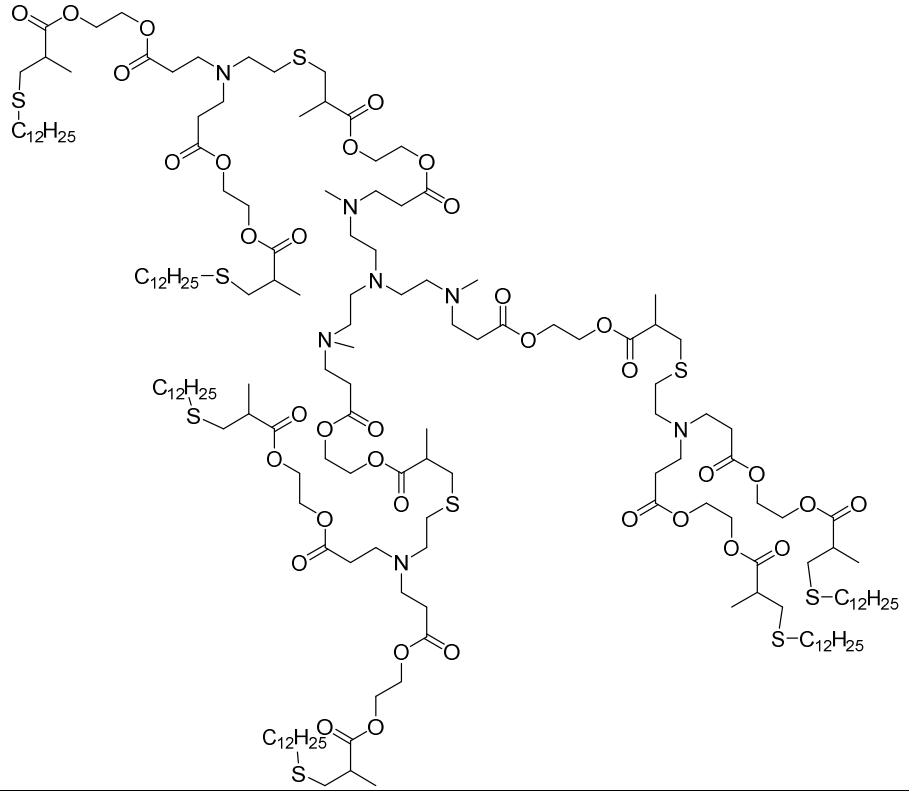
ID #	Structure
5A2-4-SC12 (6-arm)	
5A2-4-SC10 (6-arm)	

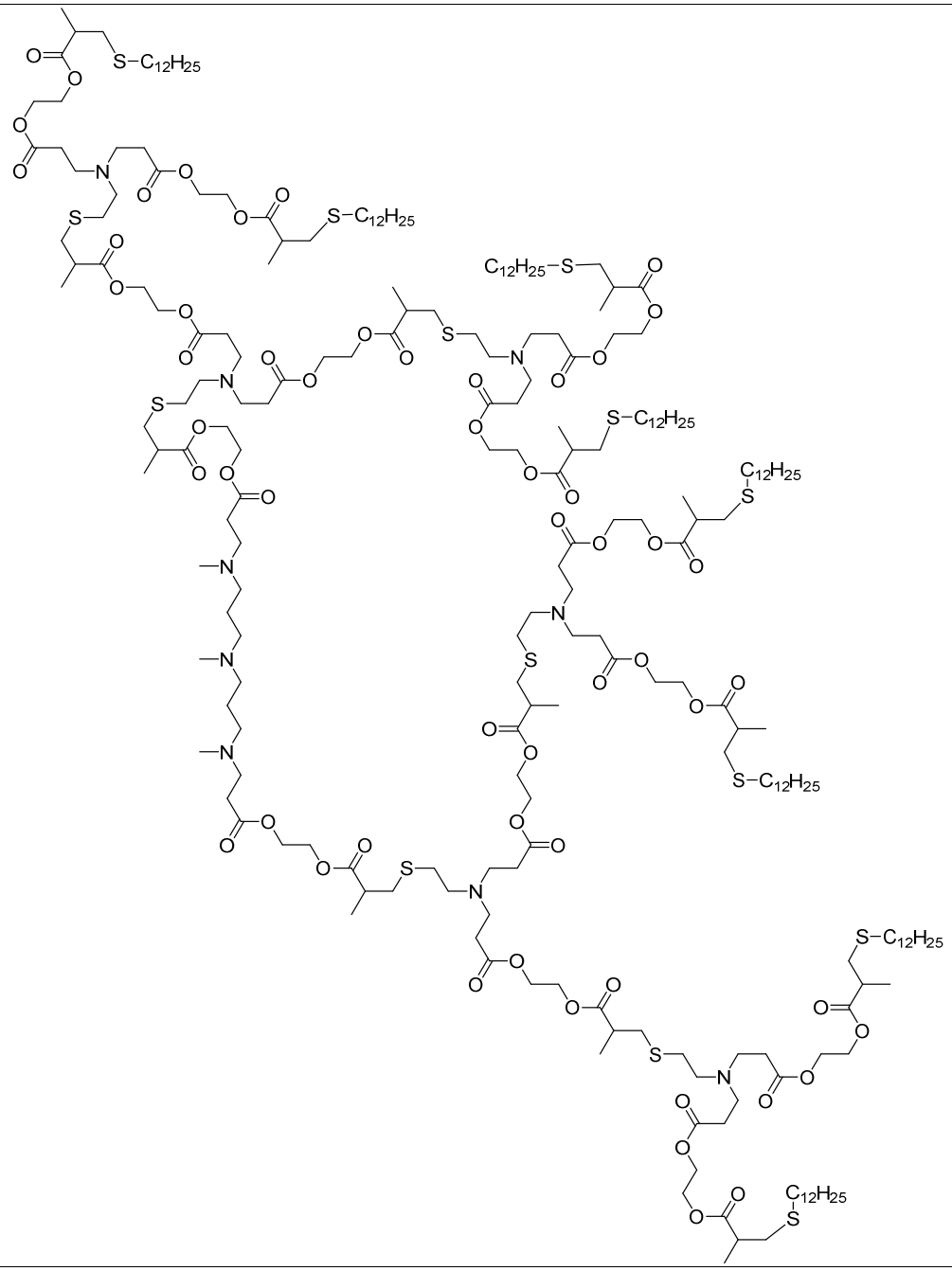
ID #	Structure
5A3-2--SC8 (6-arm)	
5A3-2-SC12 (6-arm)	
5A4-2-SC8 (6-arm)	

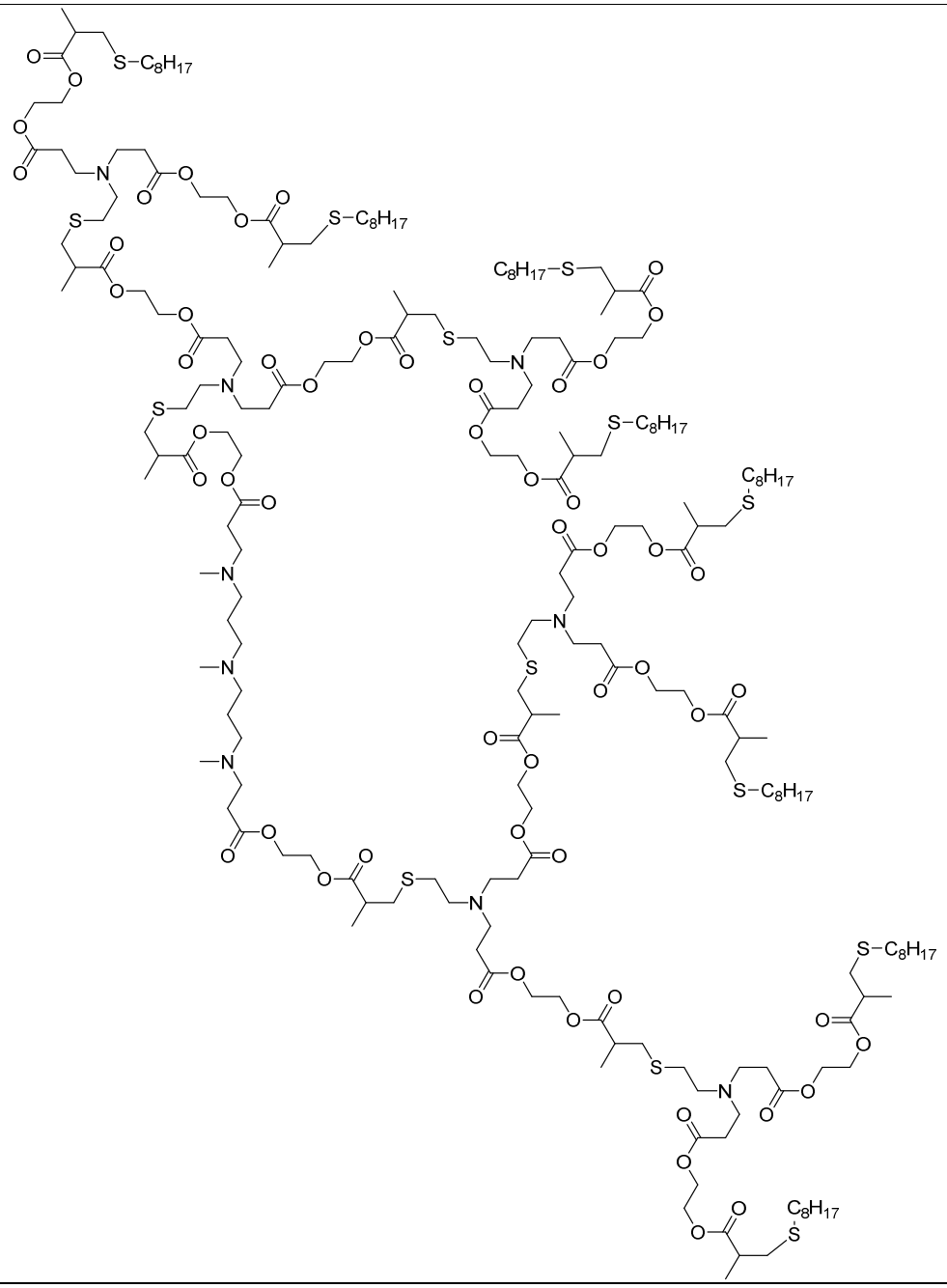
ID #	Structure
5A4-2-SC12 (6-arm)	
6A4-SC8	
6A4-SC12	

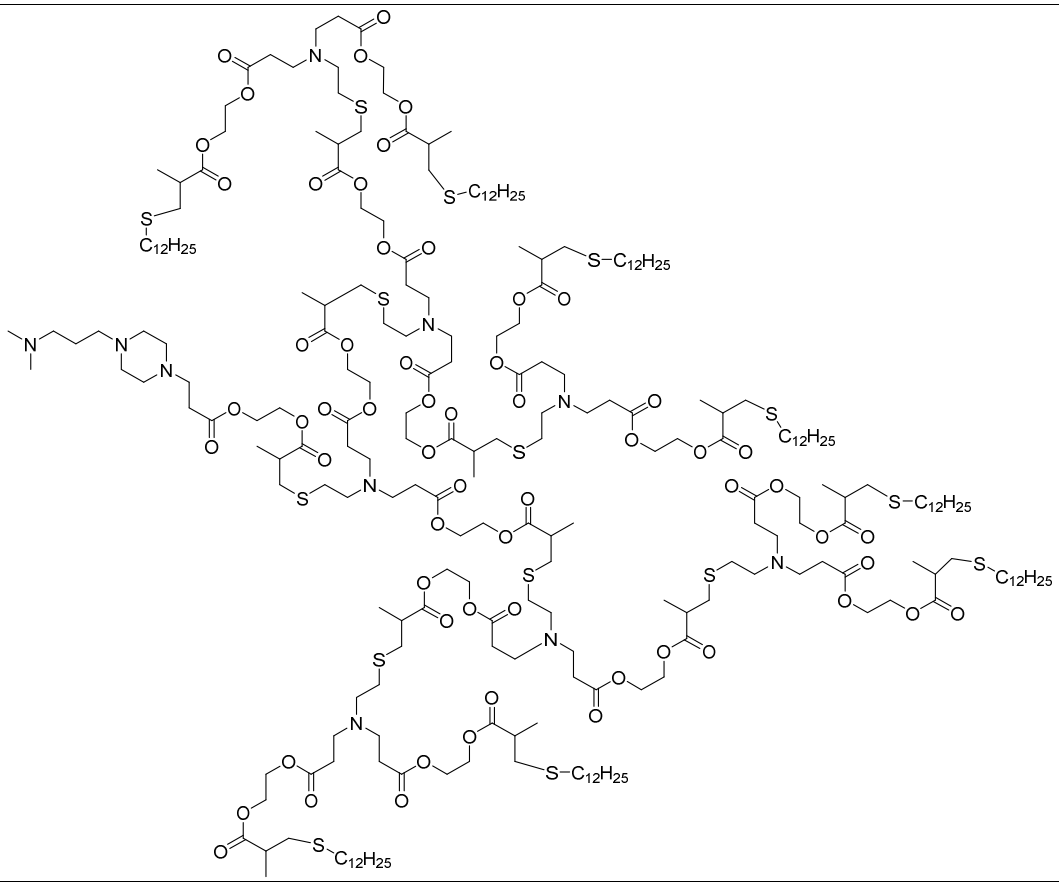
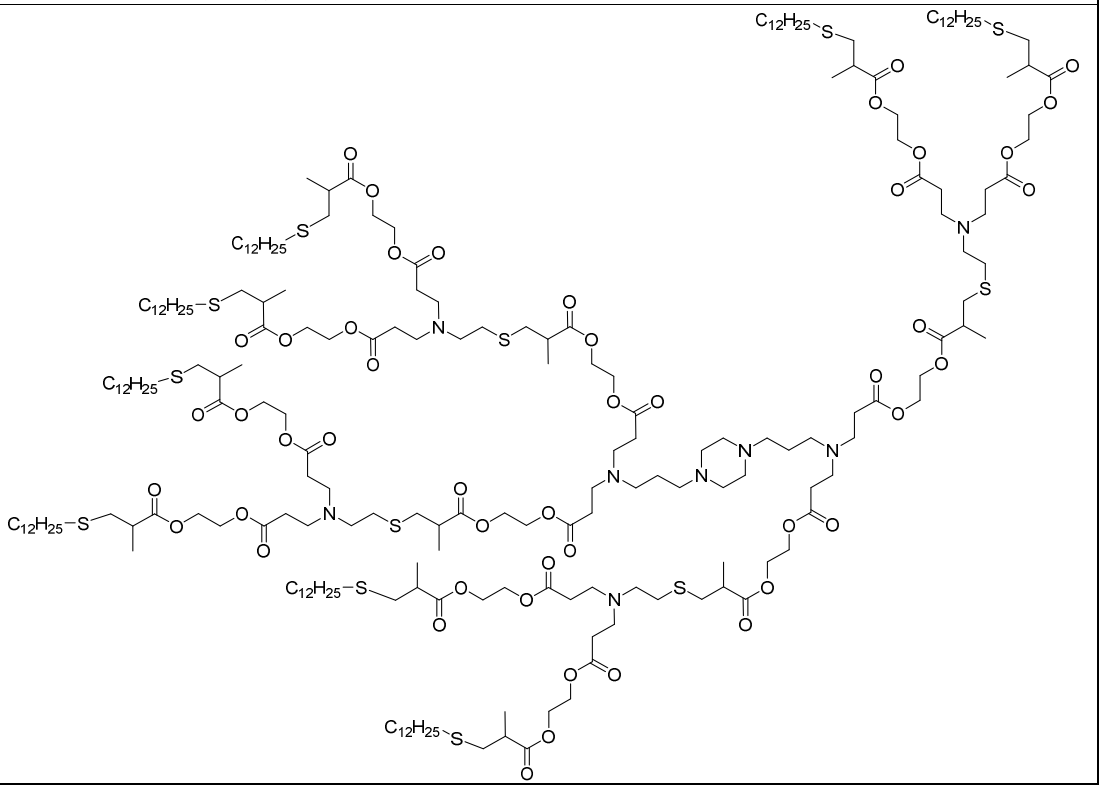
ID #	Structure
2A2-g2-SC12	
2A2-g2-SC8	
2A11-g2-SC12	

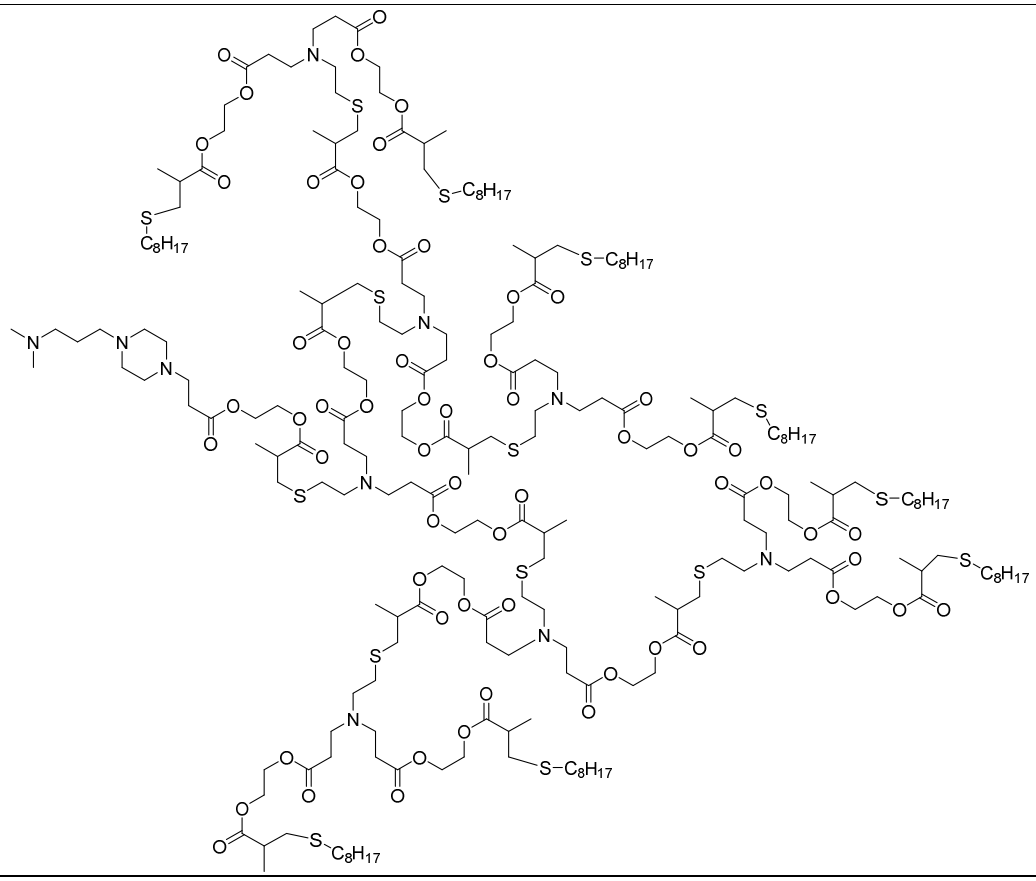
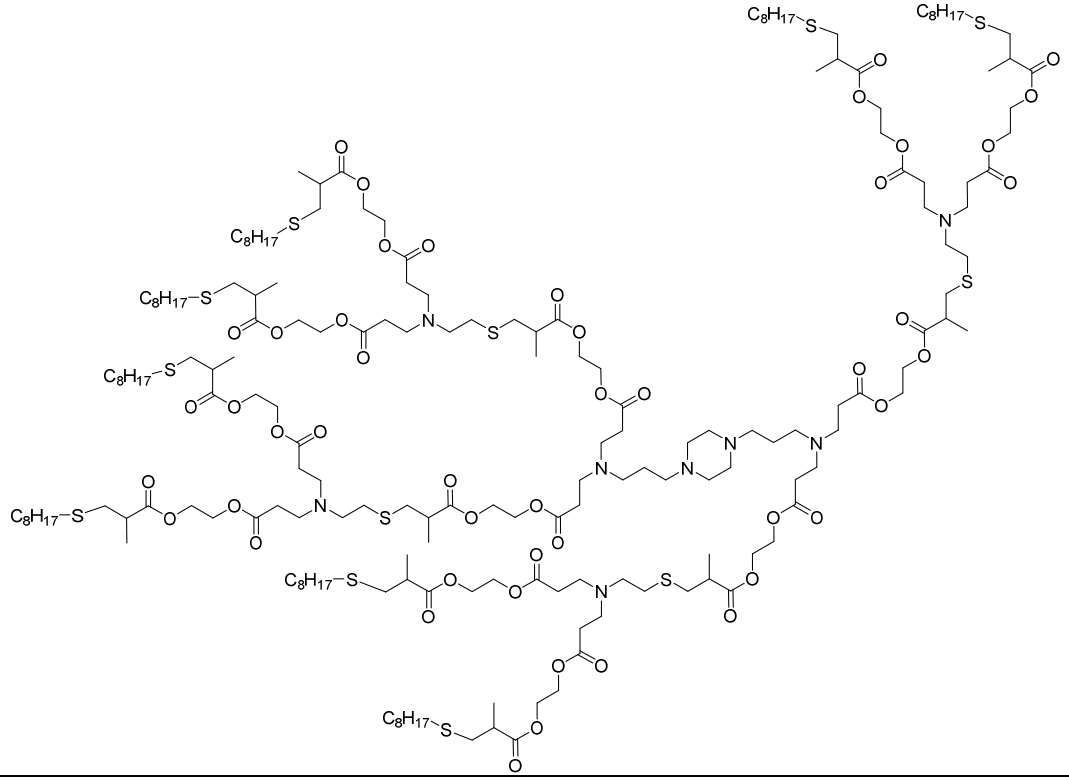
ID #	Structure
2A11-g2-SC8	
3A3-g2-SC12	

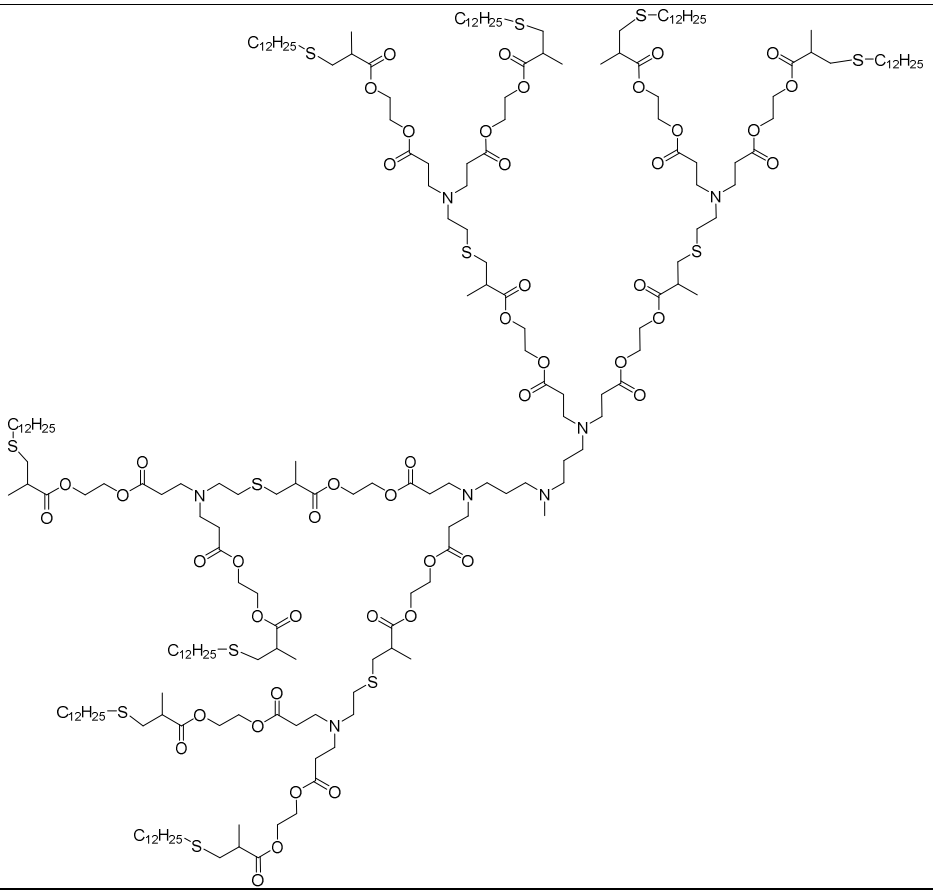
ID #	Structure
3A3-g2-SC8	
3A5-g2-SC12	

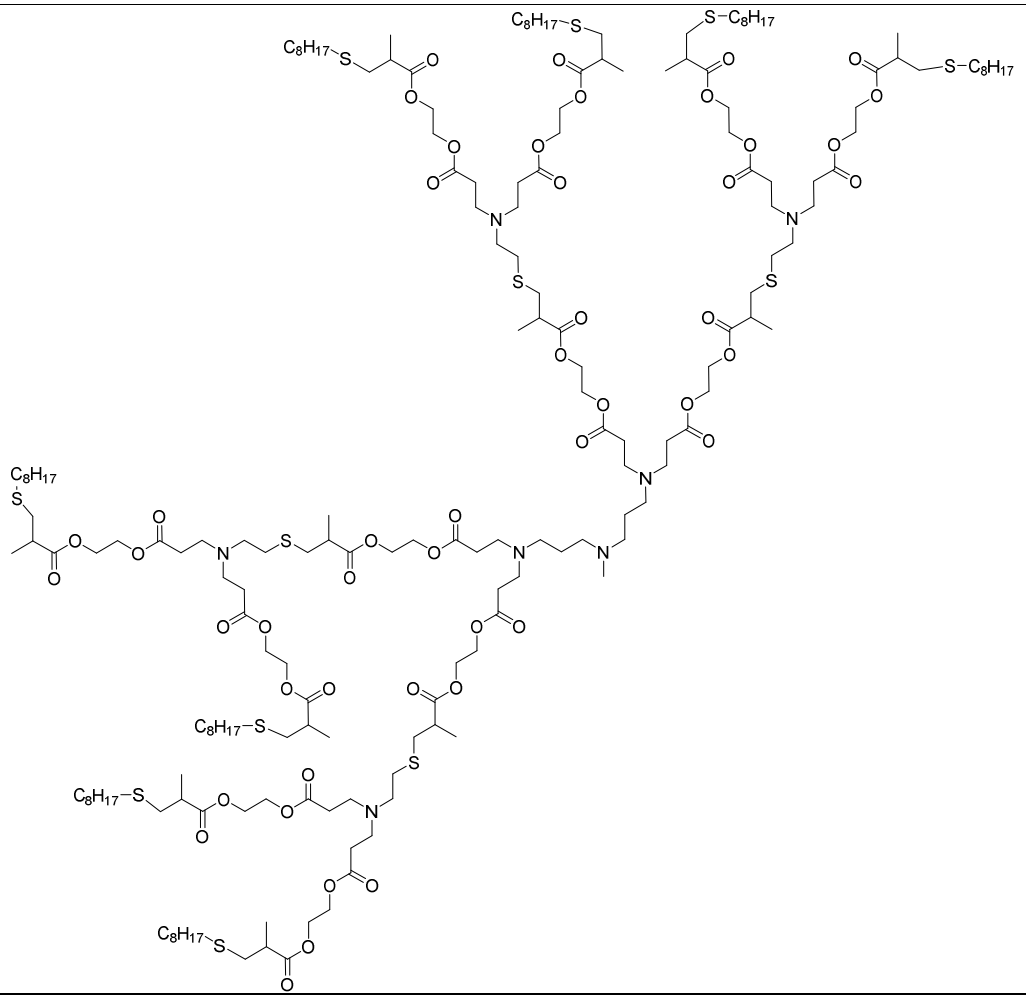
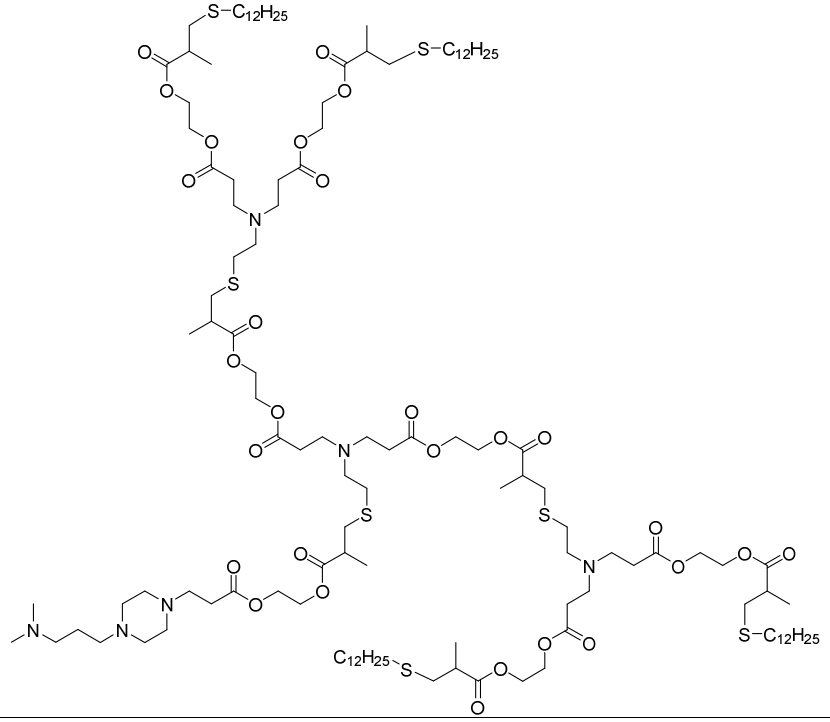
ID #	Structure
2A11-g3-SC12	

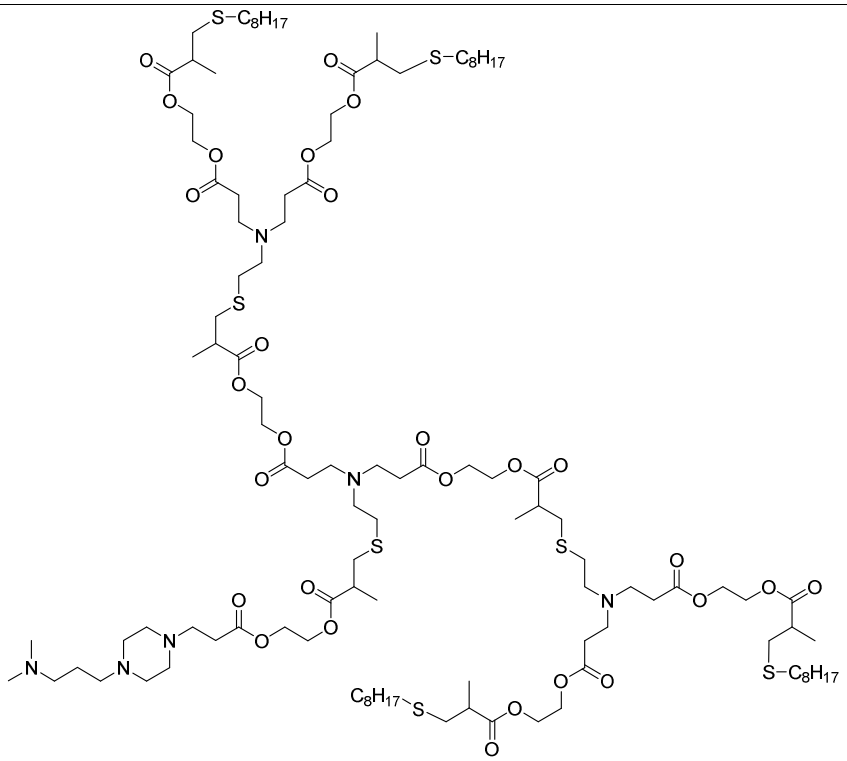
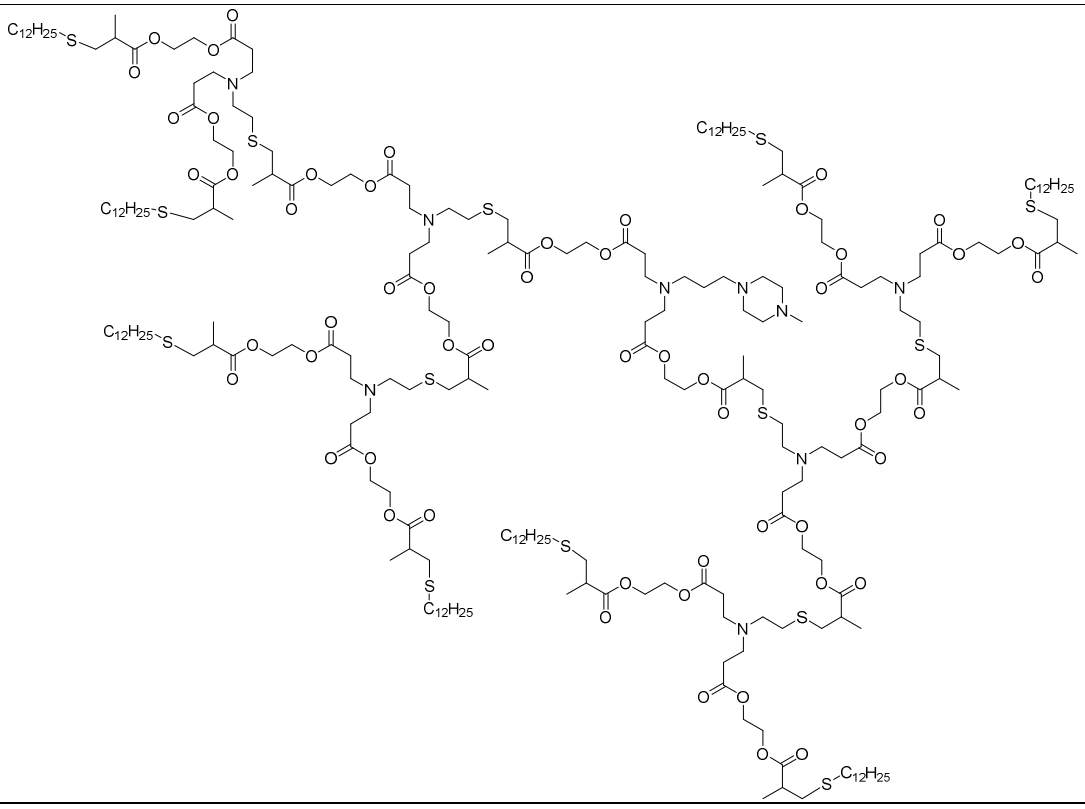
ID #	Structure
2A11-g3-SC8	

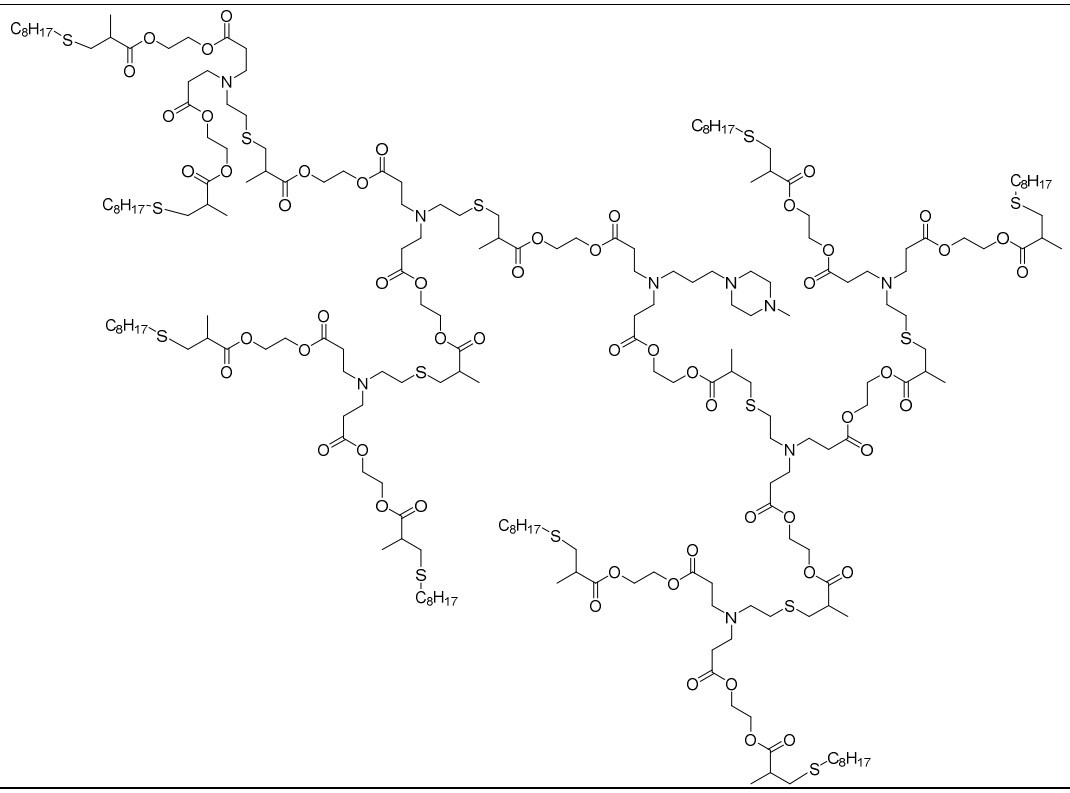
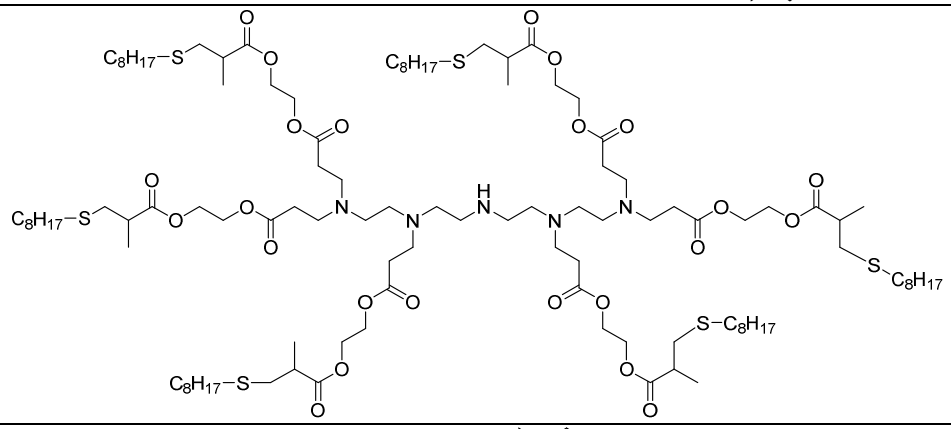
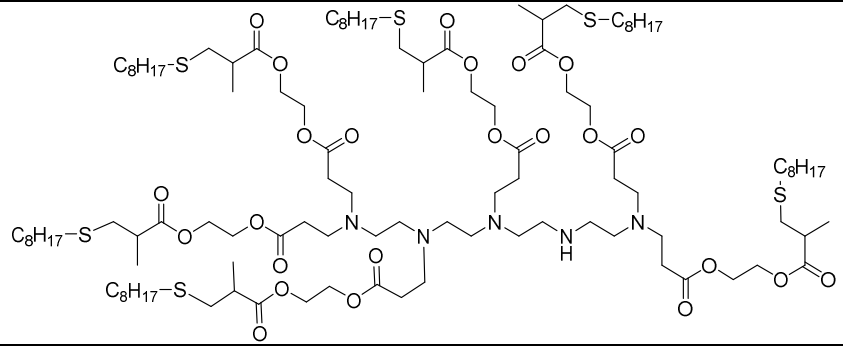
ID #	Structure
1A2-g4-SC12	
4A1-g2-SC12	

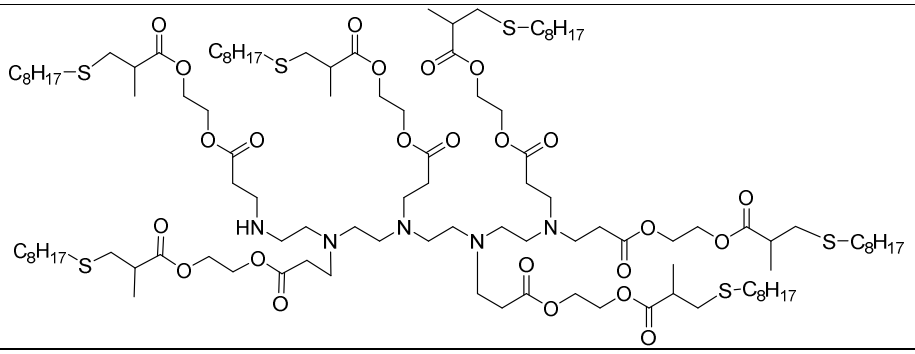
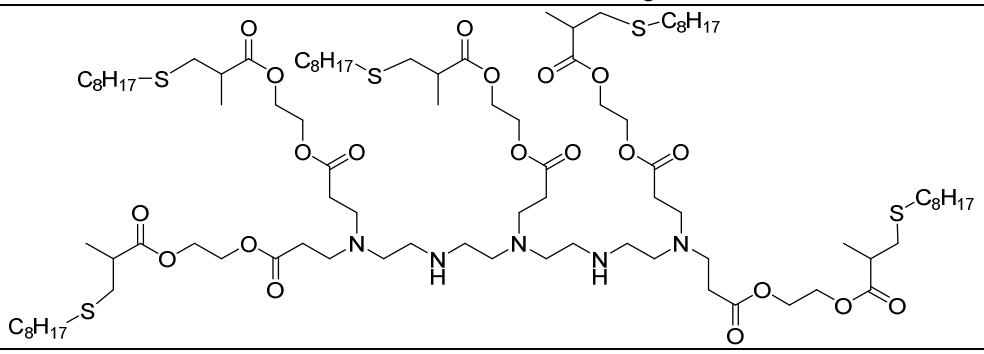
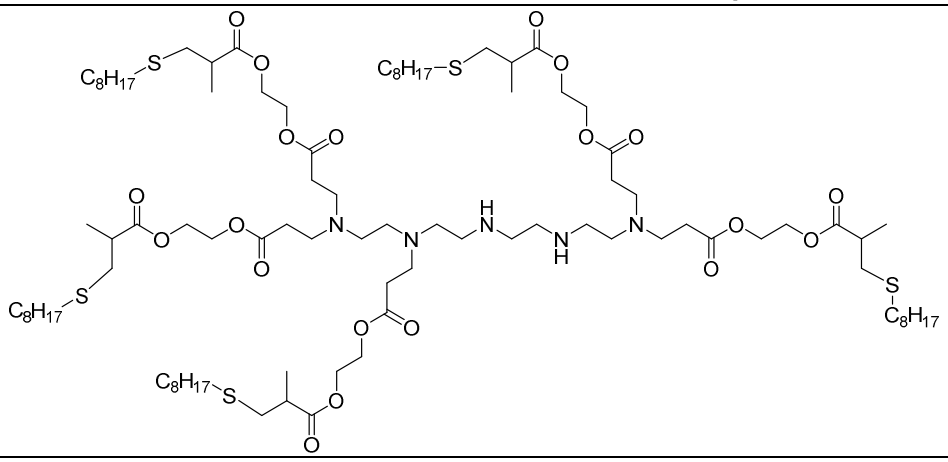
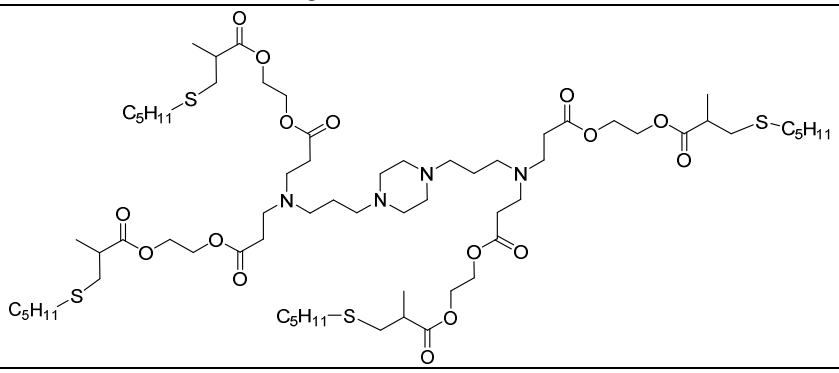
ID #	Structure
1A2-g4-SC8	
4A1-g2-SC8	

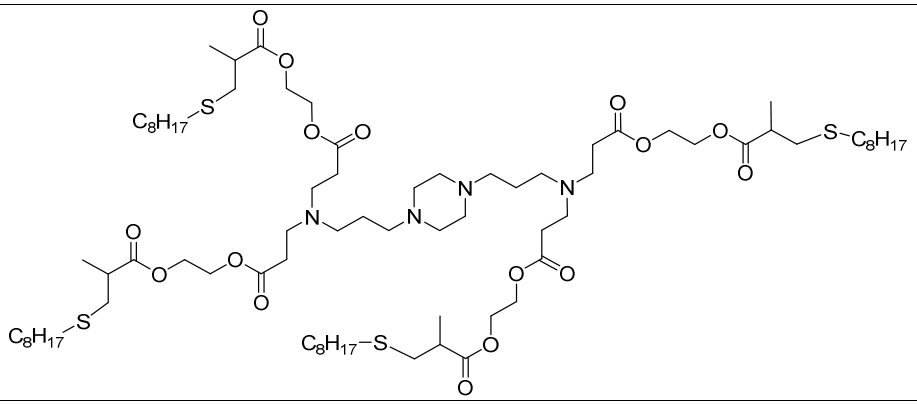
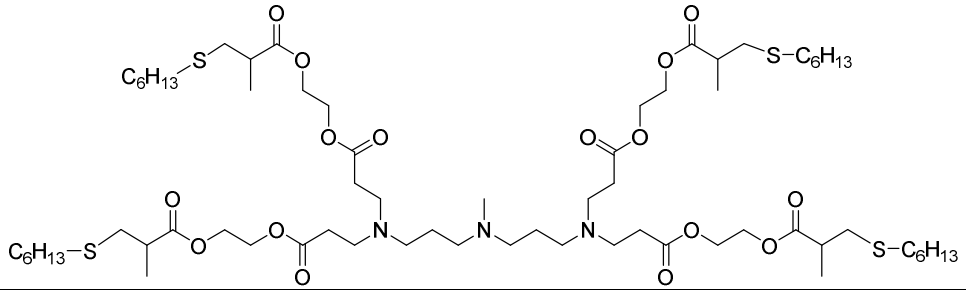
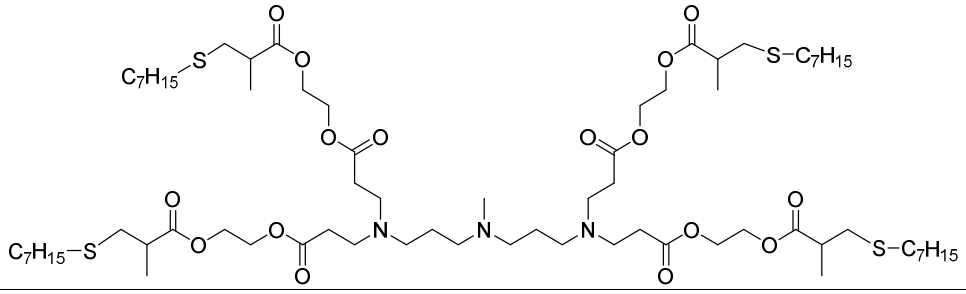
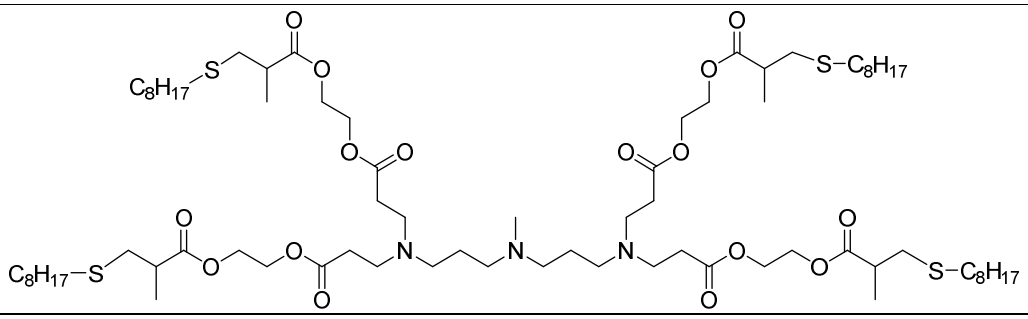
ID #	Structure
4A3-g2-SC12	

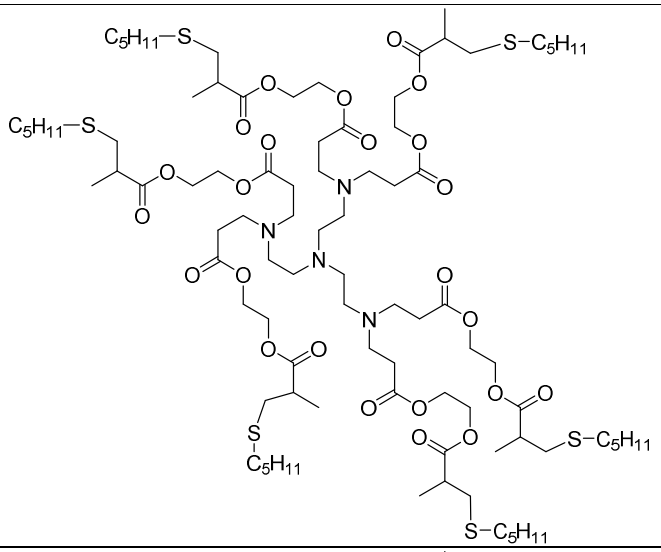
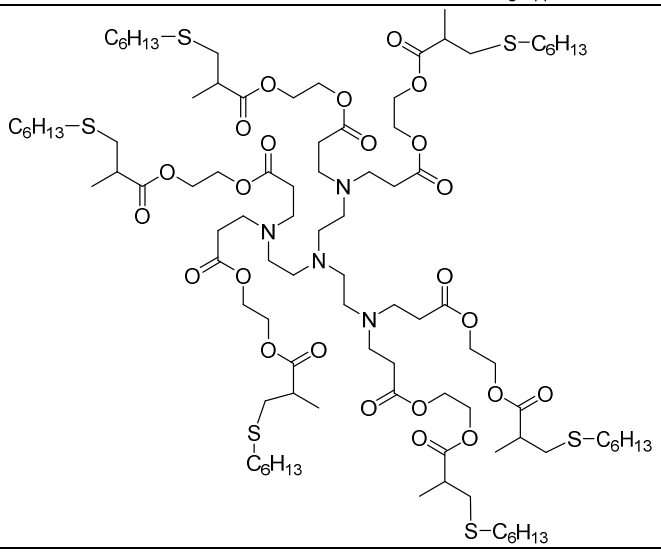
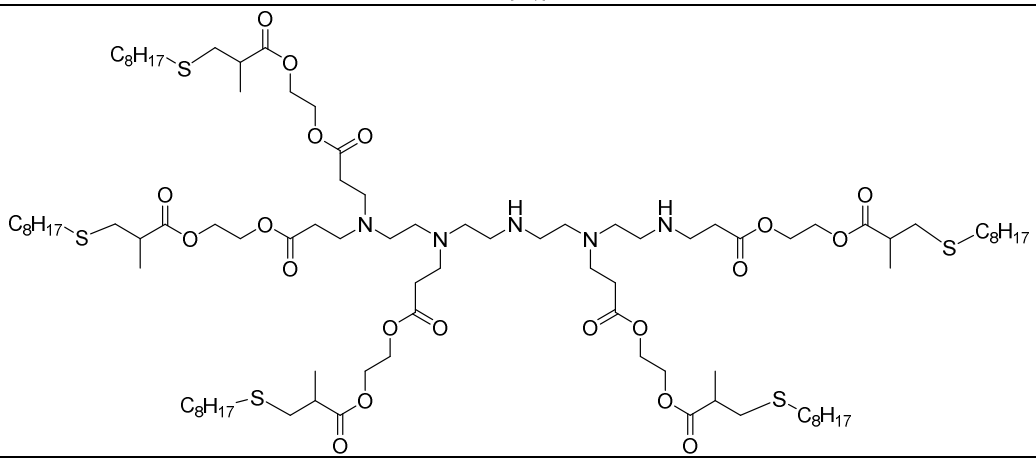
ID #	Structure
4A3-g2-SC8	
1A2-g3-SC12	

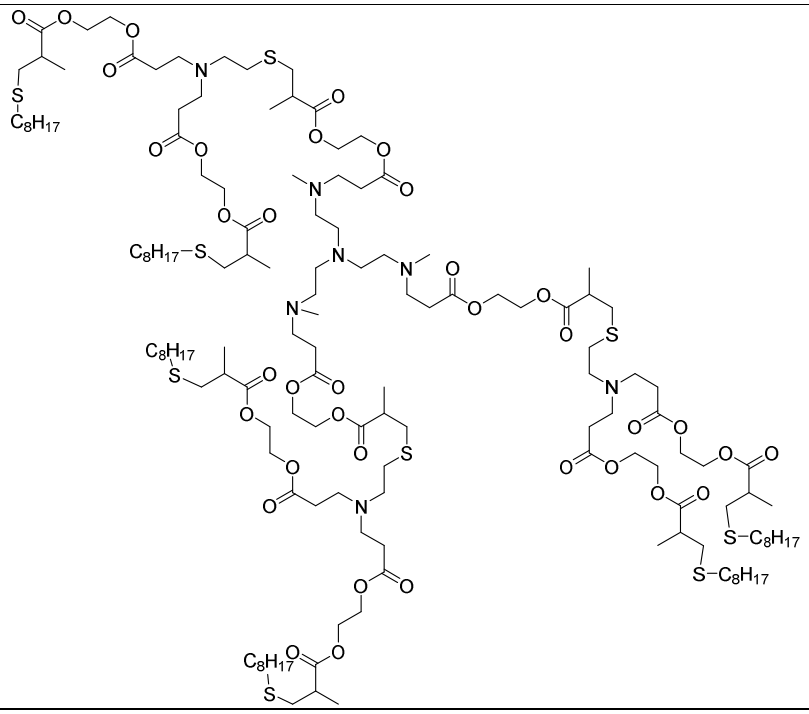
ID #	Structure
1A2-g3-SC8	
2A2-g3-SC12	

ID #	Structure
2A2-g3-SC8	
5A2-4-SC8 (6-arm)	
5A-5-SC8 (6 arm)	

ID #	Structure
5A2-6-SC8 (6 arm)	
5A2-1-SC8 (5-arm)	
5A2-2-SC8	
4A1-SC5	

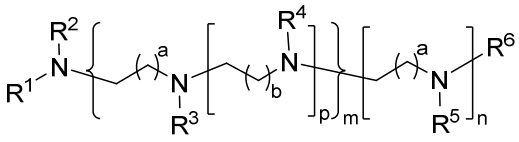
ID #	Structure
4A1-SC8	
4A3-SC6	
4A3-SC7	
4A3-SC8	

ID #	Structure
5A4-2-SC5 (6 arm)	
5A4-2-SC6 (6 arm)	
5A2-4-SC8 (5-arm)	

ID #	Structure
3A5-g2-SC8	

Other Ionizable Cationic Lipids

[00217] In some embodiments of the lipid composition, the cationic lipid comprises a structural formula (D-I'):



(D-I'), wherein:

a is 1 and b is 2, 3, or 4; or, alternatively, b is 1 and a is 2, 3, or 4;

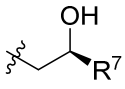
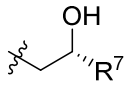
m is 1 and n is 1; or, alternatively, m is 2 and n is 0; or, alternatively, m is 2 and n is 1; and

R^1, R^2, R^3, R^4, R^5, and R^6 are each independently selected from the group consisting of H, $-CH_2CH(OH)R^7$, $-CH(R^7)CH_2OH$, $-CH_2CH_2C(=O)OR^7$, $-CH_2CH_2C(=O)NHR^7$, and $-CH_2R^7$, wherein R^7 is independently selected from C_3-C_{18} alkyl, C_3-C_{18} alkenyl having one C=C double bond, a protecting group for an amino group, $-C(=NH)NH_2$, a poly(ethylene glycol) chain, and a receptor ligand;

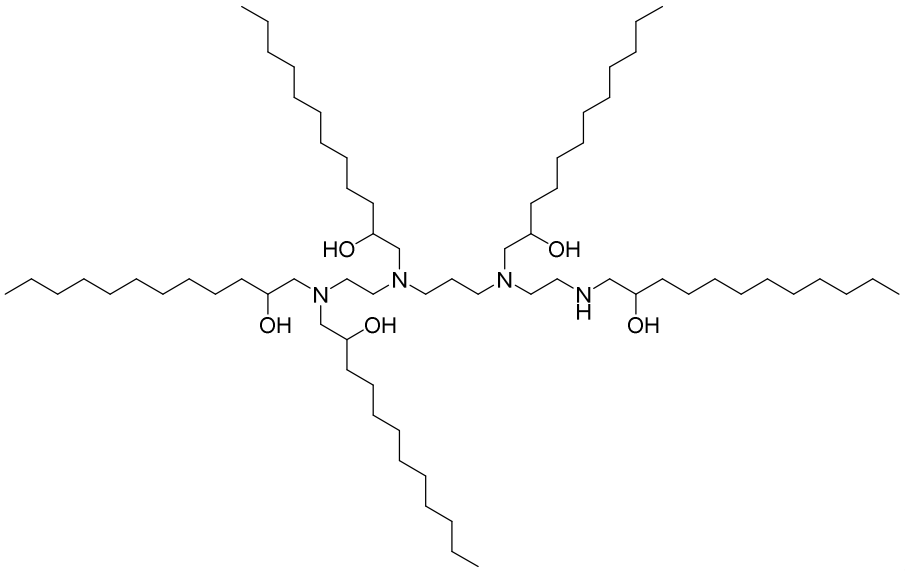
provided that at least two moieties among R^1 to R^6 are independently selected from $-CH_2CH(OH)R^7$, $-CH(R^7)CH_2OH$, $-CH_2CH_2C(=O)OR^7$, $-CH_2CH_2C(=O)NHR^7$, or $-CH_2R^7$, wherein R^7 is independently selected from C_3-C_{18} alkyl or C_3-C_{18} alkenyl having one C=C double bond; and

wherein one or more of the nitrogen atoms indicated in formula (D-I') may be protonated to provide a cationic lipid.

[00218] In some embodiments of the cationic lipid of formula (D-I'), a is 1. In some embodiments of the cationic lipid of formula (D-I'), b is 2. In some embodiments of the cationic lipid of formula (D-I'), m is 1. In some embodiments of the cationic lipid of formula (D-I'), n is 1. In some embodiments of the cationic

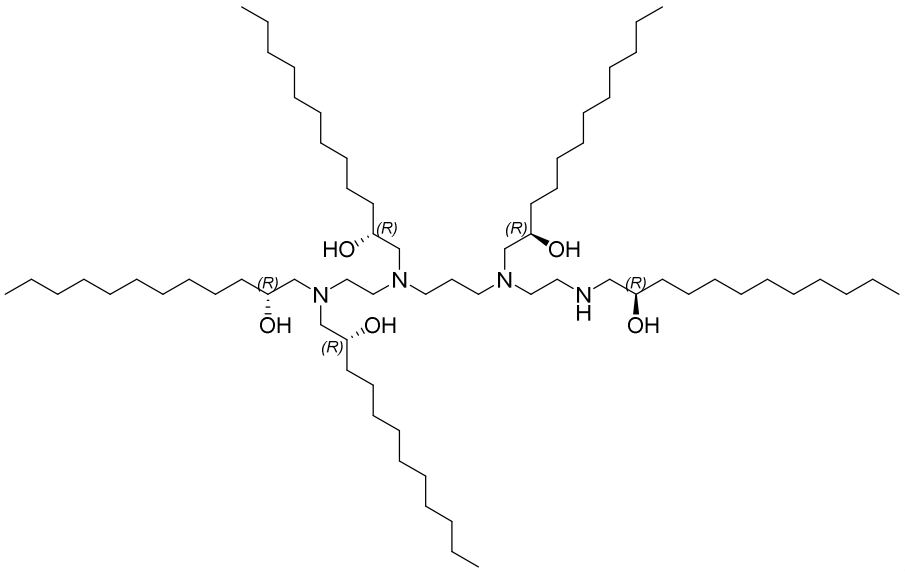
lipid of formula (D-I'), R^1, R^2, R^3, R^4, R^5, and R^6 are each independently H or $-CH_2CH(OH)R^7$. In some embodiments of the cationic lipid of formula (D-I'), R^1, R^2, R^3, R^4, R^5, and R^6 are each independently H or . In some embodiments of the cationic lipid of formula (D-I'), R^1, R^2, R^3, R^4, R^5, and R^6 are each independently H or . In some embodiments of the cationic lipid of formula (D-I'), R^7 is C_3-C_{18} alkyl (e.g., C_6-C_{12} alkyl).

[00219] In some embodiments, the cationic lipid of formula (D-I') is 13,16,20-tris(2-hydroxydodecyl)-13,16,20,23-tetraazapentatricontane-11,25-diol:



.

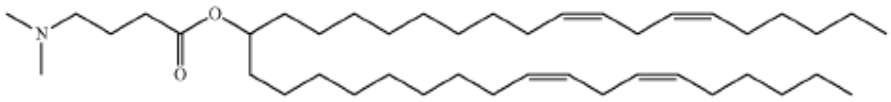
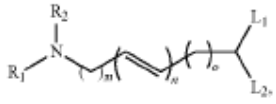
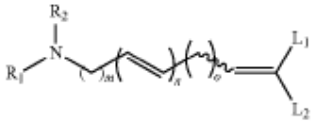
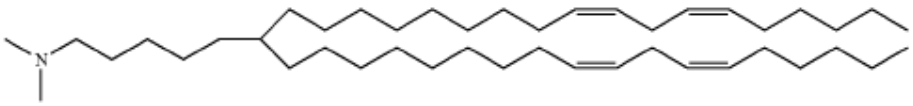
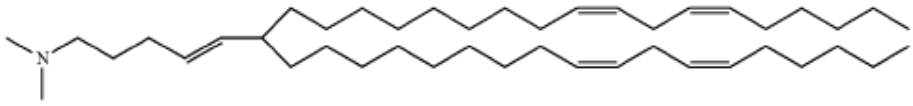
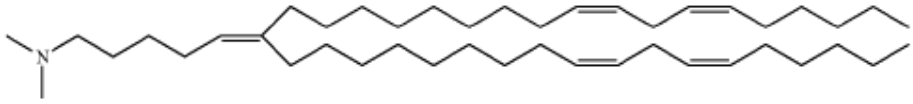
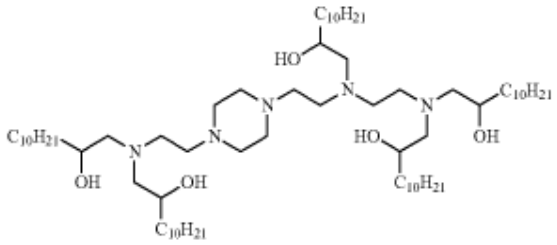
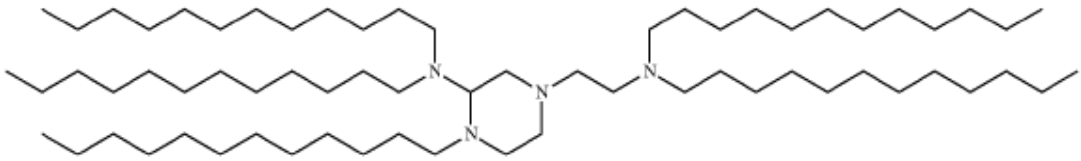
[00220] In some embodiments, the cationic lipid of formula (D-I') is (11*R*,25*R*)-13,16,20-tris((*R*)-2-hydroxydodecyl)-13,16,20,23-tetraazapentatricontane-11,25-diol:

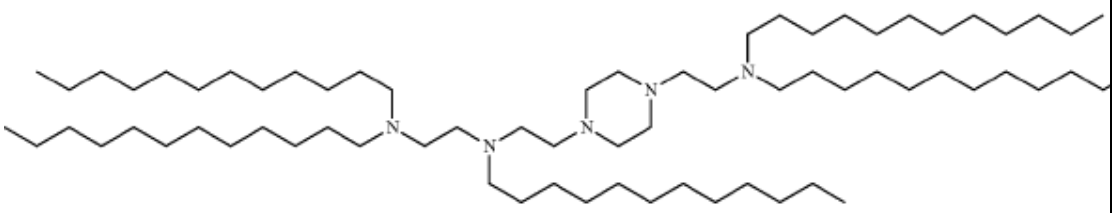
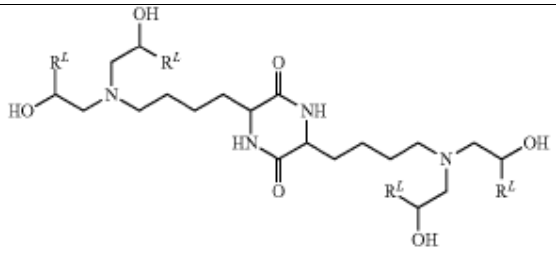
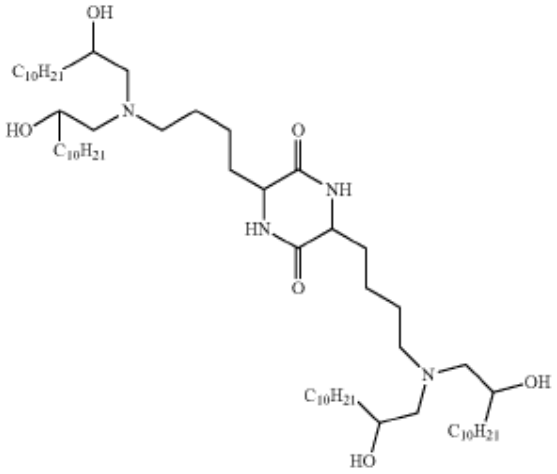
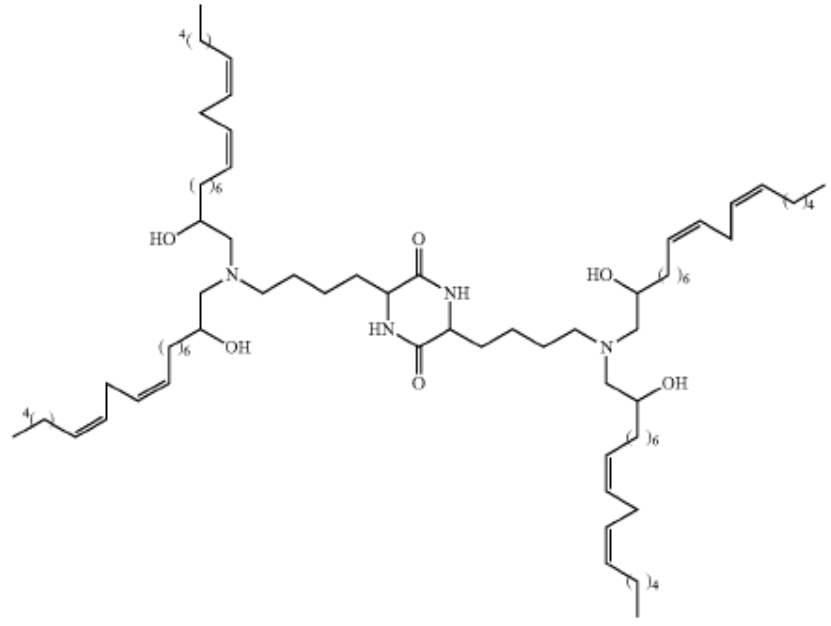


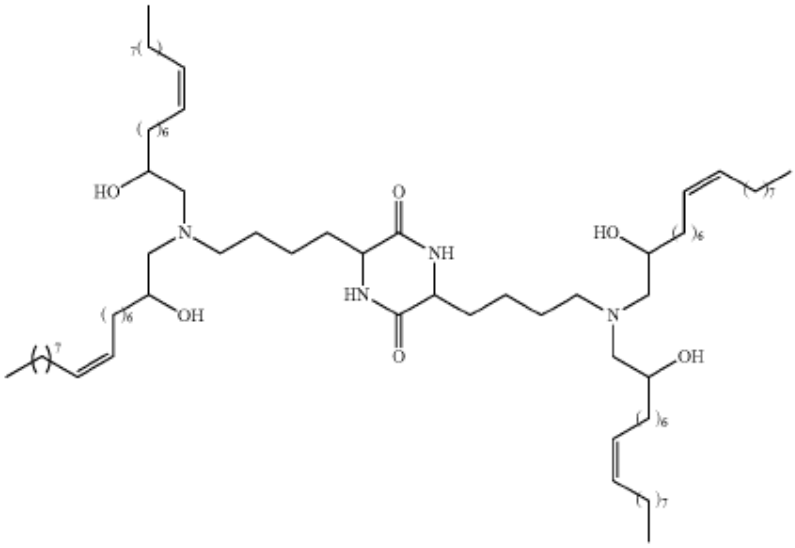
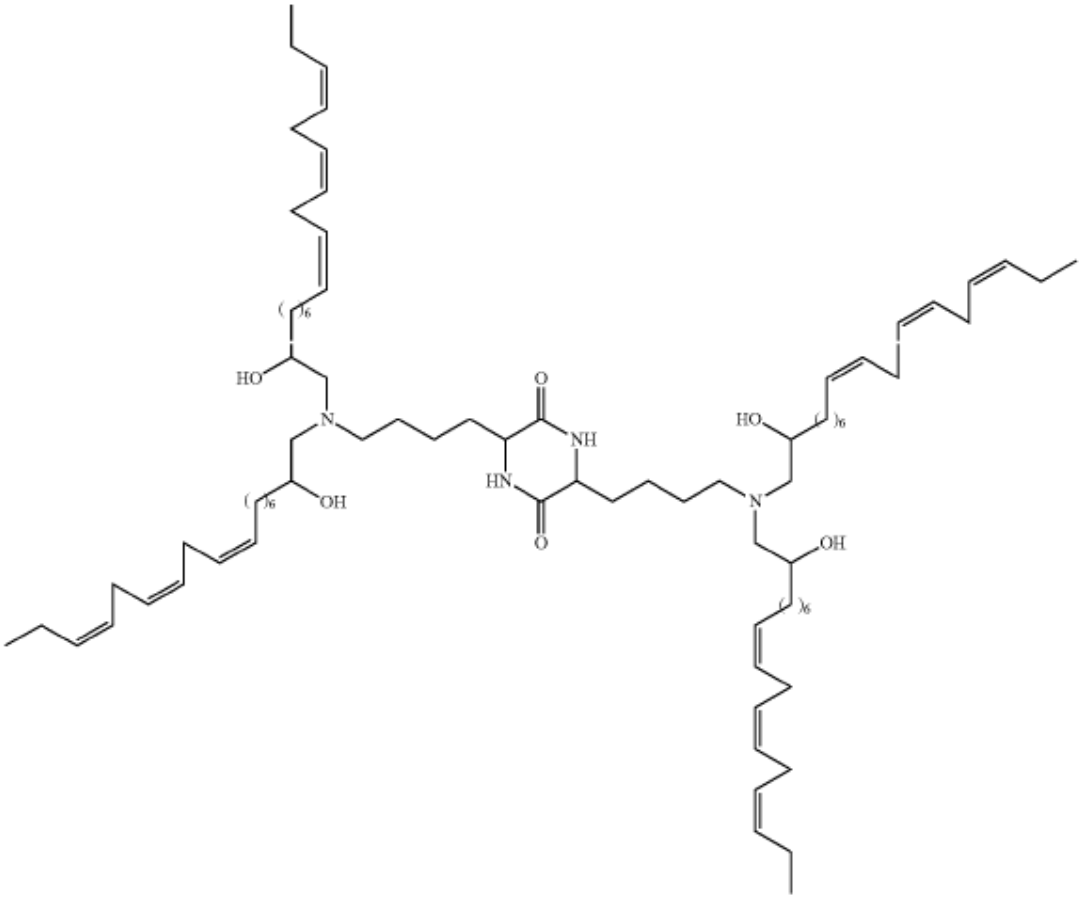
.

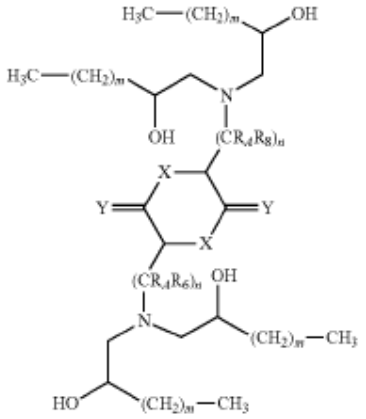
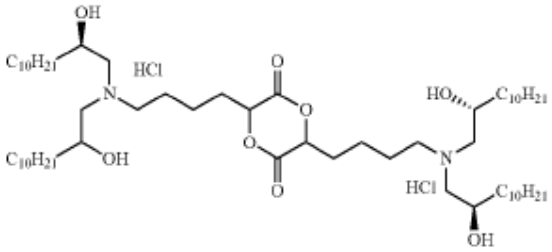
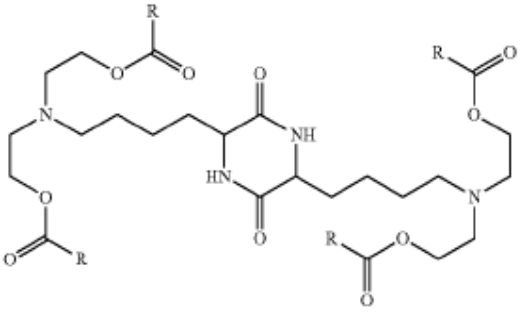
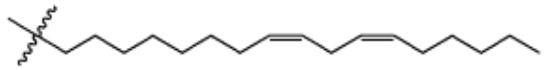
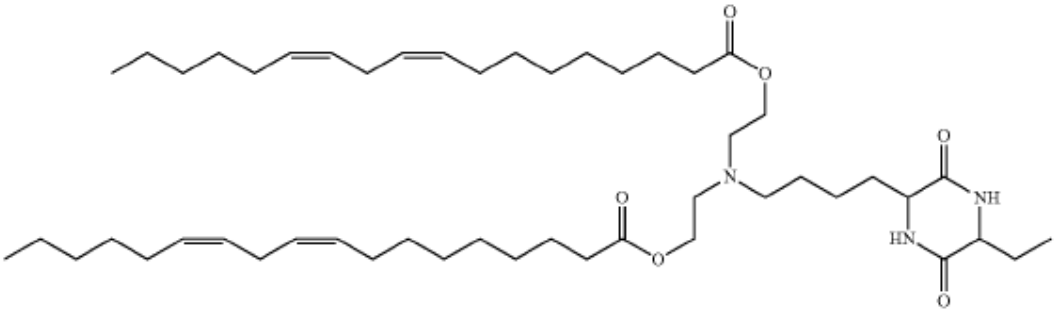
[00221] Additional cationic lipids that can be used in the compositions and methods of the present application include those cationic lipids as described in J. McClellan, M. C. King, Cell 2010, 141, 210-217, and International Patent Publication WO 2010/144740, WO 2013/149140, WO 2016/118725, WO 2016/118724, WO 2013/063468, WO 2016/205691, WO 2015/184256, WO 2016/004202, WO 2015/199952, WO 2017/004143, WO 2017/075531, WO 2017/117528, WO 2017/049245, WO 2017/173054 and WO 2015/095340, which are incorporated herein by reference for all purposes. Exemples of those ionizable cationic lipids include but are not limited to those as shown in **Table 7**.

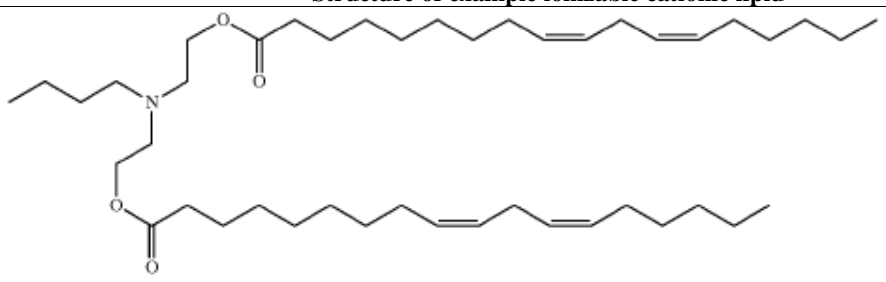
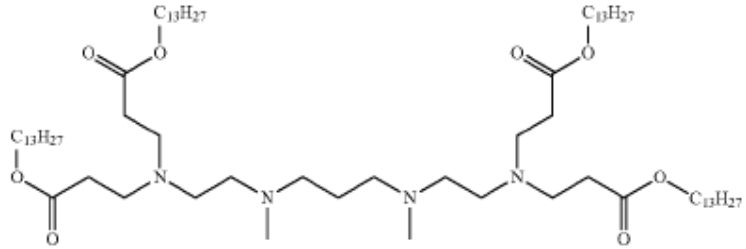
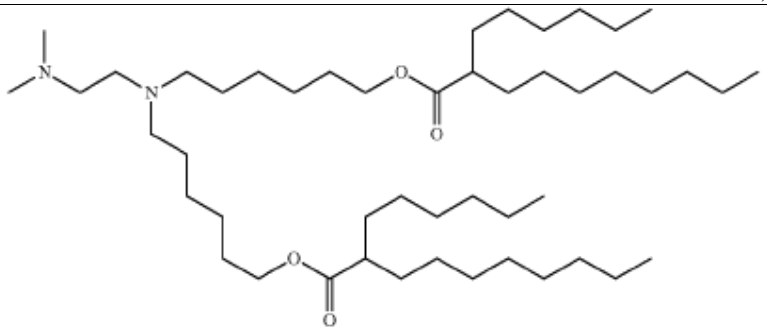
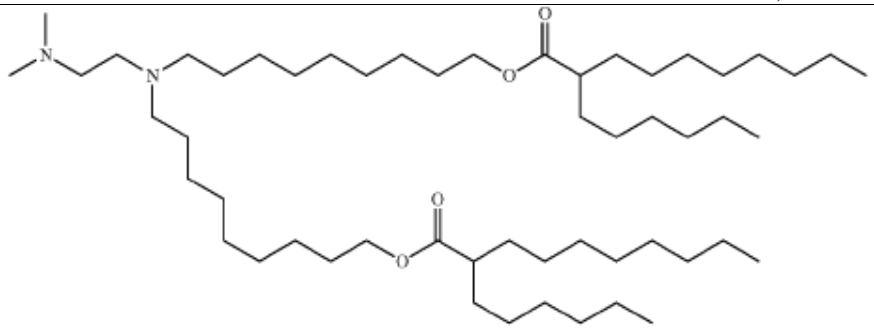
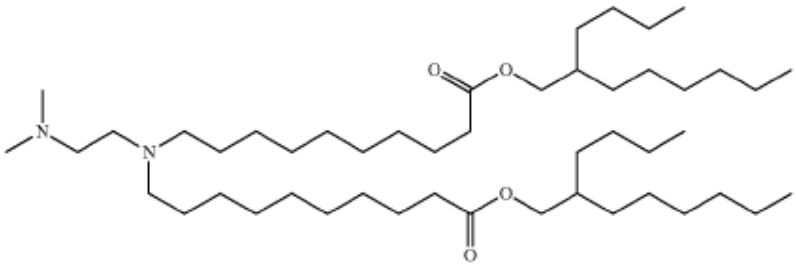
Table 7: Example ionizable cationic lipids

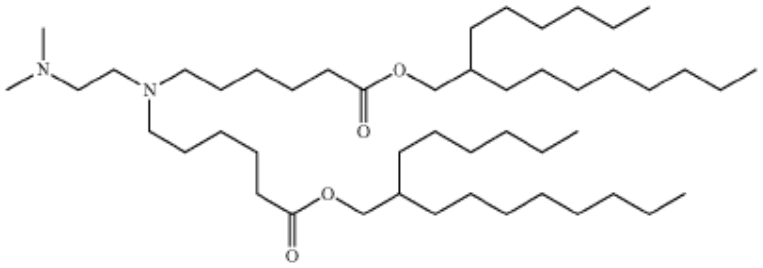
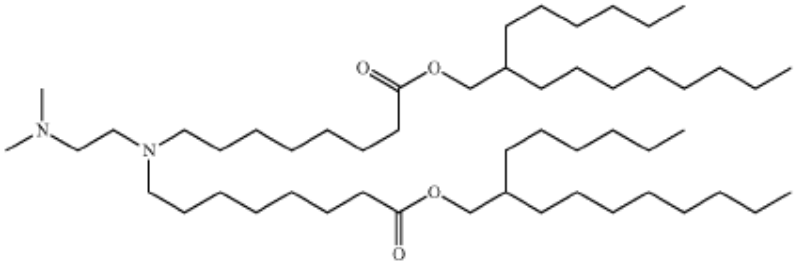
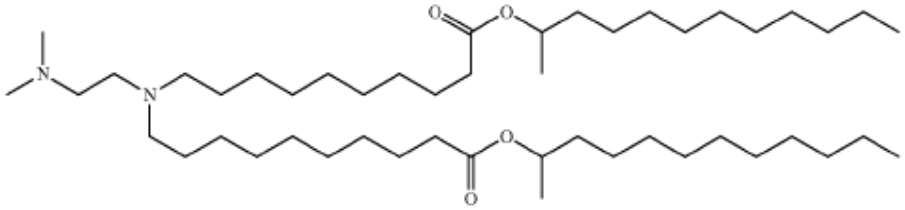
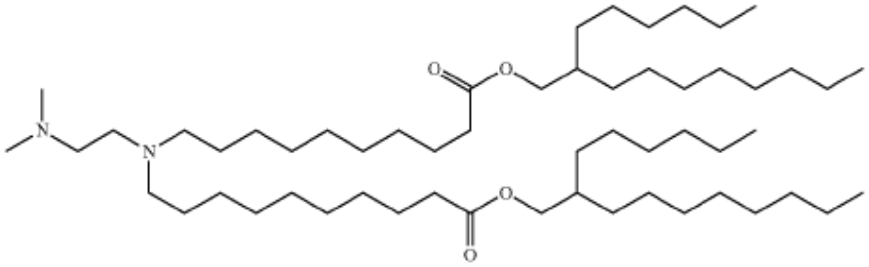
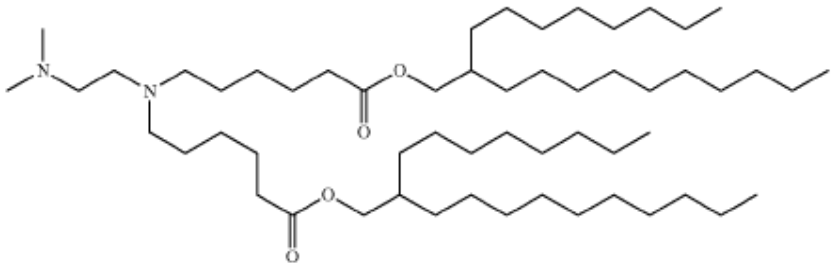
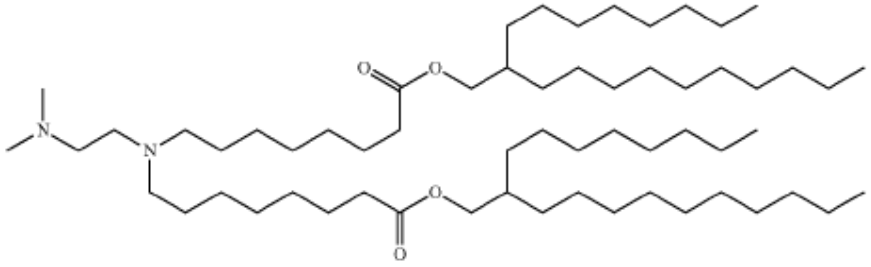
#	Structure of example ionizable cationic lipid
1	 ;
2	 ;
3	 ;
4	 (HGT-5000) ;
5	 (HGT-5001) ;
6	 (HGT-5002) ;
7	 ;
8	 ;

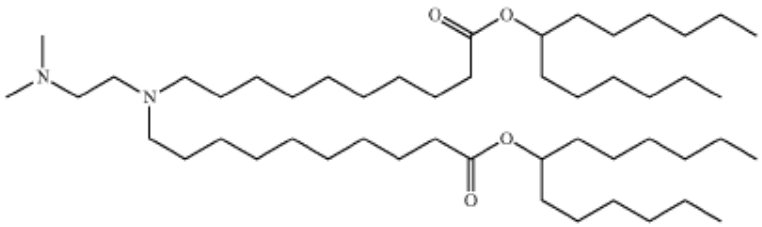
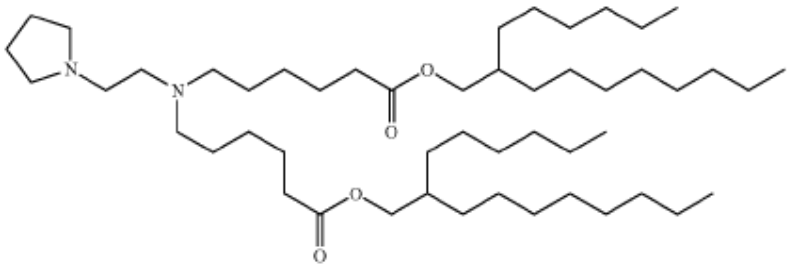
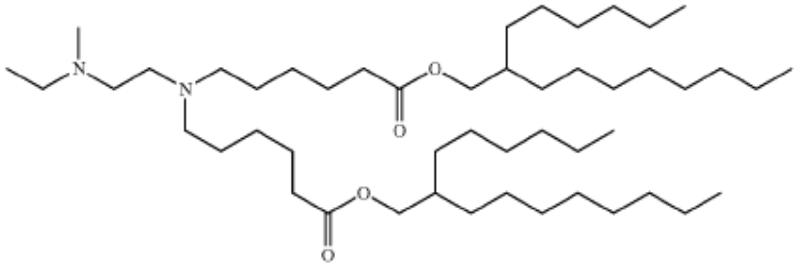
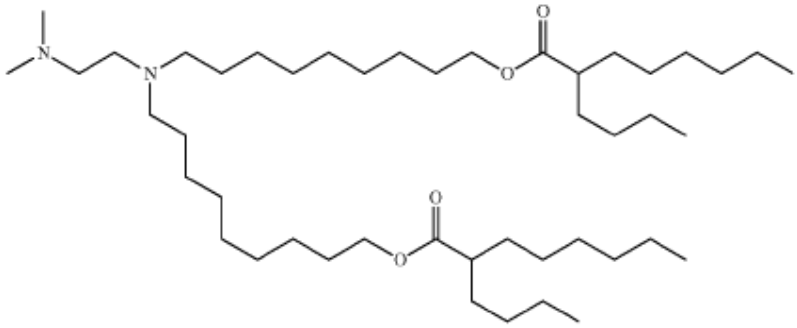
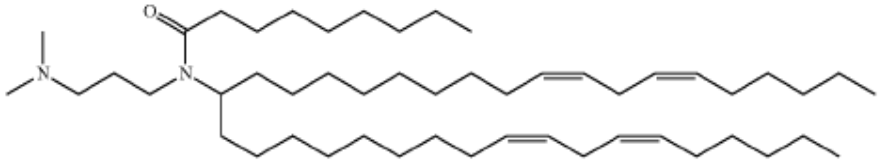
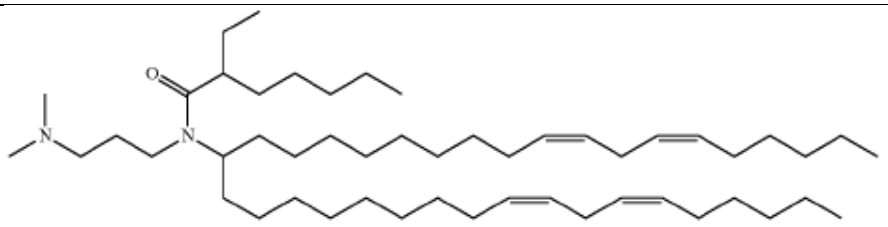
#	Structure of example ionizable cationic lipid
9	
10	 ;
11	 ;
12	 ;

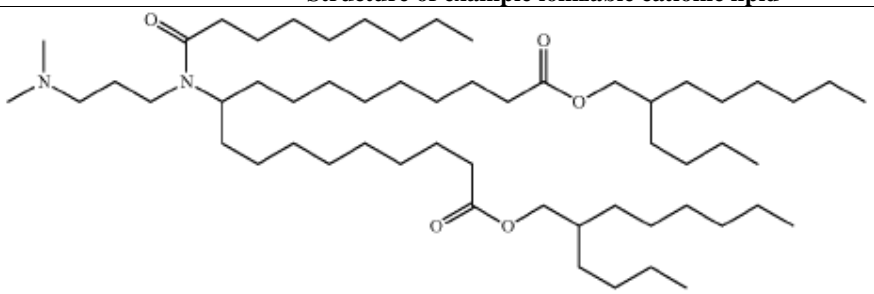
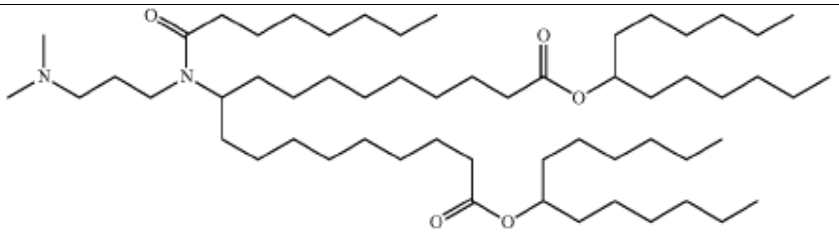
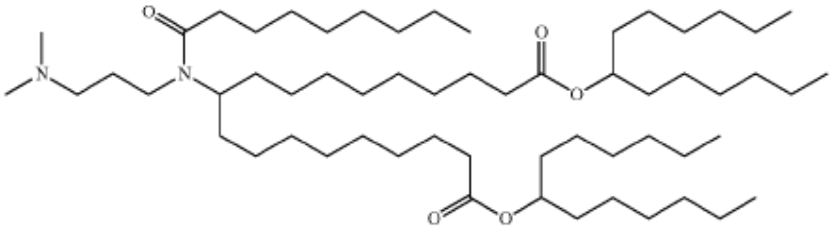
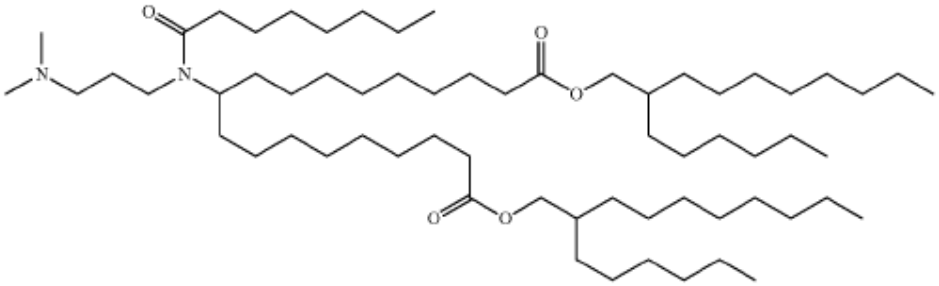
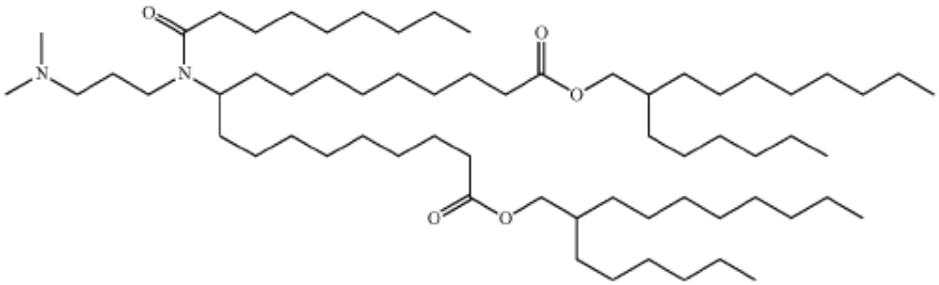
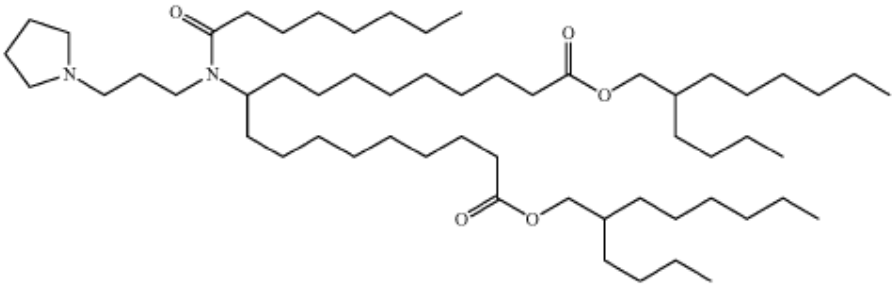
#	Structure of example ionizable cationic lipid
13	 ;
14	 ;

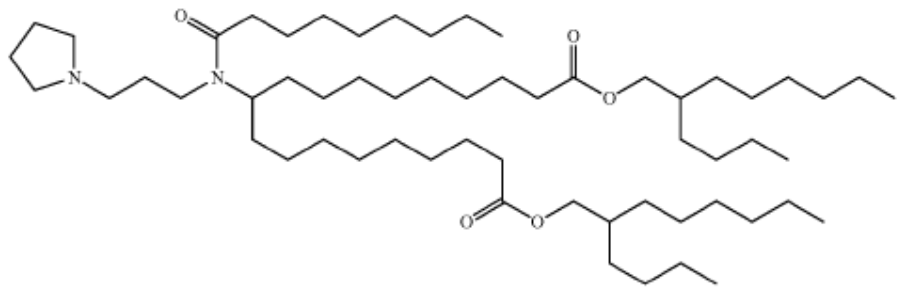
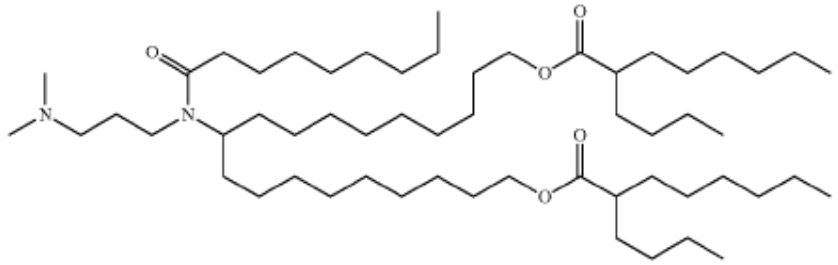
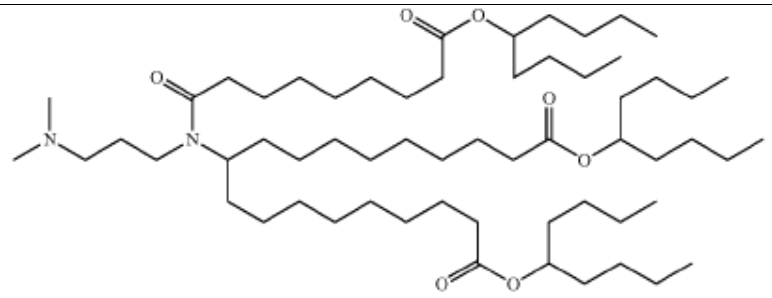
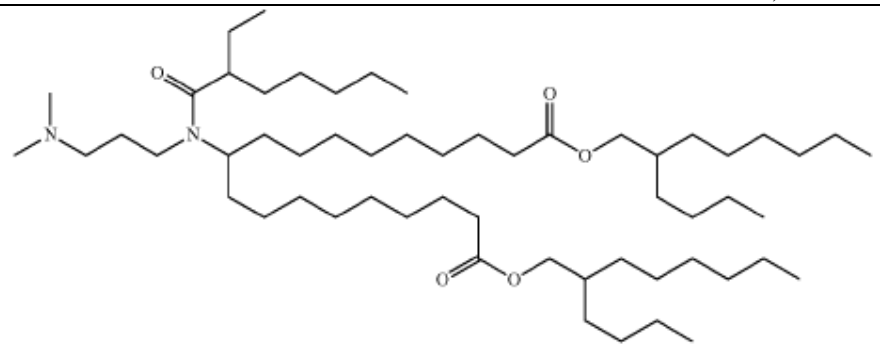
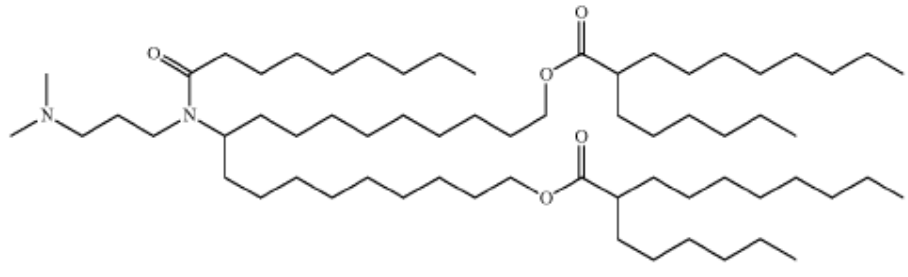
#	Structure of example ionizable cationic lipid
15	 ;
16	(Target 23)  ;
17	 ;
18	R =  ;
19	 ;

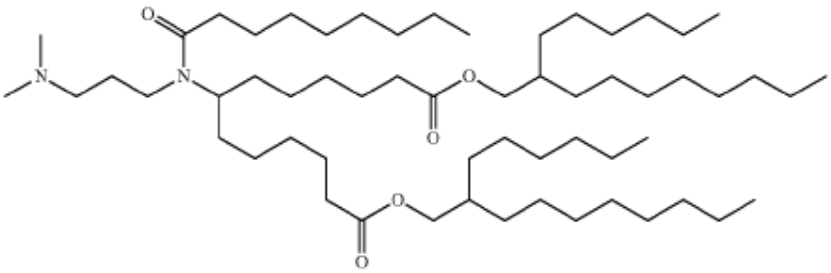
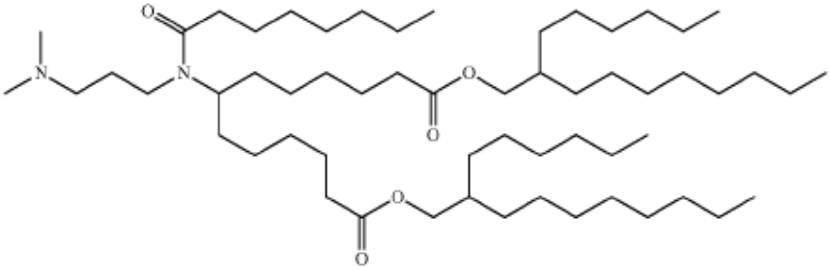
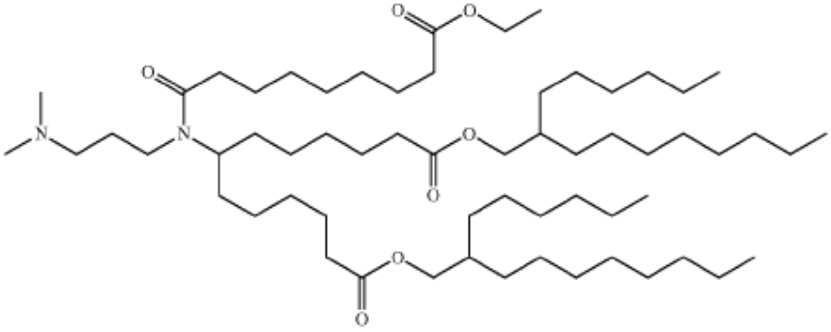
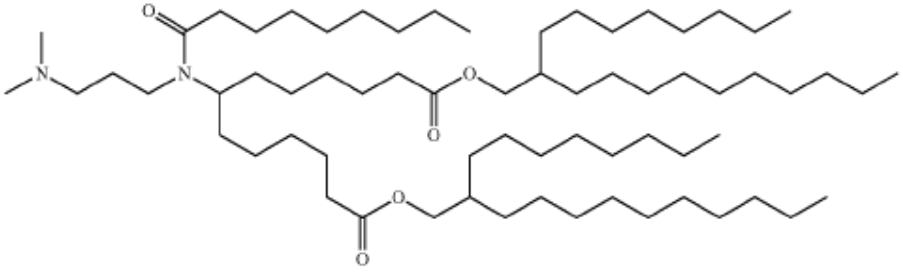
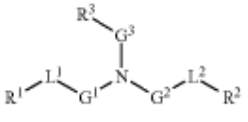
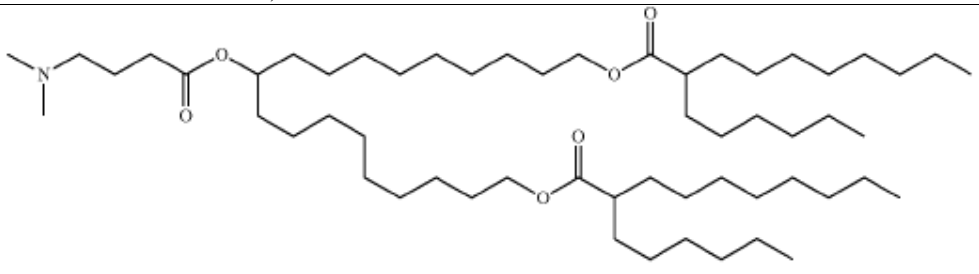
#	Structure of example ionizable cationic lipid
20	 ;
21	 ;
22	 ;
23	 ;
24	 ;

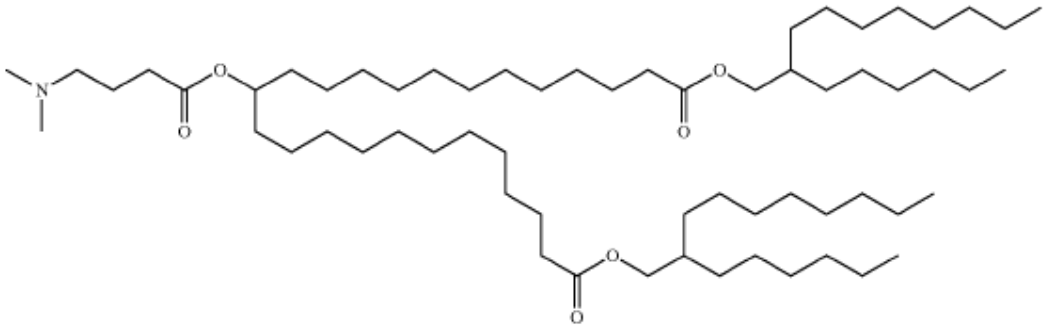
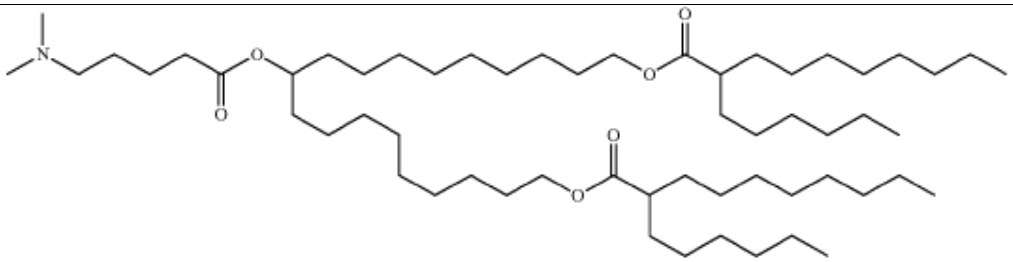
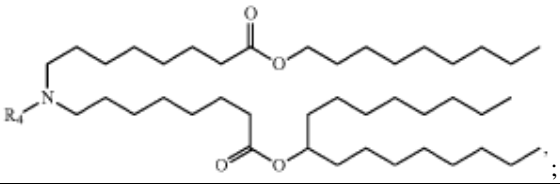
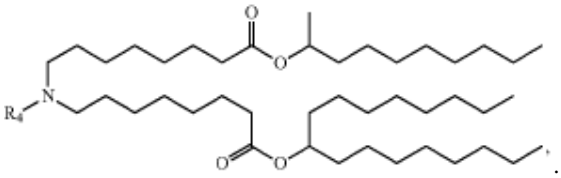
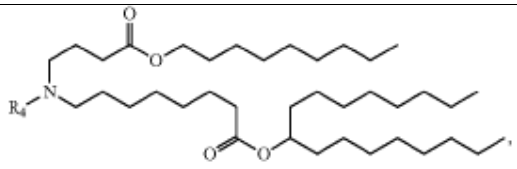
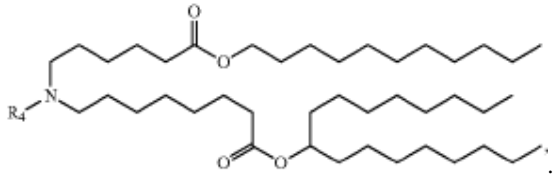
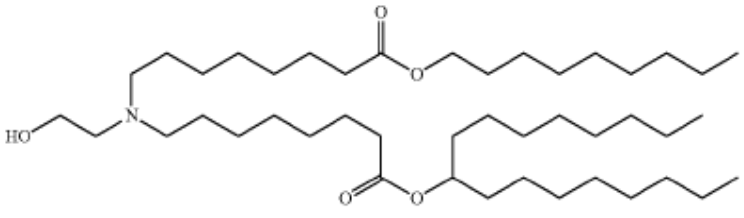
#	Structure of example ionizable cationic lipid
25	 ;
26	 ;
27	 ;
28	 ;
29	 ;
30	 ;

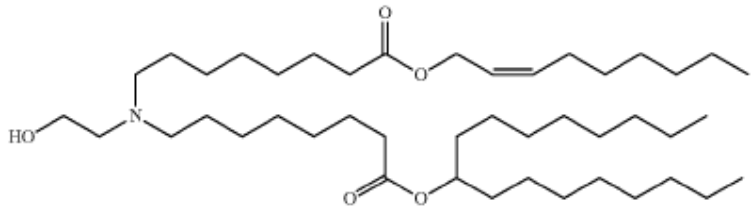
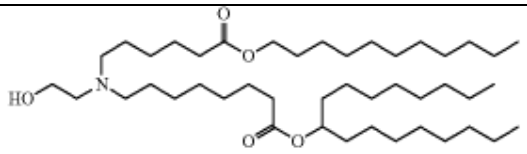
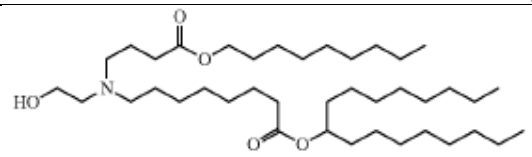
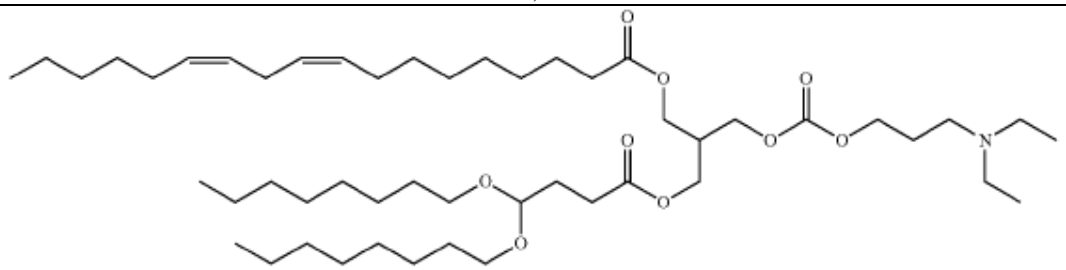
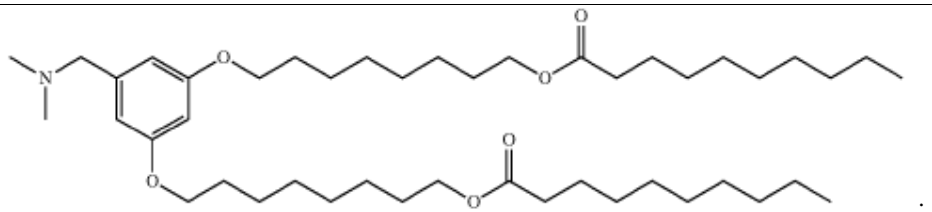
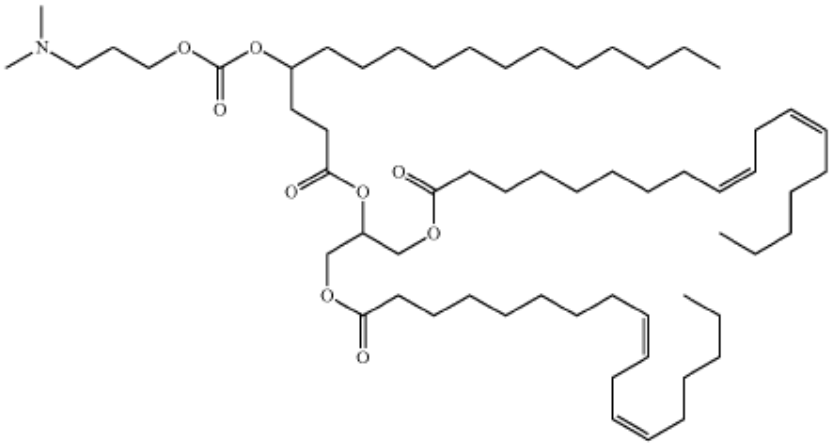
#	Structure of example ionizable cationic lipid
31	 ;
32	 ;
33	 ;
34	 ;
35	 ;
36	 ;

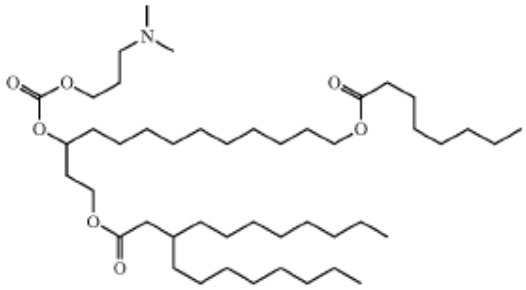
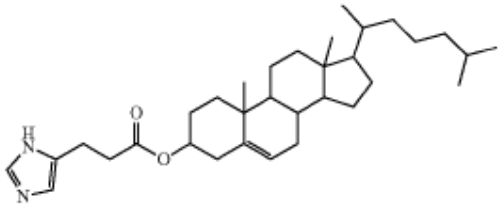
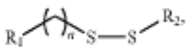
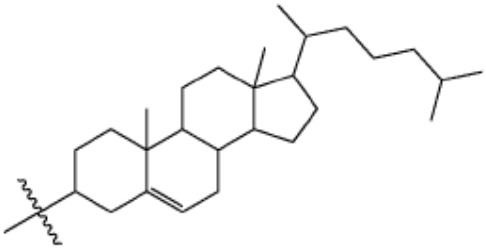
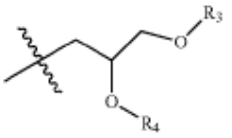
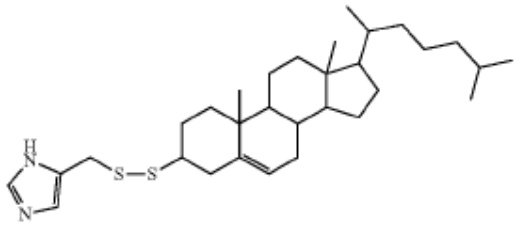
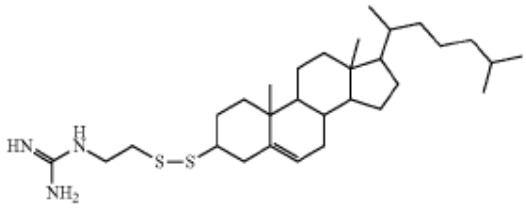
#	Structure of example ionizable cationic lipid
37	 ;
38	 ;
39	 ;
40	 ;
41	 ;
42	 ;

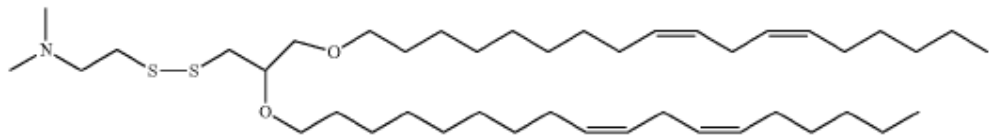
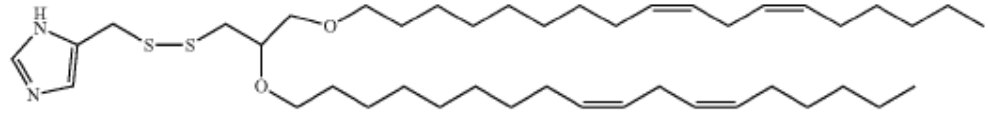
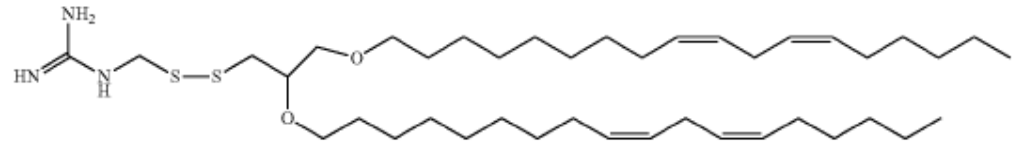
#	Structure of example ionizable cationic lipid
43	 ;
44	 ;
45	 ;
46	 ;
47	 ;

#	Structure of example ionizable cationic lipid
48	 ;
49	 ;
50	 ;
51	 ;
52	 ;
53	 ;

#	Structure of example ionizable cationic lipid
54	 ;
55	 ;
56	 ;
57	 ;
58	 ;
59	 ;
60	 ;

#	Structure of example ionizable cationic lipid
61	 ;
62	 ;
63	 ;
64	 ;
65	 ;
66	 ;

#	Structure of example ionizable cationic lipid
67	 ;
68	 ;
69	 ;
70	 ;
71	 ;
72	(HGT4001)  ;
73	 (HGT4002) ;

#	Structure of example ionizable cationic lipid
74	 (HGT4003) ;
75	 (HGT4004) ;
76	 (HGT4005

[00222] In some embodiments of the lipid composition of the present application, the ionizable cationic lipid is present in an amount from about from about 20 to about 23. In some embodiments, the molar percentage is from about 20, 20.5, 21, 21.5, 22, 22.5, to about 23 or any range derivable therein. In other embodiments, the molar percentage is from about 7.5 to about 20. In some embodiments, the molar percentage is from about 7.5, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, to about 20 or any range derivable therein.

[00223] In some embodiments of the lipid composition of the present application, said lipid composition comprises said ionizable cationic lipid at a molar percentage from about 5% to about 30%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said ionizable cationic lipid at a molar percentage from about 10% to about 25%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said ionizable cationic lipid at a molar percentage from about 15% to about 20%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said ionizable cationic lipid at a molar percentage from about 10% to about 20%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said ionizable cationic lipid at a molar percentage from about 20% to about 30%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said ionizable cationic lipid at a molar percentage of at least (about) 5%, at least (about) 10%, at least (about) 15%, at least (about) 20%, at least (about) 25%, or at least (about) 30%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said ionizable cationic lipid at a molar percentage of at most (about) 5%, at most (about) 10%, at most (about) 15%, at most (about) 20%, at most (about) 25%, or at most (about) 30%.

[00224] In some embodiments, a (e.g., mass or weight) ratio of said ionizable cationic lipid to said synthetic polynucleotide is of no more than about 100:1, 90:1, 80:1, 70:1, 60:1, 50:1, 40:1, or 30:1. In some embodiments, a (e.g., mass or weight) ratio of said ionizable cationic lipid to said synthetic polynucleotide

is of at least about 1:1, 2:1, 3:1, 4:1, or 5:1. In some embodiments, a (e.g., mass or weight) ratio of said ionizable cationic lipid to said synthetic polynucleotide is of about 1:1 to about 80:1, about 2:1 to about 80:1, about 3:1 to about 80:1, about 4:1 to about 80:1, or about 5:1 to about 80:1. In some embodiments, a (e.g., mass or weight) ratio of said ionizable cationic lipid to said synthetic polynucleotide is of about 1:1 to about 70:1, about 2:1 to about 70:1, about 3:1 to about 70:1, about 4:1 to about 70:1, or about 5:1 to about 70:1. In some embodiments, a (e.g., mass or weight) ratio of said ionizable cationic lipid to said synthetic polynucleotide is of about 1:1 to about 60:1, about 2:1 to about 60:1, about 3:1 to about 60:1, about 4:1 to about 60:1, or about 5:1 to about 60:1. In some embodiments, a (e.g., mass or weight) ratio of said ionizable cationic lipid to said synthetic polynucleotide is of about 1:1 to about 50:1, about 2:1 to about 50:1, about 3:1 to about 50:1, about 4:1 to about 50:1, or about 5:1 to about 50:1.

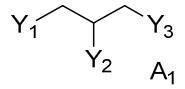
Selective Organ Targeting (SORT) Lipids

[00225] In some embodiments of the lipid composition of the present application, the lipid (e.g., nanoparticle) composition is preferentially delivered to a target organ. In some embodiments, the target organ is a lung, a lung tissue or a lung cell. As used herein, the term "preferentially delivered" is used to refer to a composition, upon being delivered, which is delivered to the target organ (e.g., lung), tissue, or cell in at least 25% (e.g., at least 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, or 75%) of the amount administered.

[00226] In some embodiments of the lipid composition, the lipid composition comprises one or more selective organ targeting (SORT) lipid which leads to the selective delivery of the composition to a particular organ. In some embodiments, the SORT lipid may have two or more alkyl or alkenyl chains of C_6-C_{24}.

[00227] In some embodiments of the lipid compositions, the SORT lipid comprises permanently positively charged moiety. The permanently positively charged moiety may be positively charged at a physiological pH such that the SORT lipid comprises a positive charge upon delivery of a polynucleotide to a cell. In some embodiments the positively charged moiety is quaternary amine or quaternary ammonium ion. In some embodiments, the SORT lipid comprises, or is otherwise complexed to or interacting with, a counterion.

[00228] In some embodiments, of the lipid compositions, the SORT lipid is a second ionizable cationic lipid. The SORT lipid may be an ionizable cationic lipid as described elsewhere in this disclosure.

[00229] In some embodiments of the lipid compositions, the SORT lipid is a permanently cationic lipid (i.e., comprising one or more hydrophobic components and a permanently cationic group). The permanently cationic lipid may contain a group which has a positive charge regardless of the pH. One permanently cationic group that may be used in the permanently cationic lipid is a quaternary ammonium group. The permanently cationic lipid may comprise a structural formula:  (S-I), wherein:

Y_1, Y_2, or Y_3 are each independently $X_1C(O)R_1$ or $X_2N^+R_3R_4R_5$;

provided at least one of Y_1, Y_2, and Y_3 is $X_2N^+R_3R_4R_5$;

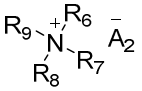
R_1 is C_1-C_{24} alkyl, C_1-C_{24} substituted alkyl, C_1-C_{24} alkenyl, C_1-C_{24} substituted alkenyl;

X_1 is O or NR_a, wherein R_a is hydrogen, C_1-C_4 alkyl, or C_1-C_4 substituted alkyl;

X_2 is C_1-C_6 alkanediyl or C_1-C_6 substituted alkanediyl;

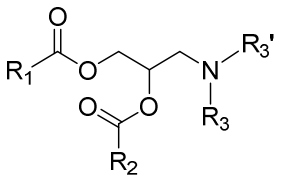
R_3, R_4, and R_5 are each independently C_1-C_{24} alkyl, C_1-C_{24} substituted alkyl, C_1-C_{24} alkenyl, C_1-C_{24} substituted alkenyl; and

A_1 is an anion with a charge equal to the number of $X_2N^+R_3R_4R_5$ groups in the compound.

[00230] In some embodiments of the SORT lipids, the permanently cationic SORT lipid has a structural formula:  (S-II), wherein:

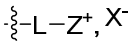
R_6-R_9 are each independently C_1-C_{24} alkyl, C_1-C_{24} substituted alkyl, C_1-C_{24} alkenyl, C_1-C_{24} substituted alkenyl; provided at least one of R_6-R_9 is a group of C_8-C_{24}; and

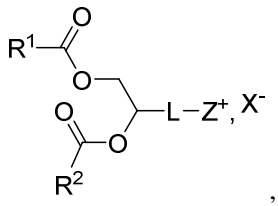
A_2 is a monovalent anion.

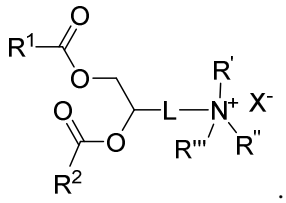
[00231] In some embodiments of the lipid compositions, the SORT lipid is a second ionizable cationic lipid (i.e., comprising one or more hydrophobic components and an ionizable cationic group). The ionizable positively charged moiety may be positively charged at a physiological pH. One ionizable cationic group that may be used in the second ionizable cationic lipid is a tertiary ammine group. In some embodiments of the lipid compositions, the SORT lipid has a structural formula:  (S-I'a), wherein:

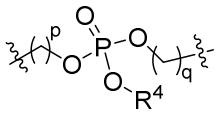
R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group; and

R_3 and R_3' are each independently $alkyl_{(C\leq 6)}$ or substituted $alkyl_{(C\leq 6)}$.

[00232] In some embodiments of the lipid compositions, the SORT lipid comprises a head group of a particular structure. In some embodiments, the SORT lipid comprises a headgroup having a structural formula: , wherein L is a linker; Z^+ is positively charged moiety and X^- is a counterion. In some embodiment, the linker is a biodegradable linker. The biodegradable linker may be degradable under physiological pH and temperature. The biodegradable linker may be degraded by proteins or enzymes from a subject. In some embodiments, the positively charged moiety is a quaternary ammonium ion or quaternary amine.

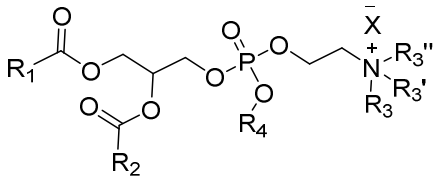
[00233] In some embodiments of the lipid compositions, the SORT lipid has a structural formula: , wherein R^1 and R^2 are each independently an optionally substituted C_6-C_{24} alkyl, or an optionally substituted C_6-C_{24} alkenyl.

[00234] In some embodiments of the lipid compositions, the SORT lipid has a structural formula: .

[00235] In some embodiments of the lipid compositions, the SORT lipid comprises a Linker (L). In some embodiments, L is , wherein:

p and q are each independently 1, 2, or 3; and

R^4 is an optionally substituted C_1-C_6 alkyl

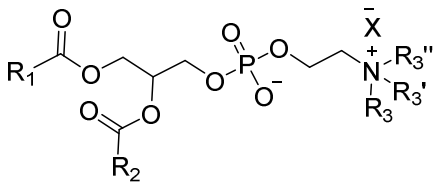
[00236] In some embodiments of the lipid compositions, the SORT lipid has a structural formula:  (IA), wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$;

R_4 is $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$; and

X^- is a monovalent anion.

[00237] In some embodiments of the lipid compositions, the SORT lipid is a phosphotidylcholine (e.g., 14:0 EPC). In some embodiments, the phophotidylcholine compound is further defined as:  (IA), wherein:

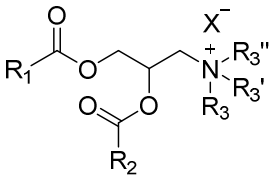
R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$; and

X^- is a monovalent anion.

[00238] In some embodiments of the lipid compositions, the SORT lipid is a phosphocholine lipid. In some embodiments, the SORT lipid is an ethylphosphocholine. The ethylphosphocholine may be, by way of example, without being limited to, 1,2-dimyristoleoyl-sn-glycero-3-ethylphosphocholine, 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine, 1,2-distearoyl-sn-glycero-3-ethylphosphocholine, 1,2-dipalmitoyl-sn-glycero-3-ethylphosphocholine, 1,2-dimyristoyl-sn-glycero-3-ethylphosphocholine, 1,2-dilauroyl-sn-glycero-3-ethylphosphocholine, 1-palmitoyl-2-oleoyl-sn-glycero-3-ethylphosphocholine.

[00239] In some embodiments of the lipid compositions, the SORT lipid has a structural formula:



(S-I'), wherein:

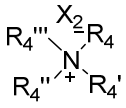
R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq 6)}$ or substituted $alkyl_{(C\leq 6)}$;

X^- is a monovalent anion.

[00240] By way of example, and without being limited thereto, a SORT lipid of the structural formula of the immediately preceding paragraph is 1,2-dioleoyl-3-trimethylammonium-propane (18:1 DOTAP) (e.g., chloride salt).

[00241] In some embodiments of the lipid compositions, the SORT lipid has a structural formula:



(S-II'), wherein:

R_4 and R_4' are each independently $alkyl_{(C6\text{-}C24)}$, $alkenyl_{(C6\text{-}C24)}$, or a substituted version of either group;

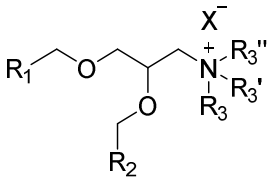
R_4'' is $alkyl_{(C\leq 24)}$, $alkenyl_{(C\leq 24)}$, or a substituted version of either group;

R_4''' is $alkyl_{(C1\text{-}C8)}$, $alkenyl_{(C2\text{-}C8)}$, or a substituted version of either group; and

X_2 is a monovalent anion.

[00242] By way of example, and without being limited thereto, a SORT lipid of the structural formula of the immediately preceding paragraph is dimethyldioctadecylammonium (DDAB) (e.g., bromide salt).

[00243] In some embodiments of the lipid compositions, the SORT lipid has a structural formula:



(S-III), wherein:

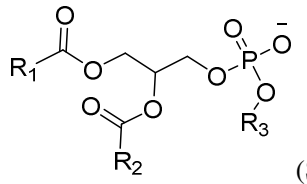
R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$; and

X^- is a monovalent anion.

[00244] By way of example, and without being limited thereto, a SORT lipid of the structural formula of the immediately preceding paragraph is N-[1-(2, 3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTMA).

[00245] In some embodiments of the lipid compositions, the SORT lipid is an anionic lipid. In some embodiments of the lipid compositions, the SORT lipid has a structural formula:



(S-IV), wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

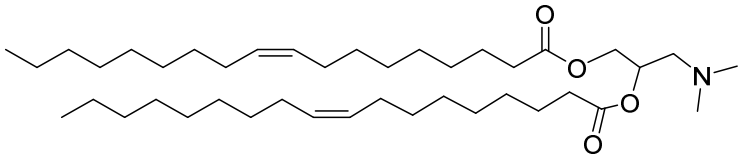
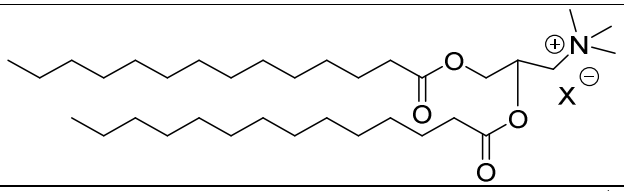
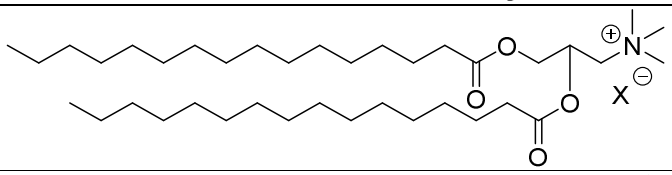
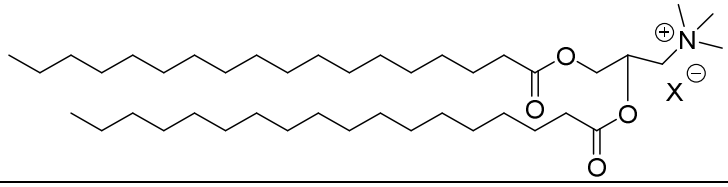
R_3 is hydrogen, $alkyl_{(C\leq6)}$, or substituted $alkyl_{(C\leq6)}$, or $-Y_1-R_4$, wherein:

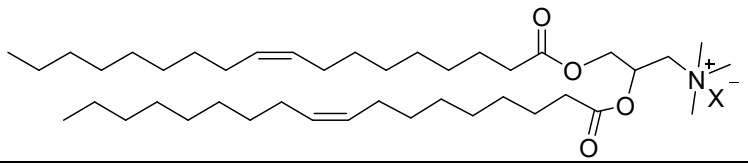
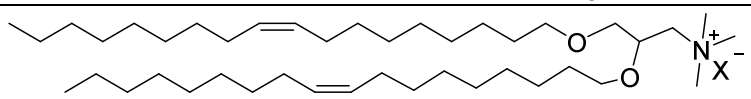
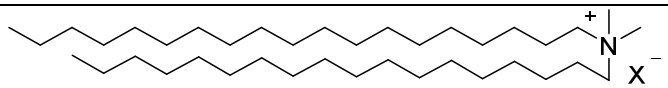
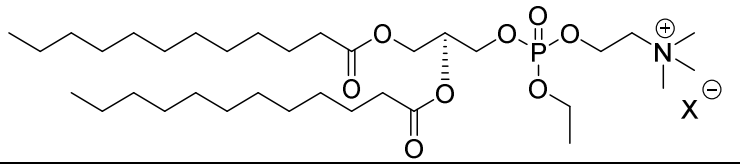
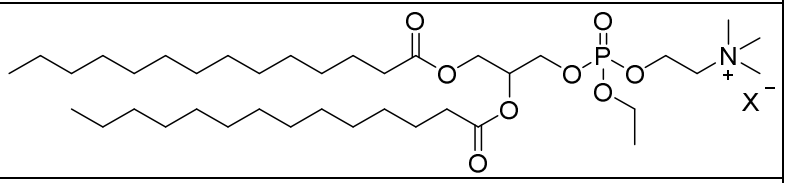
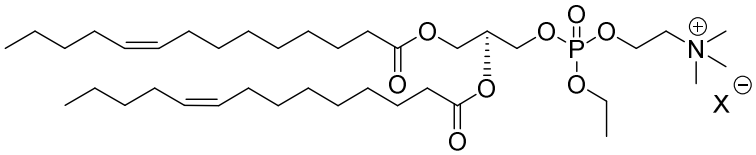
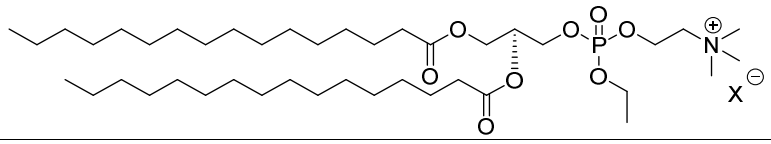
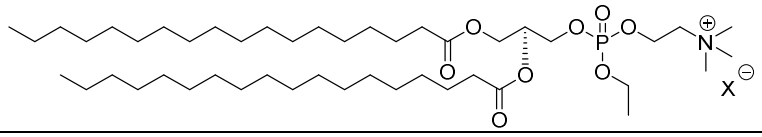
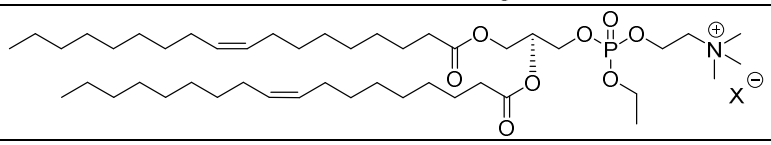
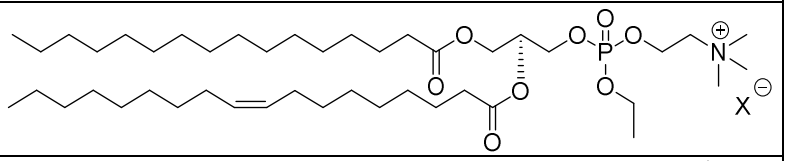
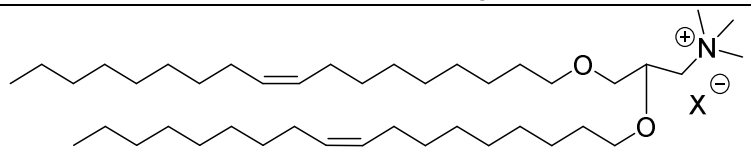
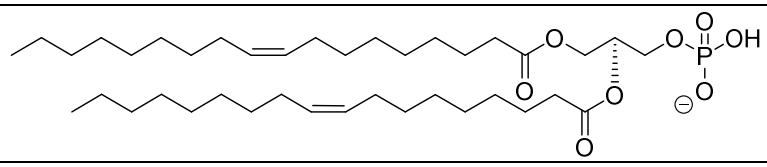
Y_1 is $alkanediyl_{(C\leq6)}$ or substituted $alkanediyl_{(C\leq6)}$; and

R_4 is $acyloxy_{(C\leq8\text{-}24)}$ or substituted $acyloxy_{(C\leq8\text{-}24)}$.

[00246] In some embodiments of the lipid compositions, the SORT lipid comprises one or more selected from the lipids set forth in **Table 8**.

Table 8. Example SORT lipids

Lipid Name	Structure
1,2-Dioleoyl-3-dimethylammonium-propane (18:1 DODAP)	
1,2-dimyristoyl-3-trimethylammonium-propane (14:0 TAP) (e.g., chloride salt)	
1,2-dipalmitoyl-3-trimethylammonium-propane (16:0 TAP) (e.g., chloride salt)	
1,2-stearoyl-3-trimethylammonium-propane (18:0 TAP) (e.g., chloride salt)	

1,2-Dioleoyl-3-trimethylammonium-propane (18:1 DOTAP) (e.g., chloride salt)	
1,2-Di-O-octadecenyl-3-trimethylammonium propane (DOTMA) (e.g., chloride salt)	
Dimethyldioctadecylammonium (DDAB) (e.g., bromide salt)	
1,2-dilauroyl-sn-glycero-3-ethylphosphocholine (12:0 EPC) (e.g., chloride salt)	
1,2-Dioleoyl-sn-glycero-3-ethylphosphocholine (14:0 EPC) (e.g., chloride salt)	
1,2-dimyristoleoyl-sn-glycero-3-ethylphosphocholine (14:1 EPC) (e.g., triflate salt)	
1,2-dipalmitoyl-sn-glycero-3-ethylphosphocholine (16:0 EPC) (e.g., chloride salt)	
1,2-distearoyl-sn-glycero-3-ethylphosphocholine (18:0 EPC) (e.g., chloride salt)	
1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (18:1 EPC) (e.g., chloride salt)	
1-palmitoyl-2-oleoyl-sn-glycero-3-ethylphosphocholine (16:0-18:1 EPC) (e.g., chloride salt)	
1,2-di-O-octadecenyl-3-trimethylammonium propane (18:1 DOTMA) (e.g., chloride salt)	
1,2-dioleoyl-sn-glycero-3-phosphate (18.1 PA)	

X^- is a counterion (e.g., Cl^-, Br^-, etc.)

[00247] In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 5% to about 65%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 5% to about 30%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 30% to about 55%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 20% to about 50%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 30% to about 60%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 25% to about 60%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 10% to about 20%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 20% to about 30%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 10% to about 30%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 10% to about 15%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage from about 15% to about 20%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage of at least (about) 25%, at least (about) 30%, at least (about) 35%, at least (about) 40%, at least (about) 45%, at least (about) 50%, at least (about) 55%, at least (about) 60%, or at least (about) 65%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said SORT lipid at a molar percentage of at most (about) 25%, at most (about) 30%, at most (about) 35%, at most (about) 40%, at least (about) 45%, at most (about) 50%, at most (about) 55%, at most (about) 60%, or at most (about) 65%.

Additional Lipids

[00248] In some embodiments of the lipid composition of the present application, the lipid composition further comprises an additional lipid including but not limited to a zwitterionic lipid (e.g., a phospholipid), a steroid or a steroid derivative, a polymer-conjugated lipid (e.g., polyethylene glycol (PEG)-conjugated lipid), or a combination thereof.

[00249] In some embodiments, a molar ratio of nitrogen in the lipid composition to phosphate in the synthetic polynucleotide (N/P ratio) is of no more than about 50:1, no more than about 40:1, no more than about 30:1, or no more than about 20:1. In some embodiments, a molar ratio of nitrogen in the lipid composition to phosphate in the synthetic polynucleotide (N/P ratio) is of at least about 1:1, at least about 2:1, at least about 3:1, at least about 4:1, or at least about 5:1. In some embodiments, a molar ratio of nitrogen in the lipid composition to phosphate in the synthetic polynucleotide (N/P ratio) is of about 1:1 to

about 50:1, at least about 2:1 to about 50:1, at least about 3:1 to about 50:1, at least about 4:1 to about 50:1, or at least about 5:1 to about 50:1. In some embodiments, a molar ratio of nitrogen in the lipid composition to phosphate in the synthetic polynucleotide (N/P ratio) is of about 1:1 to about 40:1, at least about 2:1 to about 40:1, at least about 3:1 to about 40:1, at least about 4:1 to about 40:1, or at least about 5:1 to about 40:1. In some embodiments, a molar ratio of nitrogen in the lipid composition to phosphate in the synthetic polynucleotide (N/P ratio) is of about 1:1 to about 30:1, at least about 2:1 to about 30:1, at least about 3:1 to about 30:1, at least about 4:1 to about 30:1, or at least about 5:1 to about 30:1.

Zwitterionic Lipids

[00250] In some embodiments of the lipid composition of the present application, the lipid composition further comprises a zwitterion lipid or a phospholipid. In some embodiments, the zwitterion lipid or phospholipid may contain one or two long chain (e.g., C_6-C_{24}) alkyl or alkenyl groups, a glycerol or a sphingosine, one or two phosphate groups, and, optionally, a small organic molecule. The small organic molecule may be an amino acid, a sugar, or an amino substituted alkoxy group, such as choline or ethanolamine. In some embodiments, the zwitterion lipid or phospholipid is a phosphatidylcholine. In some embodiments, the zwitterion lipid or phospholipid is distearoylphosphatidylcholine or dioleoylphosphatidylethanolamine. In some embodiments, other zwitterionic lipids are used, where zwitterionic lipid defines lipid and lipid-like molecules with both a positive charge and a negative charge.

[00251] In some embodiments of the lipid compositions, the zwitterion lipid or phospholipid is not an ethylphosphocholine.

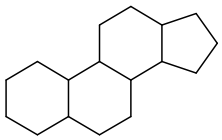
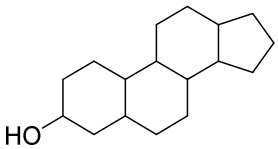
[00252] In some embodiments of the lipid composition of the present application, the compositions may further comprise a molar percentage of the zwitterion lipid or phospholipid to the total lipid composition from about 20 to about 23. In some embodiments, the molar percentage is from about 20, 20.5, 21, 21.5, 22, 22.5, to about 23 or any range derivable therein. In other embodiments, the molar percentage is from about 7.5 to about 60. In some embodiments, the molar percentage is from about 7.5, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, to about 20 or any range derivable therein.

[00253] In some embodiments of the lipid composition of the present application, said lipid composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage from about 5% to about 25%. In some embodiments of the lipid composition of the present application, said lipid composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage from about 10% to about 20%. In some embodiments of the lipid composition of the present application, said lipid composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage from about 15% to about 20%. In some embodiments of the lipid composition of the present application, said lipid composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage from about 8% to about 15%. In some embodiments of the lipid composition of the present application, said lipid composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage from about 10% to about 15%. In some embodiments of the lipid composition of the present application, said lipid

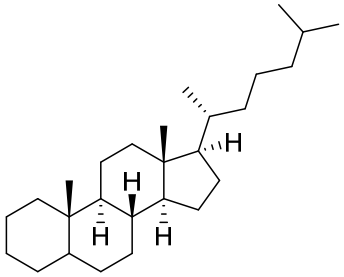
composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage from about 12% to about 18%. In some embodiments of the lipid composition of the present application, said lipid composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage of at least (about) 8%, at least (about) 10%, at least (about) 12%, at least (about) 15%, at least (about) 18%, at least (about) 20%, or at least (about) 25%. In some embodiments of the lipid composition of the present application, said lipid composition comprises the zwitterionic lipid (e.g., phospholipid or zwitterionic phospholipid) at a molar percentage of at most (about) 8%, at most (about) 10%, at most (about) 12%, at most (about) 15%, at most (about) 18%, at most (about) 20%, or at most (about) 25%.

[00254] In some embodiments, a (e.g., mass or weight) ratio of the zwitterionic lipid to the synthetic polynucleotide is of no more than about 50:1, 40:1, 30:1, 20:1, 10:1, or 7:1. In some embodiments, a (e.g., mass or weight) ratio of the zwitterionic lipid to the synthetic polynucleotide is of at least about 1:1, 2:1, 3:1, 4:1, or 5:1. In some embodiments, a (e.g., mass or weight) ratio of the zwitterionic lipid to the synthetic polynucleotide is of about 1:1 to about 10:1, about 1:1 to about 20:1, about 1:1 to about 30:1, about 1:1 to about 40:1, or about 1:1 to about 50:1.

Steroids or Steroid Derivatives

[00255] In some embodiments of the lipid composition of the present application, the lipid composition further comprises a steroid or steroid derivative. In some embodiments, the steroid or steroid derivative comprises any steroid or steroid derivative. As used herein, in some embodiments, the term "steroid" is a class of compounds with a four ring 17 carbon cyclic structure which can further comprises one or more substitutions including alkyl groups, alkoxy groups, hydroxy groups, oxo groups, acyl groups, or a double bond between two or more carbon atoms. In some embodiments, the ring structure of a steroid comprises three fused cyclohexyl rings and a fused cyclopentyl ring as shown in the formula:  . In some embodiments, a steroid derivative comprises the ring structure above with one or more non-alkyl substitutions. In some embodiments, the steroid or steroid derivative is a sterol wherein the formula is further defined as: 

. In some embodiments of the present application, the steroid or steroid derivative is a cholestane or cholestane derivative. In a cholestane, the ring structure is further

defined by the formula:  . As described above, a cholestane derivative includes one or more non-alkyl substitution of the above ring system. In some embodiments, the cholestane or cholestane derivative is a cholestene or cholestene derivative or a sterol or a sterol derivative. In other embodiments, the cholestane or cholestane derivative is both a cholestere and a sterol or a derivative thereof.

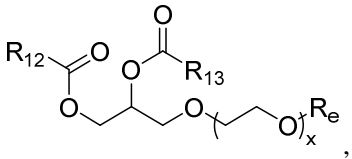
[00256] In some embodiments of the lipid composition, the compositions may further comprise a molar percentage of the steroid to the total lipid composition from about 40 to about 46. In some embodiments, the molar percentage is from about 40, 41, 42, 43, 44, 45, to about 46 or any range derivable therein. In other embodiments, the molar percentage of the steroid relative to the total lipid composition is from about 15 to about 40. In some embodiments, the molar percentage is 15, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, or 40, or any range derivable therein.

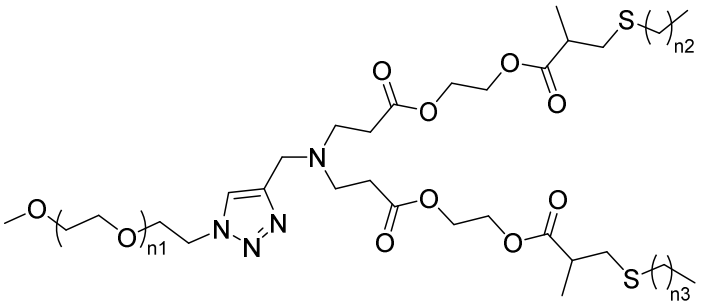
[00257] In some embodiments of the lipid composition of the present application, said lipid composition comprises said steroid or steroid derivative at a molar percentage from about 15% to about 46%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said steroid or steroid derivative at a molar percentage from about 20% to about 40%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said steroid or steroid derivative at a molar percentage from about 25% to about 35%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said steroid or steroid derivative at a molar percentage from about 30% to about 40%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said steroid or steroid derivative at a molar percentage from about 20% to about 30%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said steroid or steroid derivative at a molar percentage of at least (about) 15%, of at least (about) 20%, of at least (about) 25%, of at least (about) 30%, of at least (about) 35%, of at least (about) 40%, of at least (about) 45%, or of at least (about) 46%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said steroid or steroid derivative at a molar percentage of at most (about) 15%, of at most (about) 20%, of at most (about) 25%, of at most (about) 30%, of at most (about) 35%, of at most (about) 40%, of at most (about) 45%, or of at most (about) 46%.

Polymer-Conjugated Lipids

[00258] In some embodiments of the lipid composition of the present application, the lipid composition further comprises a polymer conjugated lipid. In some embodiments, the polymer conjugated lipidis a PEG lipid. In some embodiments, the PEG lipid is a diglyceride which also comprises a PEG chain attached to the glycerol group. In other embodiments, the PEG lipid is a compound which contains one or more C_6-C_{24}

long chain alkyl or alkenyl group or a C_6-C_{24} fatty acid group attached to a linker group with a PEG chain. Some non-limiting examples of a PEG lipid includes a PEG modified phosphatidylethanolamine and phosphatidic acid, a PEG ceramide conjugated, PEG modified dialkylamines and PEG modified 1,2-diacyloxypropan-3-amines, PEG modified diacylglycerols and dialkylglycerols. In some embodiments, PEG modified Phosphatidylethanolamine (PE). In some embodiments, PEG modified diastearoylphosphatidylethanolamine or PEG modified dimyristoyl-*sn*-glycerol. In some embodiments, the PEG modification is measured by the molecular weight of PEG component of the lipid. In some embodiments, the PEG modification has a molecular weight from about 100 to about 15,000. In some embodiments, the molecular weight is from about 200 to about 500, from about 400 to about 5,000, from about 500 to about 3,000, or from about 1,200 to about 3,000. The molecular weight of the PEG modification is from about 100, 200, 400, 500, 600, 800, 1,000, 1,250, 1,500, 1,750, 2,000, 2,250, 2,500, 2,750, 3,000, 3,500, 4,000, 4,500, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 12,500, to about 15,000. Some non-limiting examples of lipids that may be used in the present application are taught by U.S. Patent 5,820,873, WO 2010/141069, or U.S. Patent 8,450,298, which is incorporated herein by reference.

[00259] In some embodiments of the lipid composition of the present application, the PEG lipid has a structural formula: , wherein: R_{12} and R_{13} are each independently $alkyl_{(C\leq 24)}$, $alkenyl_{(C\leq 24)}$, or a substituted version of either of these groups; R_e is hydrogen, $alkyl_{(C\leq 8)}$, or substituted $alkyl_{(C\leq 8)}$; and x is 1-250. In some embodiments, R_e is $alkyl_{(C\leq 8)}$ such as methyl. R_{12} and R_{13} are each independently $alkyl_{(C\leq 4\text{-}20)}$. In some embodiments, x is 5-250. In one embodiment, x is 5-125 or x is 100-250. In some embodiments, the PEG lipid is 1,2-dimyristoyl-*sn*-glycerol, methoxypolyethylene glycol.

[00260] In some embodiments of the lipid composition of the present application, the PEG lipid has a structural formula: , wherein: n_1 is an integer between 1 and 100 and n_2 and n_3 are each independently selected from an integer between 1 and 29. In some embodiments, n_1 is 5, 10, 15, 20, 25, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100, or any range derivable therein. In some embodiments, n_1 is from about 30 to about 50. In some embodiments, n_2 is from 5 to 23. In some embodiments, n_2 is 11 to about 17. In some embodiments, n_3 is from 5 to 23. In some embodiments, n_3 is 11 to about 17.

[00261] In some embodiments of the lipid composition of the present application, the compositions may further comprise a molar percentage of the PEG lipid to the total lipid composition from about 4.0 to about 4.6. In some embodiments, the molar percentage is from about 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, to about 4.6 or any range derivable therein. In other embodiments, the molar percentage is from about 1.5 to about 4.0. In some embodiments, the molar percentage is from about 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 3.25, 3.5, 3.75, to about 4.0 or any range derivable therein.

[00262] In some embodiments of the lipid composition of the present application, said lipid composition comprises said polymer-conjugated lipid at a molar percentage from about 0.5% to about 12%, or from about 0.5% to about 10%. In some embodiments of the lipid composition of the present application, the lipid composition comprises the polymer-conjugated lipid at a molar percentage from about 1% to about 10%. In some embodiments of the lipid composition of the present application, the lipid composition comprises the polymer-conjugated lipid at a molar percentage from about 2% to about 10%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said polymer-conjugated lipid at a molar percentage from about 1% to about 8%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said polymer-conjugated lipid at a molar percentage from about 2% to about 7%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said polymer-conjugated lipid at a molar percentage from about 3% to about 5%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said polymer-conjugated lipid at a molar percentage from about 5% to about 10%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said polymer-conjugated lipid at a molar percentage of at least (about) 0.5%, at least (about) 1%, at least (about) 1.5%, at least (about) 2%, at least (about) 2.5%, at least (about) 3%, at least (about) 3.5%, at least (about) 4%, at least (about) 4.5%, at least (about) 5%, at least (about) 5.5%, at least (about) 6%, at least (about) 6.5%, at least (about) 7%, at least (about) 7.5%, at least (about) 8%, at least (about) 8.5%, at least (about) 9%, at least (about) 9.5%, or at least (about) 10%. In some embodiments of the lipid composition of the present application, said lipid composition comprises said polymer-conjugated lipid at a molar percentage of at most (about) 0.5%, at most (about) 1%, at most (about) 1.5%, at most (about) 2%, at most (about) 2.5%, at most (about) 3%, at most (about) 3.5%, at most (about) 4%, at most (about) 4.5%, at most (about) 5%, at most (about) 5.5%, at most (about) 6%, at most (about) 6.5%, at most (about) 7%, at most (about) 7.5%, at most (about) 8%, at most (about) 8.5%, at most (about) 9%, at most (about) 9.5%, or at most (about) 10%.

[00263] Some embodiments of the (e.g., pharmaceutical) composition disclosed herein comprise a particular molar ratio of the components or atoms. In some embodiments, the (e.g., pharmaceutical) composition comprises a particular molar ratio of nitrogen in the lipid composition to the phosphate in the polynucleotide (N/P ratio). In some embodiments, the molar ratio of nitrogen in the lipid composition to phosphate in the polynucleotide (N/P ratio) is no more than about 30:1. In some embodiments, the N/P ratio is from about 5:1 to about 30:1. In some embodiments, the N/P ratio is no more than 1:1, 2:1, 3:1,

4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, 50:1, or less. In some embodiments, the N/P ratio is at least 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, 50:1, or more. In some embodiments, the N/P ratio is of any one of the following values or within a range of any two of the following values: 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, and 50:1.

[00264] In some embodiments, composition comprises a particular (e.g., mass or weight) ratio of the polynucleotide to total lipids of the lipid composition. In some embodiments, the (e.g., mass or weight) ratio of the polynucleotide to total lipids of the lipid composition is no more than about 1:1, 1:10, 1:50, or 1:100. In some embodiments, the (e.g., mass or weight) ratio of the polynucleotide to total lipids of the lipid composition is no more than about 1:1, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:15, 1:20, 1:25, 1:30, 1:35, 1:40, 1:45, 1:50, 1:75, or 1:100 or less. In some embodiments, the (e.g., mass or weight) ratio of the polynucleotide to total lipids of the lipid composition is at least about 1:1, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:15, 1:20, 1:25, 1:30, 1:35, 1:40, 1:45, 1:50, 1:75, or 1:100 or more. In some embodiments, the (e.g., mass or weight) ratio of the polynucleotide to total lipids of the lipid composition is of any one of the following values or within a range of any two of the following values: 1:1, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:15, 1:20, 1:25, 1:30, 1:35, 1:40, 1:45, 1:50, 1:75, and 1:100.

[00265] In some embodiments of the composition, the composition can be formulated as any suitable dosage form known in the art. In some embodiments, the composition is formulated in a nanoparticle or a nanocapsule. In some embodiments, the composition is formulated for administration by any suitable route known in the art including, for example, oral, rectal, vaginal, transmucosal, pulmonary including intratracheal or inhaled, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections.

[00266] In some embodiments, the composition of the present application is formulated for administration by a local rather than systemic manner, for example, via injection of the pharmaceutical composition directly into a targeted tissue, such as in a sustained release formulation. Local delivery can be affected in various ways, depending on the tissue to be targeted. The composition may be formulated for aerosol administration. The aerosol administration may be delivered to the respiratory epithelium. In some embodiments, the aerosol composition has a droplet size from 0.5 micron (μm) to 10 μm. In some embodiments, the aerosol composition has a median droplet size from 0.5 μm to 10 μm. In some embodiments, the aerosol composition has an average droplet size from 0.5 μm to 10 μm. The droplet size may be measured using cascade impactor analysis or laser diffraction, or other suitable techniques for measuring aerosol droplets.

[00267] In some embodiments, the composition of the present application can be injected into the site of injury, disease manifestation, or pain, for example. In some embodiments, the composition of the present application can be provided in lozenges for oral, tracheal, or esophageal application. In some embodiments,

the composition of the present application can be supplied in liquid, tablet or capsule form for administration to the stomach or intestines. In some embodiments, the composition of the present application can be supplied in suppository form for rectal or vaginal application. In some embodiments, the composition of the present application can even be delivered to the eye by use of creams, drops, or even injection.

METHODS

Methods for enhancing CFTR expression or activity in cell(s)

[00268] In some embodiments, provided herein is a method for enhancing an expression or activity of cystic fibrosis transmembrane conductance regulator (CFTR) protein in a cell, comprising: contacting said cell with a synthetic polynucleotide as described herein assembled with a lipid composition, wherein said synthetic polynucleotide encodes a CFTR protein; and wherein said lipid composition comprises: (1) an ionizable cationic lipid; (2) a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid and said phospholipid, thereby resulting in an expression or activity of said CFTR protein in said cell. The lipid composition may further comprise a zwitterionic lipid or phospholipid.

[00269] In some embodiments, provided herein in a method for enhancing an expression or activity of cystic fibrosis transmembrane conductance regulator (CFTR) protein in a cell, the method comprising: contacting said cell with a composition comprising a synthetic polynucleotide assembled with a lipid composition, wherein said synthetic polynucleotide encodes a CFTR protein; and wherein said lipid composition comprises: an ionizable cationic lipid; and a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid, thereby yielding a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after contacting, optionally wherein said therapeutically effective activity of said functional variant of CFTR protein is determined by measuring a change in a transepithelial ion transport characteristic of a plurality of cells comprising said cell as compared to that of a reference plurality of cells in absence of said contacting. The lipid composition may further comprise a zwitterionic lipid or phospholipid.

[00270] In some embodiments of the method, the composition of the present application is formulated for administration by a local rather than systemic manner, for example, via injection of the pharmaceutical composition directly into a targeted tissue, such as in a sustained release formulation. Local delivery can be affected in various ways, depending on the tissue to be targeted. In some embodiments of the method, aerosols containing the composition of the present application can be inhaled (for nasal, tracheal, or bronchial delivery). The composition may be formulated for aerosol administration.

[00271] In some embodiments, the contacting is repeated. The contacting may be repeated 1, 2, 3, or more times. In some embodiments, the contacting is at least once a week. In some embodiments, the contacting is at least twice a week. In some embodiments, the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after each contacting. In some embodiments, a second contacting is performed, optionally at least about 1, 2, or 3 day(s) after the first contacting. In some embodiments, the method further comprises a third contacting wherein said third

contacting is performed optionally at least about 1, 2, or 3 day(s) after the second contacting. In embodiments, the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after a second contacting. In some embodiments, the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after a third contacting. The composition in each contacting may be the same or identical. The therapeutically effective amount or activity of a functional variant of CFTR protein may increase after repeated contacting.

[00272] The contacting(s) may be performed *in vivo.* The contacting(s) may be performed *in vitro.* The contacting(s) may be performed *ex vivo*.

[00273] In some embodiments, the methods achieve a therapeutically effective activity of said functional variant of CFTR protein. In some embodiments, therapeutically effective activity may be measured by a transepithelial assay. The transepithelial assay may measure a voltage or a current which may correspond to the function of a functional protein. In some embodiments, the therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least 5 micro-Amperes(μA). In some embodiments, the therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from at least 5 micro-Amperes (μA) to about 30 μA. In some embodiments, therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$). In some embodiments, said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$) to about 20 $\mu A \cdot cm^{-2} \cdot min^{-1}$. The transepithelial current may be determined via an *in vitro* assay, such as those described elsewhere herein.

[00274] In some embodiments, the methods achieve a therapeutically effective activity of said functional variant of CFTR protein can be the measurement of the forced expiratory volume (FEV) of a subject. The FEV measures how much air a subject can exhale during a forced breath. The amount of air exhaled may be measured during the first (FEV1), second (FEV2), or third (FEV3) second(s) of the forced breath. In some embodiments, the method results in the subject having a FEV1, FEV2, or FEV3 of about 40% to about 90%, about 50% to about 90%, about 60% to about 90%, about 70% to about 90%, about 80% to about 90%, about 40% to about 80%, about 40% to about 70%, about 40% to about 60%, or about 40% to about 50%. In some embodiments, the method results in the subject having a FEV1, FEV2, or FEV3 of about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%. In some embodiments, the method results in the subject having a FEV1, FEV2, or FEV3 of about 40% to about 90%.

[00275] In some embodiments of the methods, the method increases an amount of a functional variant of CFTR protein in the cell relative to a corresponding control. In some embodiments, the method increases an amount of WT CFTR protein in said cell relative to a corresponding control. In some embodiments, said control comprises a corresponding cell absent said contacting. In some embodiments, the method increases

an amount of said functional variant of CFTR protein by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.1-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 2.6-fold, at least about 2.7-fold, at least about 2.8-fold, at least about 2.9-fold, at least about 3.0-fold, at least about 3.1-fold, at least about 3.2-fold, at least about 3.3-fold, at least about 3.4-fold, at least about 3.5-fold, at least about 3.6-fold, at least about 3.7-fold, at least about 3.8-fold, at least about 3.9-fold, at least about 4.0-fold, at least about 4.1-fold, at least about 4.2-fold, at least about 4.3-fold, at least about 4.4-fold, at least about 4.5-fold, at least about 4.6-fold, at least about 4.7-fold, at least about 4.8-fold, at least about 4.9-fold, or at least about 5.0-fold, in said cell relative to a corresponding control. In some embodiments, the method increases an amount of WT CFTR protein by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.1-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 2.6-fold, at least about 2.7-fold, at least about 2.8-fold, at least about 2.9-fold, at least about 3.0-fold, at least about 3.1-fold, at least about 3.2-fold, at least about 3.3-fold, at least about 3.4-fold, at least about 3.5-fold, at least about 3.6-fold, at least about 3.7-fold, at least about 3.8-fold, at least about 3.9-fold, at least about 4.0-fold, at least about 4.1-fold, at least about 4.2-fold, at least about 4.3-fold, at least about 4.4-fold, at least about 4.5-fold, at least about 4.6-fold, at least about 4.7-fold, at least about 4.8-fold, at least about 4.9-fold, or at least about 5.0-fold, in said cell relative to a cell absent said contacting.

[00276] In some embodiments, the method results in a therapeutically effective amount of said functional variant of CFTR protein in said cell. In some embodiments, the method results in a therapeutically effective amount of WT CFTR protein in said cell.

[00277] In some embodiment, the method enhances ion transport in said cell relative to a corresponding control. In some embodiment, the method enhances chloride transport in said cell relative to a corresponding control. In some embodiments, said control comprises a corresponding cell absent said contacting. In some embodiment, the method enhances ion transport by at least about 1.1-fold, at least about 1.2-fold, at least about 1.3-fold, at least about 1.4-fold, at least about 1.5-fold, at least about 1.6-fold, at least about 1.7-fold, at least about 1.8-fold, at least about 1.9-fold, at least about 2.0-fold, at least about 2.1-fold, at least about 2.2-fold, at least about 2.3-fold, at least about 2.4-fold, at least about 2.5-fold, at least about 2.6-fold, at least about 2.7-fold, at least about 2.8-fold, at least about 2.9-fold, at least about 3.0-fold, at least about 3.1-fold, at least about 3.2-fold, at least about 3.3-fold, at least about 3.4-fold, at least about 3.5-fold, at least about 3.6-fold, at least about 3.7-fold, at least about 3.8-fold, at least about 3.9-fold, at least about 4.0-fold, at least about 4.1-fold, at least about 4.2-fold, at least about 4.3-fold, at least about 4.4-fold, at least about 4.5-fold, at least about 4.6-fold, at least about 4.7-fold, at least about 4.8-fold, at least about 4.9-fold, or at least about 5.0-fold, in said cell relative to a corresponding control.

Methods for treating cystic fibrosis

[00278] In some embodiments, provided herein is a method for treating a subject having or suspected of having a cystic fibrosis transmembrane conductance regulator (CFTR)-associated condition. The method may comprise administering to the subject a composition as described herein. In some embodiments, the CFTR-associated condition is cystic fibrosis, hereditary emphysema, or chronic obstructive pulmonary disease (COPD), or a combination thereof. In some embodiments, said subject is a mammal, such as a human, monkey, cow, sheep, goat, dog, cat, mouse, rat, guinea pig, or transgenic species thereof. The subject may be a mammal. The subject may be a human. In some embodiments, the administering comprises pulmonary administration. In some embodiments, the administering comprises inhalation by nebulization. In some embodiments, the administering comprises apical administration. In some embodiments, said subject is human. In some embodiments, said subject exhibits or is determined to exhibit a mutation in CFTR gene. In some embodiments, the mutation is R553X, G542X or F508del, or a combination thereof. In some embodiments, the mutation is R1162X. In some embodiments, the mutation is R553X, G542X, F508del, or R1162X, or a combination thereof.

[00279] The methods of the disclosure may be able to treat a subject with cystic fibrosis based on properties of the formulation or compositions. Specifically, the compositions described elsewhere herein may be able to penetrate the mucus associated with cystic fibrosis and thereby deliver the polynucleotides to the cells.

Cells

[00280] In some embodiments of any one method described herein, said cell is a lung cell. In some embodiments, said lung cell is a lung airway cell. Example lung airway cells that can be targeted by the delivery of the present application includes but is not limited to basal cell, secretory cell such as goblet cell and club cell, ciliated cell, ionocyte and any combination thereof. In some embodiments of the method, said cell is an airway epithelial cell. In some embodiments, said cell is a bronchial epithelial cell. In some embodiments, said cell is an airway epithelial cell. In some embodiments, said cell is a basal cell characterized by expression of p63 marker. In some embodiments, said cell is an ionocyte characterized by expression of FOXI1 marker. In some embodiments, said cell is undifferentiated. In some embodiments, said cell is differentiated.

Mutation(s)

[00281] In some embodiments of any one method described herein, said cell exhibits or is determined to exhibit a mutation in CFTR gene or transcript. In some embodiments, said cell exhibits or is determined to exhibit a mutation in one or more of exons 11-27 of CFTR gene. said cell exhibits or is determined to exhibit a nonsense or frameshift mutation in one or more of exons 11-27 of CFTR gene. In some embodiments, the mutation is located at a position in the CFTR gene at which a change can give rise to a mutant protein having a mutation at F508, e.g., F508del. In some embodiments, the mutation is located at a position in the CFTR gene at which a change can give rise to a mutant protein having a mutation at G542, e.g., G542X, in the CFTR protein, e.g., which corresponds to c.1624G>T in the CFTR gene. In some embodiments, the mutation is located at a position in the CFTR gene at which a change can give rise to a

mutant protein having a mutation at R553, e.g., R553X, in the CFTR protein. In some embodiments, the mutation is located at a position in the CFTR gene at which a change can give rise to a mutant protein having a mutation at R1162, e.g., R1162X, in the CFTR protein.

[00282] In some embodiments of any one method described herein, said mutation is associated with cystic fibrosis, hereditary emphysema, or chronic obstructive pulmonary disease (COPD).

Methods for lung cell delivery

[00283] In some embodiments, provided is a method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), the method comprising administering to a subject a composition as described herein, thereby yielding a therapeutically effective amount or activity of a synthetic polynucleotide in a lung secretory cell or lung basal cell of the subject. Optionally, the therapeutically effective activity of the synthetic polynucleotide may be determined by measuring a change in a transepithelial ion transport characteristic (e.g., a transepithelial current or voltage) of a lung comprising the lung secretory cell or lung basal cell as compared to that of a reference lung, e.g., in absence of the contacting. The composition may comprise a synthetic polynucleotide (as described herein) assembled with a lipid composition (as described herein). The synthetic polynucleotide may encode a cystic fibrosis transmembrane conductance regulator (CFTR) protein. The lipid composition may comprise an ionizable cationic lipid (as described herein); and a selective organ targeting (SORT) lipid (as described herein) separate from the ionizable cationic lipid. The lung basal cell may be a lung basal stem cell.

[00284] In some embodiments, provided is a method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), the method comprising administering to a subject a composition as described herein, thereby yielding a greater therapeutic amount or activity of a synthetic polynucleotide in a lung secretory cell or lung basal cell of the subject as compared to that in a lung non-secretory cell or lung non-basal cell of the subject. The composition may comprise a synthetic polynucleotide (as described herein) assembled with a lipid composition (as described herein). The synthetic polynucleotide may encode a cystic fibrosis transmembrane conductance regulator (CFTR) protein. The lipid composition may comprise an ionizable cationic lipid (as described herein); and a selective organ targeting (SORT) lipid (as described herein) separate from the ionizable cationic lipid. In some embodiments, the method yields an amount or activity of the synthetic polynucleotide in the lung secretory and/or basal cell that is at least 1.1-, 1.5-, 2-, 2.5-, 3-, 3.5-, 4-, 4.5- or 5-fold greater than that in the lung non-secretory and/or non-basal cell. In some embodiments, the method yields an amount or activity of the synthetic polynucleotide in the lung secretory cell that is at least 1.1-, 1.5-, 2-, 2.5-, 3-, 3.5-, 4-, 4.5- or 5-fold greater than that in the lung non-secretory. In some embodiments, the method yields an amount or activity of the synthetic polynucleotide in the lung basal cell that is at least 1.1-, 1.5-, 2-, 2.5-, 3-, 3.5-, 4-, 4.5- or 5-fold greater than that in the lung non-basal cell. The lung basal cell may be a lung basal stem cell. The lung non-secretory cell may be a lung ciliated cell. The lung non-basal cell may be a lung ciliated cell.

[00285] In some embodiments, provided is a method for targeted pulmonary delivery, such as lung secretory cell delivery, the method comprising administering to a subject a composition as described herein, thereby yielding a therapeutically effective amount or activity of a synthetic polynucleotide in a lung secretory cell of the subject. Optionally, the therapeutically effective activity of the synthetic polynucleotide may be determined by measuring a change in a transepithelial ion transport characteristic (e.g., a transepithelial current or voltage) of a lung comprising the lung secretory cell as compared to that of a reference lung, e.g., in absence of the contacting. The composition may comprise a synthetic polynucleotide (as described herein) assembled with a lipid composition (as described herein). The synthetic polynucleotide may encode a cystic fibrosis transmembrane conductance regulator (CFTR) protein. The lipid composition may comprise an ionizable cationic lipid (as described herein); and a selective organ targeting (SORT) lipid separate from the ionizable cationic lipid (as described herein).

[00286] In some embodiments, provided is a method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), the method comprising administering to a subject a composition as described herein, thereby yielding a greater therapeutic amount or activity of a synthetic polynucleotide in a lung secretory cell or lung basal cell of the subject as compared to that in a lung non-secretory cell or lung non-basal cell of the subject. The composition may comprise a synthetic polynucleotide (as described herein) assembled with a lipid composition (as described herein). The synthetic polynucleotide may encode a cystic fibrosis transmembrane conductance regulator (CFTR) protein. The lipid composition may comprise an ionizable cationic lipid (as described herein); and a selective organ targeting (SORT) lipid (as described herein) separate from the ionizable cationic lipid. In some embodiments, the method yields an amount activity of the synthetic polynucleotide in the lung secretory cell that is at least 1.1-, 1.5-, 2-, 2.5-, 3-, 3.5-, 4-, 4.5- or 5-fold greater than that in the lung non-secretory cell. The lung non-secretory cell may be a lung ciliated cell. In some embodiments, the method yields an amount activity of the synthetic polynucleotide in the lung basal cell that is at least 1.1-, 1.5-, 2-, 2.5-, 3-, 3.5-, 4-, 4.5- or 5-fold greater than that in the lung non-basal cell. The lung non-basal cell may be a lung ciliated cell. The lung basal cell may be a lung basal stem cell.

[00287] In various embodiments of the method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), at least about 50%, 55%, or 60% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected or observed in lung secretory cells, lung basal cells, or a combination thereof, e.g., as determined by measuring an amount or activity of the corresponding polypeptide encoded by the synthetic polynucleotide. In various embodiments of the method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery), no more than about 50%, 45%, or 40% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected or observed in lung non-secretory cells, lung non-basal cells, or a combination thereof, e.g., as determined by measuring an amount or activity of the corresponding polypeptide encoded by the synthetic polynucleotide. In various embodiments of the method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively,

lung secretory and/or basal cell delivery), no more than about 50%, 45%, or 40% of (e.g., pulmonary) expression of said synthetic polynucleotide is in lung ciliated cells, e.g., as determined by measuring an amount or activity of the corresponding polypeptide encoded by the synthetic polynucleotide. In various embodiments, the method for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery (alternatively, lung secretory and/or basal cell delivery) yields an amount or activity of said synthetic polynucleotide in lung secretory cell(s) or lung basal cell(s) that is at least 1.1-, 1.5-, or 2-fold greater than that in reference cell(s), which reference cell(s) are neither lung secretory cell(s) nor lung basal cell(s). The reference cell(s) may be lung ciliated cell(s). In various embodiments, the lung non-secretory cell or lung non-basal cell is a lung ciliated cell. The lung basal cell may be a lung basal stem cell. In some embodiments, the method for targeted pulmonary delivery, such as lung secretory cell delivery yields an amount or activity of said synthetic polynucleotide in lung secretory cell(s) that is at least 1.1-, 1.5-, or 2-fold greater than that in lung non-secretory cell(s). In some embodiments, the method for targeted pulmonary delivery, such as lung basal cell delivery yields an amount or activity of said synthetic polynucleotide in lung basal cell(s) that is at least 1.1-, 1.5-, or 2-fold greater than that in lung non-basal cell(s).

[00288] In some embodiments, provided is a method for targeted pulmonary delivery, such as lung secretory cell or basal cell delivery, comprising administering to a subject a composition as described herein, thereby yielding a therapeutic amount or activity of the synthetic polynucleotide in at least (about) 5% of lung secretory cells or lung basal cells of the subject. The composition may comprise a synthetic polynucleotide (as described herein) assembled with a lipid composition (as described herein). The synthetic polynucleotide may encode a cystic fibrosis transmembrane conductance regulator (CFTR) protein. The lipid composition may comprise an ionizable cationic lipid (as described herein); and a selective organ targeting (SORT) lipid (as described herein) separate from the ionizable cationic lipid.

[00289] In some embodiments of various methods for targeted pulmonary delivery, such as lung secretory cell or lung basal cell delivery, the lung secretory cell or lung basal cell exhibits or is determined to exhibit a mutation in *CFTR* gene. In some embodiments, the mutation is selected from the group consisting of G542X or F508del. In some embodiments, the mutation is R553X, G542X or F508del, or a combination thereof. In some embodiments, the mutation is R1162X. In some embodiments, the mutation is R553X, G542X, F508del, or R1162X, or a combination thereof.

[00290] A therapeutically effective activity of a functional variant of CFTR protein may be determined by measuring a change in a transepithelial ion transport characteristic (e.g., transepithelial current or voltage) of a lung of the subject as compared to that of a reference lung, e.g., prior to the administration.

[00291] In some embodiments, provided is a method for targeted pulmonary delivery, such as basal cell delivery of a synthetic polynucleotide that encodes a CFTR protein, comprising contacting a cell composition comprising a plurality of basal cells with a composition that comprises the synthetic polynucleotide assembled with a lipid composition, which lipid composition comprises: (1) an ionizable cationic lipid; and (2) a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby delivering said synthetic polynucleotide to at least 15% of said plurality of basal cells. In some embodiments, the lipid composition further comprises a zwitterionic lipid or phospholipid. In some embodiments, said cell composition comprises lung basal cells. In some embodiments, said cell composition comprises basal cell, secretory cell such as goblet cell and club cell, ciliated cell, ionocyte and any combination thereof. In some embodiments, said cell composition comprises a first cell of a first CFTR genotype and a second cell of a second CFTR genotype.

[00292] In some embodiments, provided is method for targeted pulmonary delivery, such as basal cell-targeted delivery of a synthetic polynucleotide that encodes a CFTR protein, comprising contacting a plurality of cells of a plurality of cell types with a composition that comprises said synthetic polynucleotide as described herein assembled with a lipid composition, which plurality of cells comprise a basal cell and a non-basal cell, wherein said lipid composition comprises: (1) an ionizable cationic lipid; and (2) a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid, thereby delivering said synthetic polynucleotide to said basal cell in a greater amount than that delivered to said non-basal cell. In some embodiments, the lipid composition further comprises a zwitterionic lipid or phospholipid. In some embodiments, said basal cell is a lung basal cell. In some embodiments, said non-basal cell is a lung non-basal cell. In some embodiments, said non-basal cell is a , secretory cell such as goblet cell and club cell, ciliated cell, ionocyte and any combination thereof. In some embodiments, the non-basal cell is a ciliated cell. In some embodiments, said plurality of cells comprise a first cell of a first CFTR genotype and a second cell of a second CFTR genotype.

[00293] In some embodiments, the composition can be formulated as any suitable dosage from known in the art. In some embodiments, the composition is formulated in a nanoparticle or a nanocapsule. In some embodiments, the composition is formulated for administration by any suitable route known in the art including, for example, oral, rectal, vaginal, transmucosal, pulmonary including intratracheal or inhaled, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. In some embodiments, said composition is formulated for apical delivery. In some embodiments, said composition is formulated for nebulization. In some embodiments, said composition is an aerosol. In some embodiments, said composition is formulated for intravenous administration.

[00294] In some embodiments, the composition can be formulated as any suitable dosage from known in the art. In some embodiments, the composition is formulated in a nanoparticle or a nanocapsule. In some embodiments, the composition is formulated for administration by any suitable route known in the art including, for example, oral, rectal, vaginal, transmucosal, pulmonary including intratracheal or inhaled, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. In some embodiments, said composition is formulated for apical delivery. In some embodiments, said composition is formulated for nebulization. In some embodiments, said composition is formulated for nebulization with a rate usage rate of about 0.1 ml/min to about 1.0 ml/min. In some

embodiments, said composition is formulated for nebulization with a rate usage rate of about 0.2 ml/min to about 0.7 ml/min. In some embodiments, said composition is formulated for nebulization with a rate usage rate of about 0.1 ml/min to about 0.5 ml/min. In some embodiments, said composition is formulated for nebulization with a rate usage rate of about 0.5 ml/min to about 1.0 ml/min. In some embodiments, said composition is an aerosol. In some embodiments, the mass median aerodynamic diameter (MMAD) of the aerosols ranges from about 1.0 μm to about 10.0 μm, from about 1.0 μm to about 5.0 μm, from about 2.0 μm to about 5.0 μm, or from about 3.0 μm to about 6.0 μm.

LIST OF EMBODIMENTS

[00295] The following list of embodiments of the invention are to be considered as disclosing various features of the invention, which features can be considered to be specific to the particular embodiment under which they are discussed, or which are combinable with the various other features as listed in other embodiments. Thus, simply because a feature is discussed under one particular embodiment does not necessarily limit the use of that feature to that embodiment.

[00296] Embodiment 1. A synthetic polynucleotide encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said synthetic polynucleotide comprises one or more nucleoside analogue(s).

[00297] Embodiment 2. The synthetic polynucleotide of Embodiment 1, wherein said synthetic polynucleotide comprises 1-methylpseudouridine.

[00298] Embodiment 3. A synthetic polynucleotide encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70%,75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 bases of a sequence selected from SEQ ID NOs: 1-4 and 23.

[00299] Embodiment 4. The synthetic polynucleotide of any one of Embodiment 1-3, wherein said nucleic acid sequence comprises fewer than about 115, 110, 105, 100, 95, or 90 UU or TT dinucleotide

[00300] Embodiment 5. The synthetic polynucleotide of any one of embodiments 1-4, wherein said nucleic acid sequence comprises at least two synonymous codons encoding arginine.

[00301] Embodiment 6. The synthetic polynucleotide of any one of embodiments 1-4, wherein said nucleic acid sequence comprises at least three synonymous codons encoding arginine.

[00302] Embodiment 7. The synthetic polynucleotide of any one of embodiments 1-6, wherein no more than about 70%, 65%, 60%, 55%, or 50% of all arginine encoding codons of said nucleic acid sequence is AGA codon.

[00303] Embodiment 8. The synthetic polynucleotide of any one of embodiments 1-7, wherein said nucleic acid sequence encodes a polypeptide that comprises an amino acid sequence having at least 70%, 75%, 80%,81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 contiguous amino acid residues to SEQ ID NO: 5.

[00304] Embodiment 9. The synthetic polynucleotide of any one of embodiments 1-8, wherein said synthetic polynucleotide is a messenger ribonucleic acid (mRNA)

[00305] Embodiment 10. The synthetic polynucleotide of any one of embodiments 1-9, wherein said synthetic polynucleotide further comprises a 3'- or 5'-noncoding region.

[00306] Embodiment 11. The synthetic polynucleotide of embodiment 10, wherein said 3'- or 5'-noncoding region enhances an expression or activity of said CFTR protein encoded by said synthetic polynucleotide within a cell.

[00307] Embodiment 12. The synthetic polynucleotide of any one of embodiments 1-11, wherein said synthetic polynucleotide further comprises a 5' cap structure.

[00308] Embodiment 13. The synthetic polynucleotide of any one of embodiments 1-12, wherein said 3' noncoding region comprises a poly adenosine tail.

[00309] Embodiment 14. The synthetic polynucleotide of embodiment 13, wherein said poly adenosine tail comprises at most 200 adenosines.

[00310] Embodiment 15. The synthetic polynucleotide of embodiment 13 or 14, wherein said poly adenosine tail improves a pharmacokinetic characteristic of said synthetic polynucleotide in a cell.

[00311] Embodiment 16. The synthetic polynucleotide of embodiment 15, wherein said poly adenosine tail improves a prolonged half-life of said synthetic polynucleotide in a cell.

[00312] Embodiment 17. A pharmaceutical composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid.

[00313] Embodiment 18. The pharmaceutical composition of embodiment 17, wherein said lipid composition comprises said ionizable cationic lipid at a molar percentage of about 5% to about 30%

[00314] Embodiment 19. The pharmaceutical composition of embodiment 17 or 18, wherein a mass or weight ratio of said ionizable cationic lipid to said synthetic polynucleotide is of no more than about 50:1, 40:1, 30:1, 20:1, 15:1 or 10:1

[00315] Embodiment 20. The pharmaceutical composition of any one of embodiments 17-19, wherein said SORT lipid is a permanently cationic lipid

[00316] Embodiment 21. The pharmaceutical composition of any one of embodiments 17-20, wherein said SORT lipid is a second ionizable cationic lipid

[00317] Embodiment 22. The pharmaceutical composition of embodiment 21, wherein said lipid composition comprises said SORT lipid at a molar percentage of about 5% to about 65%

[00318] Embodiment 23. The pharmaceutical composition of embodiment 21, wherein said lipid composition comprises said SORT lipid at a molar percentage of about 5% to about 30%

[00319] Embodiment 24. The pharmaceutical composition of any one of embodiments 17-23, wherein said lipid composition further comprises a zwitterionic lipid (e.g., a phospholipid)

[00320] Embodiment 25. The pharmaceutical composition of embodiment 24, wherein said lipid composition comprises said zwitterionic lipid at a molar percentage of about 5% to about 25%

[00321] Embodiment 26. The pharmaceutical composition of embodiment 24, wherein a molar ratio of said zwitterionic lipid to said synthetic polynucleotide is of no more than about 50:1, 40:1, 30:1, or 20:1

[00322] Embodiment 27. The pharmaceutical composition of any one of embodiments 17-26, wherein said lipid composition further comprises a steroid or steroid derivative

[00323] Embodiment 28. The pharmaceutical composition of embodiment 27, wherein said lipid composition comprises said steroid or steroid derivative at a molar percentage of about 15% to about 46%

[00324] Embodiment 29. The pharmaceutical composition of any one of embodiments 17-28, wherein said lipid composition further comprises a polymer-conjugated lipid (e.g., poly(ethylene glycol) (PEG)-conjugated lipid).

[00325] Embodiment 30. The pharmaceutical composition of embodiment 29, wherein said lipid composition comprises said polymer-conjugated lipid at a molar percentage of about 0.5% to about 10%, or about 1% to about 10%, or about 2% to about 10%.

[00326] Embodiment 31. The pharmaceutical composition of any one of embodiments 17-30, wherein a molar ratio of nitrogen in said lipid composition to phosphate in said synthetic polynucleotide (N/P ratio) is of no more than about 50:1, 40:1, 30:1, or 20:1

[00327] Embodiment 32. The pharmaceutical composition of embodiment 31, wherein said N/P ratio is from about 5:1 to about 30:1.

[00328] Embodiment 33. The pharmaceutical composition of any one of embodiments 17-32, wherein a mass or weight ratio of said synthetic polynucleotide to total lipids of said lipid composition is no more than about 1:20, 1:50, or 1:100.

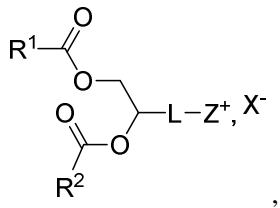
[00329] Embodiment 34. The pharmaceutical composition of any one of embodiments 17-33, wherein said SORT lipid comprises a permanently positively charged moiety (e.g., a quaternary ammonium ion).

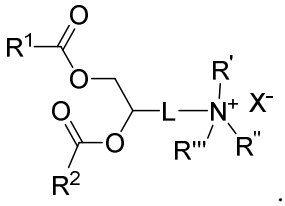
[00330] Embodiment 35. The pharmaceutical composition of embodiment 34, wherein said SORT lipid comprises a counterion.

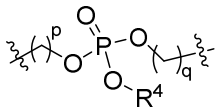
[00331] Embodiment 36. The pharmaceutical composition of any one of embodiments 17-35, wherein said SORT lipid is a phosphocholine lipid (e.g., saturated or unsaturated).

[00332] Embodiment 37. The pharmaceutical composition of any one of embodiments 36, wherein said SORT lipid is an ethylphosphocholine.

[00333] Embodiment 38. The pharmaceutical composition of any one of embodiments 17-37, wherein said SORT lipid comprises a headgroup having a structural formula: , wherein L is a (e.g., biodegradable) linker; Z^+ is positively charged moiety (e.g., a quaternary ammonium ion); and X^- is a counterion.

[00334] Embodiment 39. The pharmaceutical composition of embodiment 38, wherein said SORT lipid has a structural formula: , wherein R^1 and R^2 are each independently an optionally substituted C_6-C_{24} alkyl, or an optionally substituted C_6-C_{24} alkenyl.

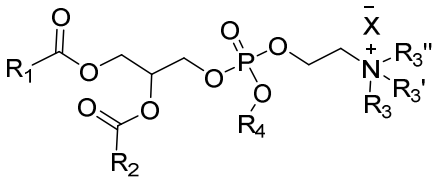
[00335] Embodiment 40. The pharmaceutical composition of embodiment 38, wherein said SORT lipid has a structural formula: .

[00336] Embodiment 41. The pharmaceutical composition of embodiment 40, wherein L is , wherein:

p and q are each independently 1, 2, or 3; and

R^4 is an optionally substituted C_1-C_6 alkyl.

[00337] Embodiment 42. The pharmaceutical composition of embodiment 38, wherein said SORT lipid has a structural formula:



(IA), wherein:

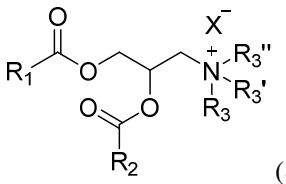
R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq 6)}$ or substituted $alkyl_{(C\leq 6)}$;

R_4 is $alkyl_{(C\leq 6)}$ or substituted $alkyl_{(C\leq 6)}$; and

X^- is a monovalent anion.

[00338] Embodiment 43. The pharmaceutical composition of any one of embodiments 17-35, wherein said SORT lipid has a structural formula:



(S-I')

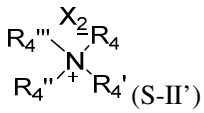
wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$;

X^- is a monovalent anion.

[00339] Embodiment 44. The pharmaceutical composition of any one of embodiments 17-35, wherein said SORT lipid has a structural formula:


(S-II')

wherein:

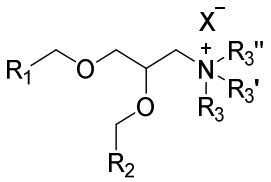
R_4 and R_4' are each independently $alkyl_{(C6\text{-}C24)}$, $alkenyl_{(C6\text{-}C24)}$, or a substituted version of either group;

R_4'' is $alkyl_{(C\leq24)}$, $alkenyl_{(C\leq24)}$, or a substituted version of either group;

R_4''' is $alkyl_{(C1\text{-}C8)}$, $alkenyl_{(C2\text{-}C8)}$, or a substituted version of either group; and

X_2 is a monovalent anion.

[00340] Embodiment 45. The pharmaceutical composition of any one of embodiments 17-35, wherein said SORT lipid has a structural formula:

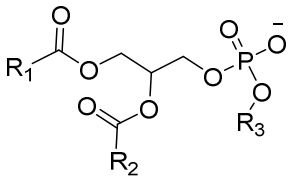

(S-III), wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$; and

X^- is a monovalent anion.

[00341] Embodiment 46. The pharmaceutical composition of any one of embodiments 17-35, wherein said SORT lipid has a structural formula:


(S-IV), wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3 is hydrogen, $alkyl_{(C\leq6)}$, or substituted $alkyl_{(C\leq6)}$, or $-Y_1-R_4$, wherein:

Y_1 is $alkanediyl_{(C\leq6)}$ or substituted $alkanediyl_{(C\leq6)}$; and

R_4 is $acyloxy_{(C\leq8\text{-}24)}$ or substituted $acyloxy_{(C\leq8\text{-}24)}$.

[00342] Embodiment 47. The pharmaceutical composition of any one of embodiments 17-46, wherein said pharmaceutical composition is an aerosol composition.

[00343] Embodiment 48. The pharmaceutical composition of embodiment 45, wherein said aerosol composition has a droplet size from 0.5 micron (μm) to 10 μm.

[00344] Embodiment 49. The pharmaceutical composition of embodiment 45, wherein said aerosol composition has a median droplet size from 0.5 μm to 10 μm.

[00345] Embodiment 50. The pharmaceutical composition of embodiment 45, wherein said aerosol composition has an average droplet size from 0.5 μm to 10 μm.

[00346] Embodiment 51. The pharmaceutical composition of any one of embodiments 17-50, wherein said pharmaceutical composition is formulated for aerosol administration

[00347] Embodiment 52. The pharmaceutical composition of any one of embodiments 17-51, wherein said pharmaceutical composition is formulated for apical delivery.

[00348] Embodiment 53. The pharmaceutical composition of any one of embodiments 17-52, wherein said pharmaceutical composition is formulated for nebulization.

[00349] Embodiment 54. A method for enhancing an expression or activity of cystic fibrosis transmembrane conductance regulator (CFTR) protein in a cell, the method comprising:

contacting said cell with a composition comprising a synthetic polynucleotide assembled with a lipid composition, wherein said synthetic polynucleotide encodes a CFTR protein; and wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after contacting, optionally wherein said therapeutically effective activity of said functional variant of CFTR protein is determined by measuring a change in a transepithelial ion transport characteristic of a plurality of cells comprising said cell as compared to that of a reference plurality of cells in absence of said contacting.

[00350] Embodiment 55. The method of embodiment 54, wherein said contacting is repeated

[00351] Embodiment 56. The method of embodiment 55, wherein said contacting is at least once a week

[00352] Embodiment 57. The method of embodiment 55, wherein said contacting is at least twice a week

[00353] Embodiment 58. The method of any one of embodiments 54-57, wherein the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after each contacting

[00354] Embodiment 59. The method of any one of embodiments 54-58, wherein said contacting is a first contacting, and wherein the method comprises a second contacting, optionally, performed at least 1, 2, or 3 day(s) after said first contacting

[00355] Embodiment 60. The method of embodiment 59, further comprising a third contacting, optionally wherein said third contacting is performed at least 1, 2, or 3 day(s) after said second contacting

[00356] Embodiment 61. The method of embodiment 59, wherein the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after said second contacting

[00357] Embodiment 62. The method of embodiment 60, wherein the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after said third contacting

[00358] Embodiment 63. The method of any one of embodiments 54-62, wherein said contacting comprises administering to a subject said composition comprising said synthetic polynucleotide assembled with said lipid composition

[00359] Embodiment 64. The method of embodiment 63, wherein said subject is a mammal.

[00360] Embodiment 65. The method of embodiment 63, wherein said subject is a human.

[00361] Embodiment 66. The method of any one of embodiments 63-65, wherein said administering comprises inhalation by nebulization.

[00362] Embodiment 67. The method of any one of embodiments 54-66, wherein said composition in each contacting is identical

[00363] Embodiment 68. The method of any one of embodiments 54-67, wherein said cell is a lung airway cell.

[00364] Embodiment 69. The method of embodiment 68, wherein said cell is a lung secretory cell

[00365] Embodiment 70. The method of embodiment 68 or 69, wherein said cell is a bronchial epithelial cell

[00366] Embodiment 71. The method of any one of embodiments 54-70, wherein said cell is undifferentiated.

[00367] Embodiment 72. The method of any one of embodiments 54-70, wherein said cell is differentiated.

[00368] Embodiment 73. The method of any one of embodiments 54-72, wherein said cell is derived from said subject.

[00369] Embodiment 74. The method of any one of embodiments 54-73, wherein said contacting is *in vivo*.

[00370] Embodiment 75. The method of any one of embodiments 54-73, wherein said contacting is *in vitro*.

[00371] Embodiment 76. The method of any one of embodiments 54-73, wherein said contacting is *ex vivo*.

[00372] Embodiment 77. The method of any one of embodiments 54-76, wherein said functional variant of CFTR protein is a wild-type CFTR protein

[00373] Embodiment 78. The method of any one of embodiments 54-77, wherein said functional variant of CFTR protein is a full-length CFTR protein

[00374] Embodiment 79. The method of any one of embodiments 54-78, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least about 5 micro-Ampere (μA), e.g., as determined in an *in vitro* assay.

[00375] Embodiment 80. The method of embodiment 79, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from about 5 micro-Ampere (μA) to about 30 μA.

[00376] Embodiment 81. The method of any one of embodiments 54-80, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$), e.g., as determined in an *in vitro* assay.

[00377] Embodiment 82. The method of embodiment 81, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$) to about 20 $\mu A \cdot cm^{-2} \cdot min^{-1}$.

[00378] Embodiment 83. The method of any one of embodiments 54-82, wherein the method increases an amount or activity of said functional variant of CFTR protein in said cell (e.g., by at least about 1.1-fold) relative to a corresponding control (e.g., that of a corresponding cell absent said contacting).

[00379] Embodiment 84. The method of any one of embodiments 54-83, wherein the method enhances (e.g., chloride) ion transport in said cell (e.g., by at least about 1.1-fold) relative to a corresponding control (e.g., that of a corresponding cell absent said contacting).

[00380] Embodiment 85. The method of any one of embodiments 54-84, wherein said subject exhibits or is determined to exhibit a mutation in a cystic fibrosis transmembrane conductance regulator (CFTR) gene.

[00381] Embodiment 86. The method of embodiment 85, wherein said mutation is a loss-of-function mutation.

[00382] Embodiment 87. The method of embodiment 85 or, wherein said mutation is a nonsense or frameshift mutation.

[00383] Embodiment 88. The method of any one of embodiments 85-87, wherein said mutation is in one or more of exons 11-27 of CFTR gene.

[00384] Embodiment 89. The method of any one of embodiments 85-88, wherein said mutation is R553X, G542X, F508del, or R1162X, or a combination thereof; for example, said mutation is G542X or F508del.

[00385] Embodiment 90. A method for lung secretory cell or lung basal cell delivery, comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a therapeutically effective amount or activity of said synthetic polynucleotide in a lung secretory cell of said subject, optionally wherein said therapeutically effective activity of said synthetic polynucleotide is determined by measuring a change in a transepithelial ion transport characteristic of a lung comprising said lung secretory cell or lung basal cell as compared to that of a reference lung in absence of said contacting.

[00386] Embodiment 91. A method for lung secretory cell or lung basal cell delivery, comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a greater therapeutic amount or activity of said synthetic polynucleotide in a lung secretory cell or lung basal cell of said subject as compared to that in a lung non-secretory cell or non basal cell of said subject.

[00387] Embodiment 92. The method of embodiment 91, wherein the method is characterized in that: (i) at least about 50%, 55%, or 60% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected in lung secretory cells, lung basal cells, or a combination thereof; or (ii) no more than about 50%, 45%, or 40% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected in lung non-secretory cells, lung non-basal cells, or a combination thereof.

[00388] Embodiment 93. The method of embodiment 91 or 92, wherein said lung non-secretory cell is a lung ciliated cell

[00389] Embodiment 94. The method of any one of embodiments 91-93, wherein said lung non-secretory cell is a lung basal cell

[00390] Embodiment 95. A method for lung secretory cell delivery, comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a therapeutic amount or activity of said synthetic polynucleotide in at least 5% of lung secretory cells of said subject.

[00391] Embodiment 96. The method of embodiment 95, wherein said administering comprises administering to a lung of said subject said composition comprising said synthetic polynucleotide assembled with said lipid composition.

[00392] Embodiment 97. The method of embodiment 95 or 96, wherein said lung secretory cell is a club cell or a goblet cell.

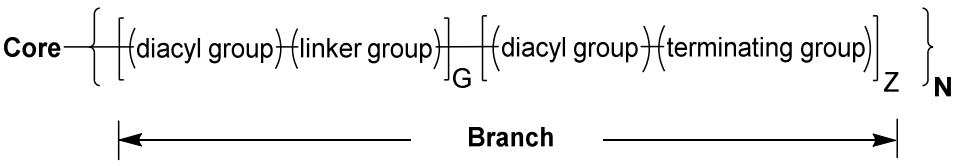
[00393] Embodiment 98. A method for treating a subject having or suspected of having a cystic fibrosis transmembrane conductance regulator (CFTR)-associated condition, the method comprising administering to said subject a pharmaceutical composition of any one of embodiments 17-53 and 101-151.

[00394] Embodiment 99. The method of embodiment 98, wherein said CFTR-associated condition is cystic fibrosis, hereditary emphysema, or chronic obstructive pulmonary disease (COPD).

[00395] Embodiment 100. The method of embodiment 98 or 99, wherein said administering comprises local administration (e.g., nebulization).

[00396] Embodiment 101. The composition of any one of Embodiments 17-53, wherein said SORT lipid is selected from those set forth in **Table 8**, or pharmaceutically acceptable salts thereof, or a subset of the lipids and the pharmaceutically acceptable salts thereof.

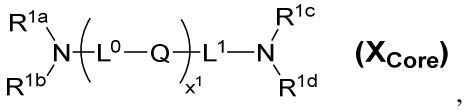
[00397] Embodiment 102. The composition of any one of Embodiments 17-53 and 101, wherein the ionizable cationic lipid is a dendrimer or dendron of a generation (g) having a structural formula:



(X),

or a pharmaceutically acceptable salt thereof, wherein:

(a) the core comprises a structural formula (X_{Core}):



(X_{Core}),

wherein:

Q is independently at each occurrence a covalent bond, -O-, -S-, -NR^2-, or -$CR^{3a}R^{3b}$-;

R^2 is independently at each occurrence R^{1g} or -L^2-$NR^{1e}R^{1f}$;

R^{3a} and R^{3b} are each independently at each occurrence hydrogen or an optionally substituted (e.g., C_1-C_6, such as C_1-C_3) alkyl;

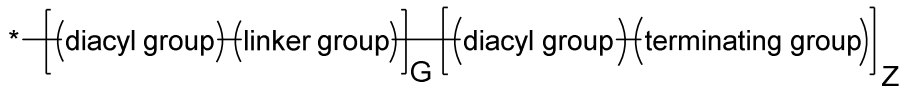
R^{1a}, R^{1b}, R^{1c}, R^{1d}, R^{1e}, R^{1f}, and R^{1g} (if present) are each independently at each occurrence a point of connection to a branch, hydrogen, or an optionally substituted (e.g., C_1-C_{12}) alkyl;

L^0, L^1, and L^2 are each independently at each occurrence selected from a covalent bond, (e.g., C_1-C_{12}, such as C_1-C_6 or C_1-C_3) alkylene, (e.g., C_1-C_{12}, such as C_1-C_8 or C_1-C_6) heteroalkylene (e.g., C_2-C_8 alkyleneoxide, such as oligo(ethyleneoxide)), [(e.g., C_1-C_6) alkylene]-[(e.g., C_4-C_6) heterocycloalkyl]-[(e.g., C_1-C_6) alkylene], [(e.g., C_1-C_6) alkylene]-(arylene)-[(e.g., C_1-C_6) alkylene] (e.g., [(e.g., C_1-C_6) alkylene]-phenylene-[(e.g., C_1-C_6) alkylene]), (e.g., C_4-C_6) heterocycloalkyl, and arylene (e.g., phenylene); or,

alternatively, part of L^1 form a (e.g., C_4-C_6) heterocycloalkyl (e.g., containing one or two nitrogen atoms and, optionally, an additional heteroatom selected from oxygen and sulfur) with one of R^{1c} and R^{1d}; and

x^1 is 0, 1, 2, 3, 4, 5, or 6; and

(b) each branch of the plurality (N) of branches independently comprises a structural formula (X_{Branch}):



(X_{Branch}),

wherein:

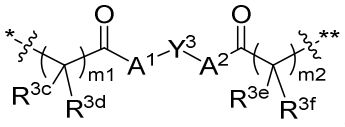
* indicates a point of attachment of the branch to the core;

g is 1, 2, 3, or 4;

$Z = 2^{(g-1)}$;

G=0, when g=1; or $G = \sum_{i=0}^{i=g-2} 2^i$, when g≠1;

(c) each diacyl group independently comprises a structural formula



, wherein:

* indicates a point of attachment of the diacyl group at the proximal end thereof;

** indicates a point of attachment of the diacyl group at the distal end thereof;

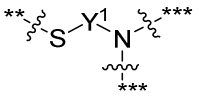
Y^3 is independently at each occurrence an optionally substituted (e.g., C_1-C_{12}); alkylene, an optionally substituted (e.g., C_1-C_{12}) alkenylene, or an optionally substituted (e.g., C_1-C_{12}) arenylene;

A^1 and A^2 are each independently at each occurrence -O-, -S-, or -NR^4-, wherein:

R^4 is hydrogen or optionally substituted (e.g., C_1-C_6) alkyl;

m^1 and m^2 are each independently at each occurrence 1, 2, or 3; and

R^{3c}, R^{3d}, R^{3e}, and R^{3f} are each independently at each occurrence hydrogen or an optionally substituted (e.g., C_1-C_8) alkyl; and

(d) each linker group independently comprises a structural formula 

, wherein:

** indicates a point of attachment of the linker to a proximal diacyl group;

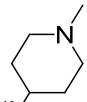
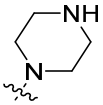
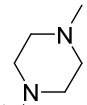
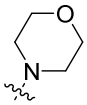
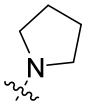
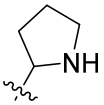
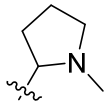
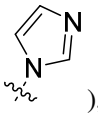
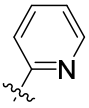
*** indicates a point of attachment of the linker to a distal diacyl group; and

Y_1 is independently at each occurrence an optionally substituted (e.g., C_1-C_{12}) alkylene, an optionally substituted (e.g., C_1-C_{12}) alkenylene, or an optionally substituted (e.g., C_1-C_{12}) arenylene; and

(e) each terminating group is independently selected from optionally substituted (e.g., C_1-C_{18}, such as C_4-C_{18}) alkylthiol, and optionally substituted (e.g., C_1-C_{18}, such as C_4-C_{18}) alkenylthiol.

[00398] Embodiment 103. The composition of Embodiment 102, wherein x^1 is 0, 1, 2, or 3.

[00399] Embodiment 104. The composition of Embodiment 102 or 103, wherein R^{1a}, R^{1b}, R^{1c}, R^{1d}, R^{1e}, R^{1f}, and R^{1g} (if present) are each independently at each occurrence a point of connection to a branch (e.g., as indicated by *), hydrogen, or C_1-C_{12} alkyl (e.g., C_1-C_8 alkyl, such as C_1-C_6 alkyl or C_1-C_3 alkyl), wherein the alkyl moiety is optionally substituted with one or more substituents each independently selected from -

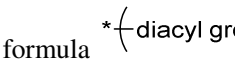
OH, C_4-C_8 (e.g., C_4-C_6) heterocycloalkyl (e.g., piperidinyl (e.g., , or ), *N*-(C_1-C_3 alkyl)-piperidinyl (e.g., ), piperazinyl (e.g., ), *N*-(C_1-C_3 alkyl)-piperadizinyl (e.g., ), morpholinyl (e.g., ), *N*-pyrrolidinyl (e.g., ), pyrrolidinyl (e.g., ), or *N*-(C_1-C_3 alkyl)-pyrrolidinyl (e.g., )), (e.g., C_6-C_{10}) aryl, and C_3-C_5 heteroaryl (e.g., imidazolyl (e.g., ), or pyridinyl (e.g., )).

[00400] Embodiment 105. The composition of Embodiment 104, wherein R^{1a}, R^{1b}, R^{1c}, R^{1d}, R^{1e}, R^{1f}, and R^{1g} (if present) are each independently at each occurrence a point of connection to a branch (e.g., as indicated by *), hydrogen, or C_1-C_{12} alkyl (e.g., C_1-C_8 alkyl, such as C_1-C_6 alkyl or C_1-C_3 alkyl), wherein the alkyl moiety is optionally substituted with one substituent -OH.

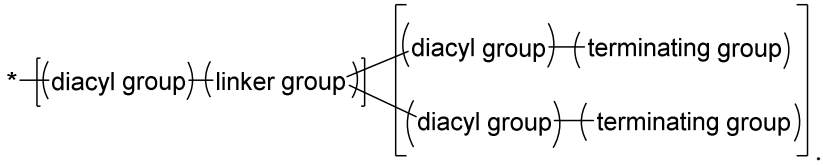
[00401] Embodiment 106. The composition of any one of Embodiments 102-105, wherein R^{3a} and R^{3b} are each independently at each occurrence hydrogen.

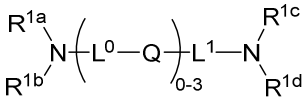
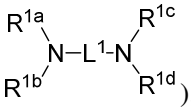
[00402] Embodiment 107. The composition of any one of Embodiments 102-106, wherein the plurality (N) of branches comprises at least 3 (e.g., at least 4, or at least 5) branches.

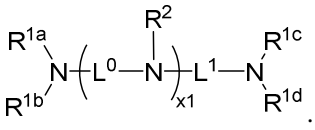
[00403] Embodiment 108. The composition of any one of Embodiments 102-107, wherein g=1; G=0; and Z=1.

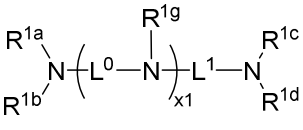
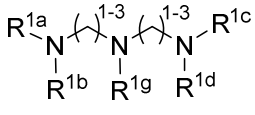
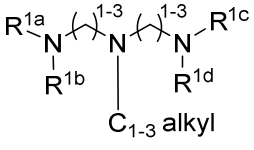
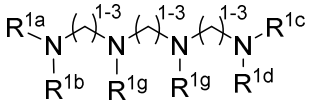
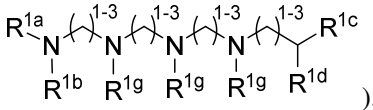
[00404] Embodiment 109. The composition of Embodiment 108, wherein each branch of the plurality of branches comprises a structural formula .

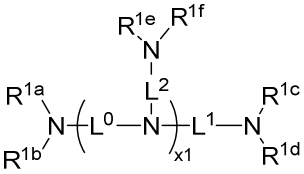
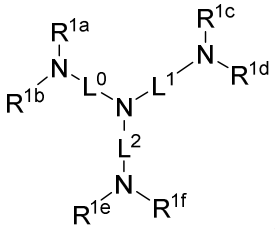
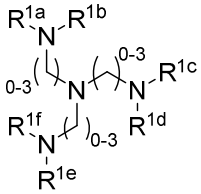
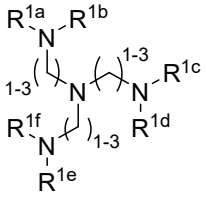
[00405] Embodiment 110. The composition of any one of Embodiments 102-107, wherein g=2; G=1; and Z=2.

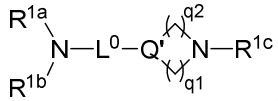
[00406] Embodiment 111. The composition of Embodiment 110, wherein each branch of the plurality of branches comprises a structural formula .

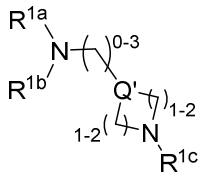
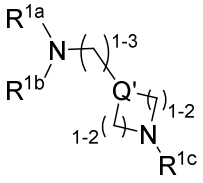
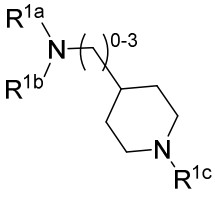
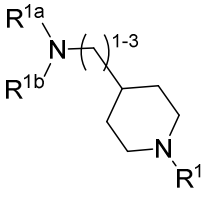
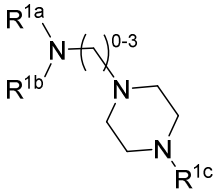
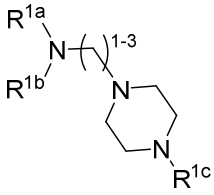
[00407] Embodiment 112. The composition of any one of Embodiments 102-111, wherein the core comprises a structural formula:  (e.g., ).

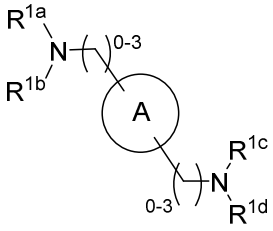
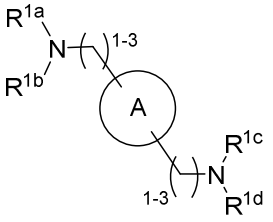
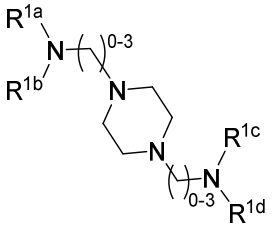
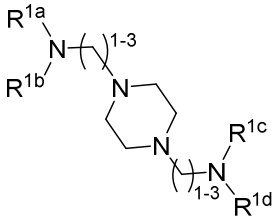
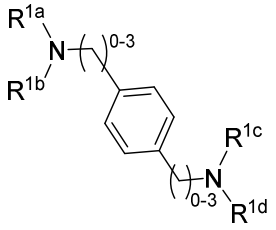
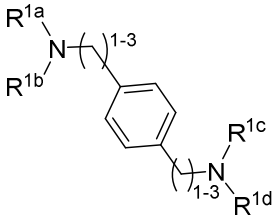
[00408] Embodiment 113. The composition of any one of Embodiments 102-111, wherein the core comprises a structural formula: .

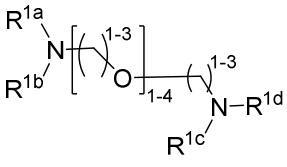
[00409] Embodiment 114. The composition of Embodiment 113, wherein the core comprises a structural formula:  (e.g., , , , or ).

[00410] Embodiment 115. The composition of Embodiment 113, wherein the core comprises a structural formula:  (e.g., , such as  or ).

[00411] Embodiment 116. The composition of any one of Embodiments 102-111, wherein the core comprises a structural formula: , wherein Q' is $-NR^{2}-$ or $-CR^{3a}R^{3b}-$; q^1 and q^2 are each independently 1 or 2.

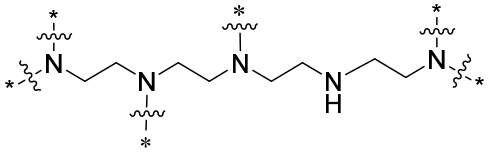
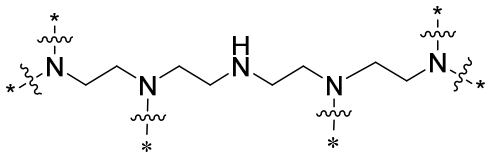
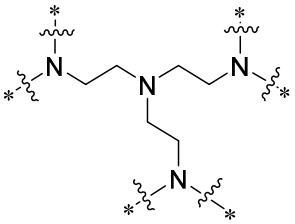
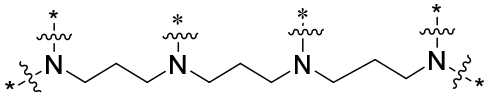
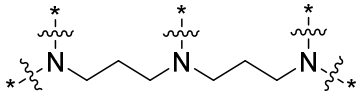
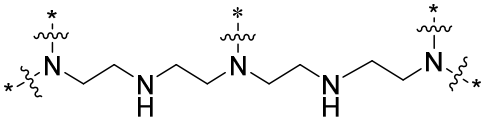
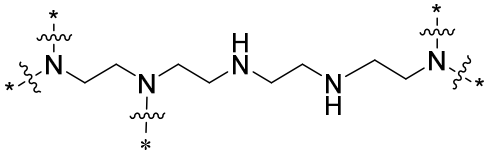
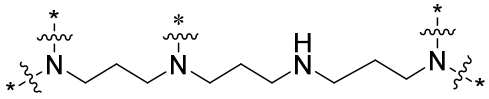
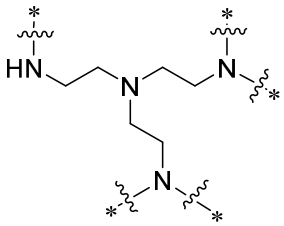
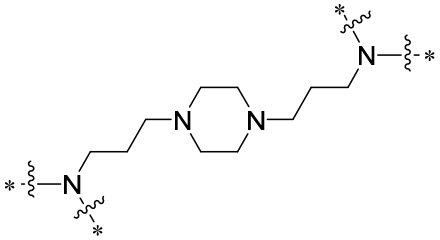
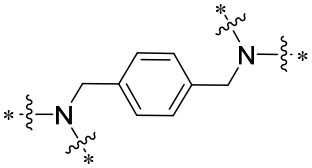
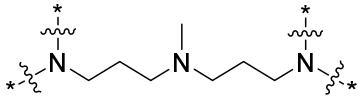
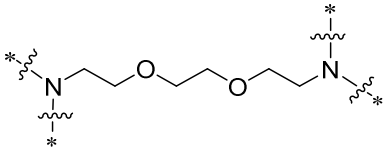
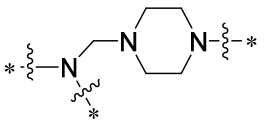
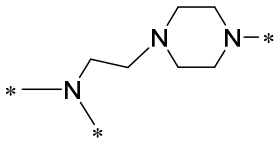
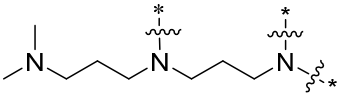
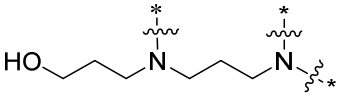
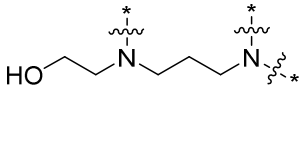
[00412] Embodiment 117. The composition of Embodiment 116, wherein the core comprises a structural formula:  or  (e.g., , , , or ).

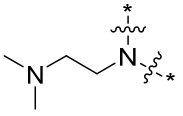
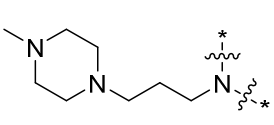
[00413] Embodiment 118. The composition of any one of Embodiments 102-111, wherein the core comprises a structural formula  or  (e.g., , , , or ), wherein ring A is an optionally substituted aryl or an optionally substituted (e.g., C_3-C_{12}, such as C_3-C_5) heteroaryl.

[00414] Embodiment 119. The composition of any one of Embodiments 102-111, wherein the core comprises has a structural formula .

[00415] Embodiment 120. The composition of any one of Embodiments 102-111, wherein the core is selected from those set forth in **Table 3** or a subset thereof.

[00416] Embodiment 121. The composition of any one of Embodiments 102-111, wherein the core comprises a structural formula selected from the group consisting of:


,

,


,
,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

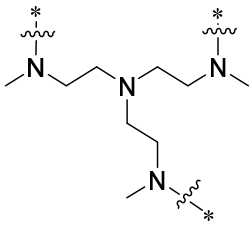
,

,

,

,

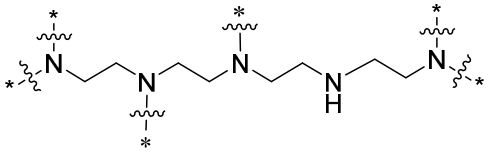
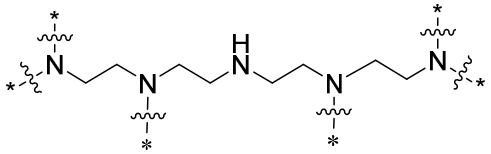
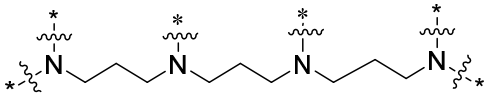
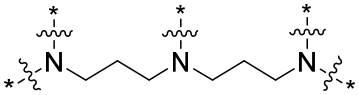
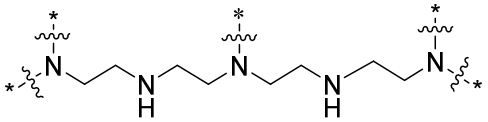
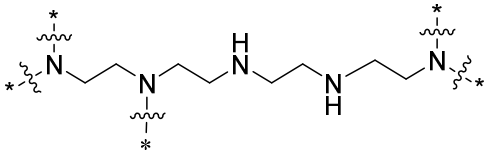
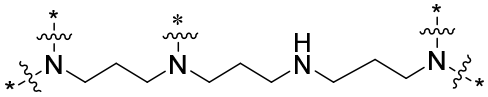
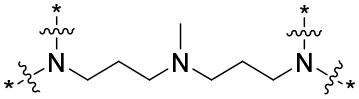
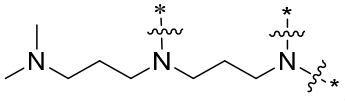
,

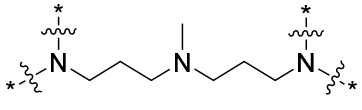
,

,


, and pharmaceutically acceptable salts thereof, wherein * indicates a point of attachment of the core to a branch of the plurality of branches.

[00417] Embodiment 122. The composition of any one of Embodiments 102-111, wherein the core comprises a structural formula selected from the group consisting of:

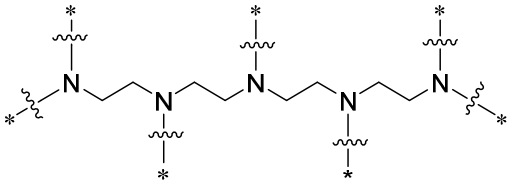

,

,

,

,

,

,

,

,

, and pharmaceutically acceptable salts thereof, wherein * indicates a point of attachment of the core to a branch of the plurality of branches.

[00418] Embodiment 123. The composition of any one of Embodiments 102-111, wherein the core has the structure , wherein * indicates a point of attachment of the core to a branch of the plurality of branches or H.

[00419] Embodiment 124. The composition of Embodiment 123, wherein at least 2 branches are attached to the core.

[00420] Embodiment 125. The composition of Embodiment 123, wherein at least 3 branches are attached to the core.

[00421] Embodiment 126. The composition of Embodiment 123, wherein at least 4 branches are attached to the core.

[00422] Embodiment 127. The composition of any one of Embodiments 102-111, wherein the core has the structure , wherein * indicates a point of attachment of the core to a branch of the plurality of branches or H.

[00423] Embodiment 128. The composition of Embodiment 127, wherein at least 4 branches are attached to the core.

[00424] Embodiment 129. The composition of Embodiment 127, wherein at least 5 branches are attached to the core.

[00425] Embodiment 130. The composition of Embodiment 127, wherein at least 6 branches are attached to the core.

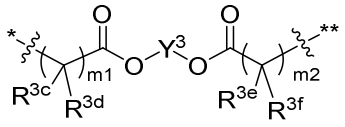
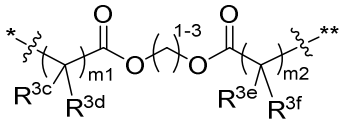
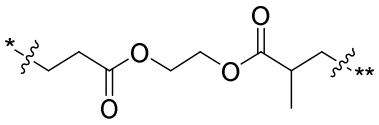
[00426] Embodiment 131. The composition of any one of Embodiments 102-130, wherein A^1 is -O- or -NH-.

[00427] Embodiment 132. The composition of Embodiment 131, wherein A^1 is -O-.

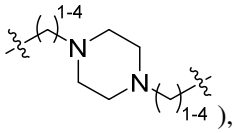
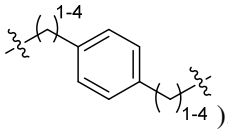
[00428] Embodiment 133. The composition of any one of Embodiments 102-132, wherein A^2 is -O- or -NH-.

[00429] Embodiment 134. The composition of any Embodiment 133, wherein A^2 is -O-.

[00430] Embodiment 135. The composition of any one of Embodiments 102-134, wherein Y^3 is C_1-C_{12} (e.g., C_1-C_6, such as C_1-C_3) alkylene.

[00431] Embodiment 136. The composition of any one of Embodiments 102-135, wherein the diacyl group independently at each occurrence comprises a structural formula  (e.g., , such as ), optionally wherein R^{3c}, R^{3d}, R^{3e}, and R^{3f} are each independently at each occurrence hydrogen or C_1-C_3 alkyl.

[00432] Embodiment 137. The composition of any one of Embodiments 102-136, wherein L^0, L^1, and L^2 are each independently at each occurrence selected from a covalent bond, C_1-C_6 alkylene (e.g., C_1-C_3 alkylene), C_2-C_{12} (e.g., C_2-C_8) alkyleneoxide (e.g., oligo(ethyleneoxide), such as -$(CH_2CH_2O)_{1-4}$-

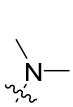
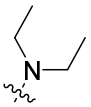
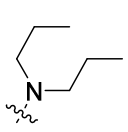
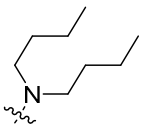
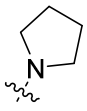
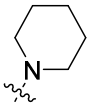
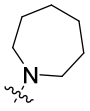
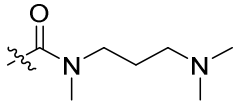
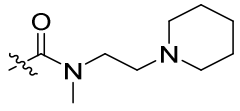
((CH_2CH_2)-), [(C_1-C_4) alkylene]-[(C_4-C_6) heterocycloalkyl]-[(C_1-C_4) alkylene] (e.g., ), and [(C_1-C_4) alkylene]-phenylene-[(C_1-C_4) alkylene] (e.g., ).

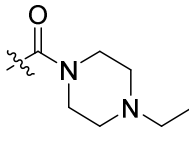
[00433] Embodiment 138. The composition of Embodiment 137, wherein L^0, L^1, and L^2 are each independently at each occurrence selected from C_1-C_6 alkylene (e.g., C_1-C_3 alkylene), -(C_1-C_3 alkylene-O)$_{1-4}$-(C_1-C_3 alkylene), -(C_1-C_3 alkylene)-phenylene-(C_1-C_3 alkylene)-, and -(C_1-C_3 alkylene)-piperazinyl-(C_1-C_3 alkylene)-.

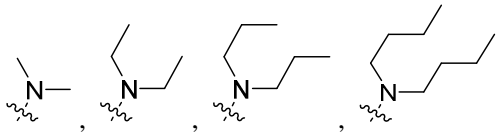
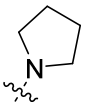
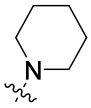
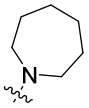
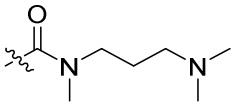
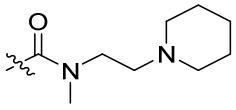
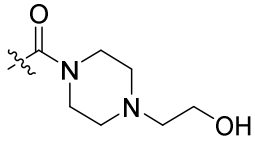
[00434] Embodiment 139. The composition of Embodiment 137, wherein L^0, L^1, and L^2 are each independently at each occurrence C_1-C_6 alkylene (e.g., C_1-C_3 alkylene).

[00435] Embodiment 140. The composition of Embodiment 137, wherein L^0, L^1, and L^2 are each independently at each occurrence C_2-C_{12} (e.g., C_2-C_8) alkyleneoxide (e.g., -(C_1-C_3 alkylene-O)$_{1-4}$-(C_1-C_3 alkylene)).

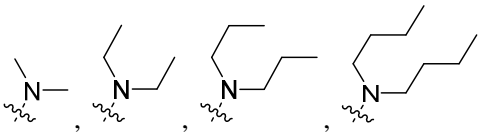
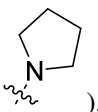
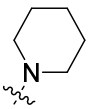
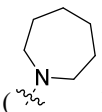
[00436] Embodiment 141. The composition of Embodiment 137, wherein L^0, L^1, and L^2 are each independently at each occurrence selected from [(C_1-C_4) alkylene]-[(C_4-C_6) heterocycloalkyl]-[(C_1-C_4) alkylene] (e.g., -(C_1-C_3 alkylene)-phenylene-(C_1-C_3 alkylene)-) and [(C_1-C_4) alkylene]-[(C_4-C_6) heterocycloalkyl]-[(C_1-C_4) alkylene] (e.g., -(C_1-C_3 alkylene)-piperazinyl-(C_1-C_3 alkylene)-).

[00437] Embodiment 142. The composition of any one of Embodiments 102-141, wherein each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkenylthiol or C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol, wherein the alkyl or alkenyl moiety is optionally substituted with one or more substituents each independently selected from halogen, C_6-C_{12} aryl (e.g., phenyl), C_1-C_{12} (e.g., C_1-C_8) alkylamino (e.g., C_1-C_6 mono-alkylamino (such as -$NHCH_2CH_2CH_2CH_3$) or C_1-C_8 di-alkylamino (such as , , , )), C_4-C_6 *N*-heterocycloalkyl (e.g., *N*-pyrrolidinyl (), *N*-piperidinyl (), *N*-azepanyl ()), -OH, -C(O)OH, –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_1-C_{12} alkyl-amino (e.g., mono- or di-alkylamino)) (e.g., ), –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_4-C_6 *N*-heterocycloalkyl) (e.g., ), –C(O)–(C_1-C_{12} alkylamino (e.g.,

mono- or di-alkylamino)), and –C(O)–(C_4-C_6 *N*-heterocycloalkyl) (e.g., ), wherein the C_4-C_6 *N*-heterocycloalkyl moiety of any of the preceding substituents is optionally substituted with C_1-C_3 alkyl or C_1-C_3 hydroxyalkyl.

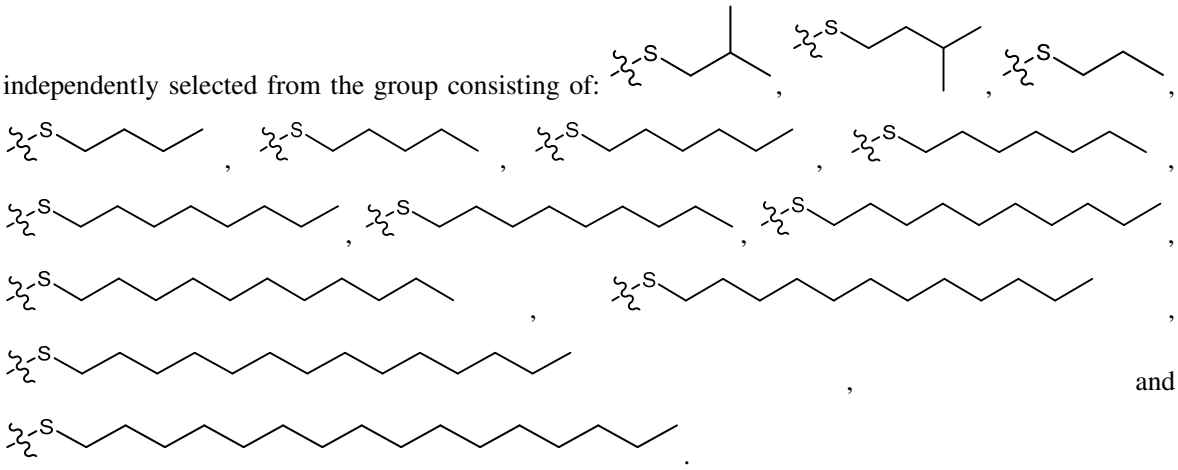
[00438] Embodiment 143. The composition of Embodiment 142, wherein each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol, wherein the alkyl moiety is optionally substituted with one or more (e.g., one) substituents each independently selected from C_6-C_{12} aryl (e.g., phenyl), C_1-C_{12} (e.g., C_1-C_8) alkylamino (e.g., C_1-C_6 mono-alkylamino (such as -$NHCH_2CH_2CH_2CH_3$) or C_1-C_8 di-alkylamino (such as )), C_4-C_6 *N*-heterocycloalkyl (e.g., *N*-pyrrolidinyl (), *N*-piperidinyl (), *N*-azepanyl ()), -OH, -C(O)OH, –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_1-C_{12} alkylamino (e.g., mono- or di-alkylamino)) (e.g., ), –C(O)N(C_1-C_3 alkyl)–(C_1-C_6 alkylene)–(C_4-C_6 *N*-heterocycloalkyl) (e.g., ), and –C(O)–(C_4-C_6 *N*-heterocycloalkyl) (e.g., ), wherein the C_4-C_6 *N*-heterocycloalkyl moiety of any of the preceding substituents is optionally substituted with C_1-C_3 alkyl or C_1-C_3 hydroxyalkyl.

[00439] Embodiment 144. The composition of Embodiment 143, wherein each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol, wherein the alkyl moiety is optionally substituted with one substituent -OH.

[00440] Embodiment 145. The composition of Embodiment 143, wherein each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol, wherein the alkyl moiety is optionally substituted with one substituent selected from C_1-C_{12} (e.g., C_1-C_8) alkylamino (e.g., C_1-C_6 mono-alkylamino (such as -$NHCH_2CH_2CH_2CH_3$) or C_1-C_8 di-alkylamino (such as )) and C_4-C_6 *N*-heterocycloalkyl (e.g., *N*-pyrrolidinyl (), *N*-piperidinyl (), *N*-azepanyl ()).

[00441] Embodiment 146. The composition of Embodiment 142, wherein each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkenylthiol or C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol.

[00442] Embodiment 147. The composition of Embodiment 146, wherein each terminating group is independently C_1-C_{18} (e.g., C_4-C_{18}) alkylthiol.

[00443] Embodiment 148. The composition of Embodiment 147, wherein each terminating group is independently selected from the group consisting of: , and .

[00444] Embodiment 149. The composition of any one of Embodiments 102-141, wherein each terminating group is independently selected from those set forth in **Table 5** or a subset thereof.

[00445] Embodiment 150. The composition of any one of Embodiments 17-53 and 101, wherein the ionizable cationic lipid is selected from those set forth in **Table 6**, or pharmaceutically acceptable salts thereof, or a subset of the lipids and the pharmaceutically acceptable salts thereof.

[00446] Embodiment 151. The composition of any one of Embodiments 17-53 and 101, wherein the ionizable cationic lipid is selected from those set forth in **Table 6** or **Table 7**, or pharmaceutically acceptable salts thereof, or a subset of the lipids and the pharmaceutically acceptable salts thereof.

[00447] Embodiment 101. The method of any one of claims 54-97, wherein the composition is according to any one of Embodiments 17-53 and 101-151

EXAMPLES

Example 1: Preparation of DOTAP or DODAP Modified Lipid Nanoparticles

[00448] Lipid nanoparticles (LNPs) are the most efficacious carrier class for *in vivo* nucleic acid delivery. Historically, effective LNPs are composed of 4 components: an ionizable cationic lipid, zwitterionic phospholipid, cholesterol, and lipid poly(ethylene glycol) (PEG). However, these LNPs result in only general delivery of nucleic acids, rather than organ or tissue targeted delivery. LNPs administered by IV typically deliver RNAs only to the liver. Therefore, new formulations of LNPs were sought in an effort to provide targeted nucleic acid delivery.

[00449] The four canonical types of lipids were mixed in a 15:15:30:3 molar ratio, with or without the addition of a permanently cationic lipid. Briefly, LNPs were prepared by mixing a dendrimer lipid (ionizable cationic), DOPE (zwitterionic), cholesterol, DMG-PEG, and DOTAP (permanently cationic).

Alternatively DOTAP can be substituted for DODAP to generate a LNP comprising DODAP. The structure of DOTAP and DODAP are shown in **FIG. 1**. Various dendrimer lipids that may be used are shown in **FIG. 2**.

[00450] For preparation of the LNP formulation, a dendrimer lipid, DOPE, Cholesterol and DMG-PEG were dissolved in ethanol at desired molar ratios. The mRNA was dissolved in citrate buffer (10 mM, pH 4.0). The mRNA was then diluted into the lipids solution to achieve a weight ratio of 40:1 (total lipids:mRNA) by rapidly mixing the mRNA into the lipids solution at a volume ratio of 3:1 (mRNA:lipids, v/v). This solution was then incubated for 10 min at room temperature. For formation of DOTAP modified LNP formulations, mRNA was dissolved in 1 × PBS or citrate buffer (10 mM, pH 4.0), and mixed rapidly into ethanol containing 5A2-SC8, DOPE, Cholesterol, DMG-PEG and DOTAP, fixing the weight ratio of 40:1 (total lipids:mRNA) and volume ratio of 3:1 (mRNA:lipids). Formulations are named X% DOTAP Y (or X%DODAP Y) where X represents the DOTAP (or DODAP) molar percentage in total lipids, and Y represents the type of dendrimer lipid. Alternatively, formulation may be named Y X%DOTAP or Y X%DODAP where X represents the DOTAP (or DODAP) molar percentage in total lipids, and Y represents the type of dendrimer lipid.

Example 2: SORT LNP Stability:

[00451] LNPs were tested for stability. 5A2-SC8 20% DODAP ("Liver-SORT) and 5A2-SC8 50% DOTAP ("Lung-SORT") were generated using either a microfluidic mixing method or a cross/tee mixing method. The different LNP formulations were characterized by size, polydispersity index (PDI) and zeta-potential, as assessed by dynamic light scattering, 3 separate times for each formulation. The characteristics of the LNPs are show in **Table 9**.

TABLE 9: SORT LNP characteristics

	Size (nm)	PDI	Zeta (mV)	Encapsulation Efficiency (%)
Lung-SORT - microfluidic	82.3	0.10	3.0	100
Lung-SORT - cross/tee mixing	78.1	0.09	2.2	100
Liver-SORT - microfluidic	59.1	0.10	-2.3	97
Liver-SORT - cross/tee mixing	60.0	0.11	-30	96

[00452] The encapsulation efficiency was tested using a Ribogreen RNA assay (Zhao *et al.*, 2016). Briefly, mRNA was encapsulated with > 95% efficiency in LNPs when the mRNA was dissolved in acidic buffer (10 mM citrate, pH 4). The characteristics were observed over 28 days for the two types of LNPs (5A2-SC8 20% DODAP ("Liver-SORT") and 5A2-SC8 50% DOTAP ("Lung-SORT")). **FIG. 6** shows the changes of the characteristics of the LNP over the course of 28 days.

[00453] In addition, to the measure of the stability of the LNPs in solution, the stability of the LNPs and resulting mRNA expression was observed in mice. Briefly, mice were injected intravenously with 0.1

mg/kg and observed in vivo. Luciferin was added 5 hrs after injection and visualized. As shown in **FIG. 7**, the Lung-SORT LNP generated tissue specific radiance in the lungs which remained high even after 14 days with a slight decay in signal by the 21st and 28th day. **FIG. 8** shows images of the organs of the mouse at specific times periods after being treated with Lung-SORT or Liver-SORT.

Example 3: Expression of TR (Tomato Red) mRNA in different cell types

[00454] Expression of TR (Tomato Red) mRNA in different cell types in hBE cultures (human bronchial epithelial cultures) was analyzed. TR mRNA was loaded into either 20%DODAP 4A3-SC7 LNP or 10%DOTAP 5A2-SC8 LNPs and delivered into well-differentiated human bronchial epithelial cultures using apical bolus dosing (upper panel) or aerosol delivery (bottom panel). TR protein expression in various cell-types was observed and the percent of TR positive cells in different cell-type was plotted. As shown in the top panel of **FIG. 3**, the 20% DODAP 4A3-SC7 LNPs preferentially caused secretory cells to express TR, while 10%DOTAP 5A2-SC8 LNPs cause the ciliated cells to preferentially express TR. This preferential delivery may allow a treatment delivered to the lungs to preferentially affect a specific cell type in the lungs. The TR mRNA was also loaded into LNPs without the SORT lipid (e.g. DODAP or DOTAP) to identify how the DODAP or DOTAP affected the potency. As shown in the bottom panel of **FIG. 3**, the LNPs comprising DOTAP or DODAP showed increased TR expression compared to their corresponding LNP without DOTAP or DODAP.

Example 4: Luciferase activity and histopathology from LNPs delivered via inhaled aerosol

[00455] Luc mRNA was loaded into a number of LNPs including LNPs comprising a SORT lipid and a dendrimer. LNPs of 4A3-SC7 20% DODAP, 4A3-SC7 10% DODAP, 5A2-SC8, and 5A2-SC8 10% DOTAP were generated and loaded with Luc mRNA. 0.4/2/8 mg of LNP-formulated Luc2 mRNA (1 mg/ml) was delivered into a pie chamber by nebulization (Aerogen solo), with an estimated (not measured) per mouse delivered dose of 0.01, 0.06 or 0.22 mg/kg. The mice were 7 week old B6 male albino mice. Luciferin was administered to the mice 5 hrs after delivery of the LNPs. The luciferase activity was detected as a measure of delivery to the target. **FIG. 4** shows the distribution and expression of the luciferase in the mice demonstrating the expression was successful and delivery of the LNPs may be performed using inhaled aerosol delivery.

Example 5: Toxicity of EPC containing LNPs

[00456] LNPs comprising ethylphosphocholine (EPC) in place of DOTAP or DODAP were tested for toxicity by using apical bolus dosing on human bronchial epithelial cells. The % of lactate dehydrogenase (LDH) that was released was used as a metric of cellular death and indicative of the toxicity of the LNP. The release of LDH was detected prior to treatment (pre-treatment) and 24 post treatment. As shown in **FIG. 5**, the treatment of 50% DOTAP LNP resulted in an ~15% LDH release whereas EPC didn't show a significant %LDH release. Importantly, DOTAP and EPC have a similar quaternary amine moiety, indicating that the activity for cell targeting may be similar, but that EPC is considerably less toxic.

Example 6: Optimization of CFTR mRNA

[00457] mRNA sequences of CFTR mRNA were optimized before production. Briefly, wild type sequence was firstly subject to optimization according to % codon-usage, elimination of UU and UA hydrolysis hotspots, elimination of U-containing codons other than TAC and TTC, and elimination of unwanted or accidental restriction to obtain a codon optimized ORF. The example structure of CFTR mRNA is as shown in **FIG. 9A**.. A variation of Kozak sequence AUGC was used to maintain the wild-type amino acid composition. Second residue after N-terminal Met is Gln (coded by CAG). ORF sequence including **SEQ ID NO. 1**. was placed after Kozak sequence, followed by polyA.

[00458] DNA corresponding to the gene of CFTR was synthesized at GenScript. pUC57/CFTR was digested with BstBI (the ORF for CFTR is codon optimized). Standard *in vitro* transcription procedure was used for RNA production utilizing either unmodified or modified nucleotides. Capping reaction was carried out using Vaccinia Virus capping system and cap 2'-O-methyl transferase. Full Length CFTR mRNA was further subjected to fragment analysis, as shown in **FIG. 9B** and demonstrated the synthesis of full length CFTR mRNA with unmodified or modified nucleotides. Additional routine QC parameters tested include concentration, sterility, mRNA function, residual plasmid and bacterial DNA, dsRNA, and endotoxin.

Example 7: Dose-dependent expression of CFTR protein in FRT cells

[00459]. Briefly, CFTR mRNAs of the present application were transfected into FRT cells. As shown in **FIG. 10A and 10B**, transfection of CFTR mRNAs induced CFTR protein expression in FRTs. The FRT cells were lysed and the lysate was collected. Gel electrophoresis was performed on the lysate samples and subsequently Western blotted with anti-CFTR. Both HA-tagged CFTR and untagged CFTR mRNAs produced detectable protein, with more detectable protein from the untagged mRNA. CFTR expression was observed to be higher at 24 h vs. 48 h post transfection.

[00460] The functional expression of CFTR mRNAs was assessed in Fisher Rat Thyroid (FRT) cell lines using TransEpithelial Current Clamp with 24 electrode manifold (TECC24) and 24 well plates with permeable membrane support inserts Transwell ® (Corning). The FRT cells were seeded on the porous membrane and grown until confluence before transfection. Lipofectamine 2000 was used as an optimal reagent for FRT transfection. 24 hours after transfection FRT plates were placed on thermostabilized (36^{O} C) platforms and transepithelial resistance (Rt) values were measured with ~1 min acquisition interval upon robot-assisted or manual transitions of the electrode manifold between cell-populated and reference plates filled with assay buffer. The FRT assay sequence includes baseline reading interval (~8 min), 10 μM Forskolin-induced CFTR activation interval (~10 min), 1 μM VX-770 – induced CFTR potentiation interval (~10 min), and 20 μM INH-172- induced CFTR inhibition interval (~10 min). The FRT conductance traces [Gt = 1/(Rt-50), mS/cm] were reconstructed vs. time. The Forskolin/VX-770-mediated responses were calculated as an Area Under the Gt Curve (Gt AUC) for time points between Forskolin and NIH-172 addition. This AUC calculation was performed after baseline subtraction calculated as a slop line between the initial baseline Gt timepoint and a plateau Gt timepoint acquired after apparent complete CFTR inhibition (**FIG. 11A, dashed line**). The changes in the Gt/min corresponding to CFTR-mediated alteration

of Cl- flux across FRT layer extrapolated as the CFTR functional activity were statistically compared for different mRNA designs. As example in **FIG. 11**, the parental FRT cells were tested with the TECC24 conductance assay after treatment with two concentrations of CFTR mRNA, with 1 µg of mRNA resulting in the Forskolin-induced conductance of ~28 mS/cm^2 and 0.5 µg of mRNA resulting in the conductance of ~7 mS/cm^2 demonstrating the ReCode mRNA-mediated functional expression of CFTR protein. **FIG. 11B** and **FIG. 11C** show an example experiment of dose-dependent CFTR function observed with optimized CFTR mRNA in FRT cells. **FIG 11B** shows a representative conductance kinetic traces in FRT monolayer after addition of CFTR modulators. 5 day-old confluent FRT cells grown on TransWell ® permeable support were transfected with optimized mRNAs using Lipofectamine 2000. MTECC24 assay of the transepithelial conductance was performed 1 day after transfection with one dose of optimized CFTR mRNA. **FIG. 11C** shows CFTR activity by forskolin induced and INH 172-supressed Cl- conductance: mRNA dose dependent transepithelial conductance (Gt) responses: bars are Gt area under the curve (AUC) per min between forskolin addition and Inhibitor-172 addition time points.

Example 8: Delivery of reporter mRNA into fully differentiated hBE cells

[00461] Expression of TR (Tomato Red) mRNA in fully differentiated hBE cells was analyzed. Briefly, hBE cells were plated in 24 well plate and allowed to proliferate. LNPs comprising Tomato Red mRNA were nebulized onto the 24-well plate by placing the plate into an enclosed chamber and allowing the nebulized LNPs to settle onto the hBE cells. As shown in **FIG. 12**, the untreated wells did not show any signal, whereas the wells treated with nebulized LNPs were observed to comprise cells with Tomato Red expression. Two different delivered doses were tested, with increased expression observed after the 400-µg dose compared to treatment with 200-µg dose.

Example 9. Transepithelial Resistance and Equivalent transepithelial current assay in hBE cells to validate cellular tolerability and CFTR rescue efficacy of selected pharmacological compounds and CFTR mRNA/LNP formulations of the present application.

[00462] Transepithelial resistance (Rt) of human Bronchial Epithelia (hBE) cells grown on permeable support in 24 well Transwell ® plate and differentiated/polarized against apical air-liquid interface (ALI) was used to assess the tolerability of human bronchial epithelia to mRNA/LNP formulations of the present application. The reduction of baseline Rt > 50% is considered significant, corresponding to obstruction of the barrier epithelial function. As shown in Fig. 13A, LNPs encapsulated Tomato Red reporter mRNA and CFTR mRNA induces no significant loss of Rt vs. vehicle, suggesting good tolerability of human bronchial epithelia to the ReCode mRNA/LNP formulations. As expected, hBE treated with TR reporter mRNA LNPs showed no detectable Forskolin induced current (**FIG. 13B**) despite strong TR protein expression (red fluorescent signal **FIG. 13C**). The results revealed that TR reporter LNPs can be used as positive transfection control along with CFTR therapeutics used as a positive control of CFTR activity in hBE.

Example 10. Compensation of F508del/F508del CFTR mutation by CFTR mRNA LNP formulations of the present application in differentiated primary hBE cells from a F508del/F508del subject.

[00463] ReCode CFTR mRNA-encapsulated LNPs showed significant rescue of CFTR in the F508del/F508del CFTR hBE model. Briefly, CFTR mRNA was encapsulated with different LNP compositions and delivered to F508del/F508del hBE cells as apical liquid bolus or apical exposure of ALI hBE to nebulized LNPs aerosol. The hBE cell isolated from a cystic fibrosis patient with F508del/F508del genotype at passage 3 were seeded on 24 wells Transwell ® plates and airlifted after 96 hours. Cells grown following a 3 days/week feeding routine with Vertex ALI media. After 5 weeks hBE cell culture were consideredfully differentiated, polarized, and to be ready for the TECC24 functional assay. 4 days prior to treatment mucus was washed from the apical side of the hBE culture with 3 mM DTT in PBS. 24 hours before treatment cells were washed with PBS, additionally washed with PBS on the treatment day, treated apically with liquid bolus or VitroCell nebulized formulations, and tested after 24 or 24+n24 hour CO2 incubation as planned. The TECC 24 assay was performed similarly as described for FRT cells, except for different pharmacological agents needed to suppress large Na^+ conductance and record small isolate CFTR mediated Cl^- current typical for hBE. Specifically, the hBE assay sequence includes background current/ resistance recording interval (~8 min), baseline Cl^- current recording interval (~8 min) after inhibition of Na^+ conductance with 6 µM Benzamil, 10 µM Forskolin + 1 µM VX-770 induced CFTR activation interval (~15 min), and 20 µM Bumetanide induced Cl - current inhibition interval (~10 min). The hBE transepithelial equivalent current traces [Ieq = Vt/(Rt-50), µA/cm] were reconstructed vs. time. The Forskolin/VX-770-induced Cl^-current responses were calculated as an Area Under the Ieq Curve (Ieq AUC) for time points between Forskolin/VX770 and NIH-172 addition. The Ieq AUC/min values were statistically validated and compared across experimental samples. As shown in **FIG. 14A**, the apical bolus treatment with the LNPs comprising HA-CFTR mRNA 5A2-SC8 and DOTAP formulation recovered forskolin-dependent Cl^- current. in F508del/F508del hBEs, suggestion rescue of CFTR function equal or better in comparison to a positive pharmacological control (VX-809). The treatment of F508del/F508del hBEs with apical aerosol of VitroCell-nebulized also recovered CFTR activity.

Example 11. Restoration of CFTR function in CFTR mutant hBEs on with CFTR mRNA formulation of the present application

[00464] ReCode CFTR mRNA encapsulated LNPs were validated to rescue of CFTR function in CFTR mutant (such as R553X/F508del) hBE. Briefly, CFTR mRNA was assembled with different LNP compositions and delivered to CFTR mutant (such as R553X/F508del) hBE cells following similar protocol as described for F508del/F508del hBE above. As shown, in **FIG. 15A**, delivery of the HA-CFTR mRNA via nebulization using LNPs (comprising 4A3-SC7 and DODAP) yielded expression of CFTR in the cells. As shown in **FIG. 15B**, delivery of the composition comprising CFTR mRNA via nebulization was also shown to be effective at rescuing function.

Example 12. Compensation of G542X/F508del CFTR mutation by CFTR mRNA LNP formulations of the present application in differentiated primary hBE cells from a G542X/F508del subject

[00465] ReCode CFTR mRNA encapsulated LNPs showed significant rescue of CFTR in the G542X /F508del CFTR hBE model. Briefly, CFTR mRNA was encapsulated with different LNP compositions and

delivered to G542X/F508del hBE cells as apical liquid bolus or apical exposure of ALI hBE to nebulized LNPs aerosol. The hBE cells isolated from a cystic fibrosis patient with G542X /F508del genotype at passage 3 were seeded on 24 wells Transwell ® plates and airlifted after 96 hours. Cells grown following a 3 days/week feeding routine with Vertex ALI media. After 5 weeks hBE cell culture were considered as fully differentiated, polarized, and to be ready for the TECC24 functional assay. 4 days prior to treatment mucus was washed from the apical side of the hBE culture with 3 mM DTT in PBS. 24 hours before treatment cells were washed with PBS, additionally washed with PBS on the treatment day, treated apically with liquid bolus or VitroCell nebulized formulations, and tested after 24 or 24+n24 12-, 24-, 48-, 72-, or 96-hour CO2 incubation post-treatment time points . Specifically, Tthe hBE assay sequence includes background (initial) current/ resistance recording interval (~8 min), baseline Cl^- current recording interval (~8 min after inhibition of Na^+ conductance with 6 µM Benzamil), 10 µM Forskolin + 1 µM VX-770 induced CFTR activation interval (~15 min), and 20 µM Bumetanide induced Cl - current inhibition interval (~10 min). The hBE transepithelial equivalent current traces [Ieq = Vt/(Rt-50), µA/cm] were reconstructed vs. time. The Forskolin/VX-770-induced Cl^-current responses were calculated as an Area Under the Ieq Curve (Ieq AUC) for time points between Forskolin/VX-770 and NIH-172 addition. The Ieq AUC/min values were statistically validated and compared across experimental samples. As shown in Figure 16A, the apical bolus treatment with the LNPs comprising CFTR mRNA recovered forskolin-dependent Cl^- current in G542X/F508del hBEs, suggesting rescue of CFTR function equal or better in comparison to a positive pharmacological control (VX-809 and Trikafta). Also shown in **FIG. 16A** are representative traces of Forskolin-induced and Bumetanide-suppressed Cl- current in ΔF508-HBE grown and ALI on Transwell ® permeable support. The MTECC24 assay was performed 1 day posttreatment with single dose aerosolized CFTR mRNA LNP, VX-809 and Trikafta equivalent. **FIG. 16B** shows relative CFTR functional activity shown as equivalent current AUC/min. **FIG. 16C** shows initial transepithelial resistance as a relative reference of treatment toxicity. **FIG. 16D** shows a protein blot for CFTR and demonstrates that CFTR protein is successfully produced in the CFTR mRNA treated cells is produced in similar or larger amounts than the positive pharmacological controls, cells treated with Lumacaftor (3 µM VX-809) or Trikafta (3.3 µM VX-661 + 3µM VX-445 and 1 µM VX-770 added with Forskolin).

[00466] In a similar assay, the CFTR mRNA LNP formulations were repeatedly administered based on a twice a week dosing schedule. Using a similar protocol to determine the CFTR function, the repeated administrations showed CFTR function after each dose. **FIG. 17A** and **17B** show that each dose was able to generate improved CFTR function over a negative control.

Example 13. mRNA delivery to specific lung cells.

[00467] Expression of TR (Tomato Red) mRNA in different cell types in hBE cultures (human bronchial epithelial cultures) was analyzed. TR mRNA was loaded in LNPs and delivered into well-differentiated human bronchial epithelial cultures using apical bolus dosing (upper panel) or aerosol delivery (bottom panel). TR protein expression in various cell-types was observed and the percent of TR positive cells in different cell-types was plotted. As shown in the top panel of **FIG. 18**, TR expression was observed strongly

in goblet cells, with less expression in basal and club cells, and minimal expression in ciliated cells. FIG. 18 also shows that well-differentiated human hBE cells were treated once with RTX0001 formulated Td Tomato mRNA (4 mg) using Vitrocell nebulization. % positive cells were determined by co-localization with the indicated markers. This demonstrates that the LNP formulations can be selectively delivered to specific selected lung cell types for delivery.

Example 14. Detection of CFTR mRNA delivery to a subject.

[00468] A subject having or suspected of having cystic fibrosis is given a treatment by administering a composition as described elsewhere herein. The subject is monitored at regular intervals for expression of CFTR in the lungs. A sample of lung tissue from the subject is taken comprising ciliated cells of the lung. The cells are harvested and prepared for RNA isolation. cDNA is produced from the RNA using a first strand synthesis kit and random hexamer. qPCR reactions are run using a set of forward and reverse primers and a fluorescent probe, specific to CFTR and a second set specific to a control or housekeeping gene for expression normalization. Expression of CFTR is detected using a fluorescent readout corresponding the CFTR probe.

Example 15. Clinical Trials on Human Subjects

[00469] ReCode CFTR mRNA encapsulated LNPs are administered to human subjects having cystic fibrosis for maximum tolerated dose, dose limiting toxicity, and safety through single ascending dose (SAD) and multiple ascending dose (MAD) studies. The SAD and MAD studies are tested on human subjects over a 3-month period. Additionally, long term extension study of tolerability and toxicity is tested in subjects in open-label extension studies (OLE). The OLE studies are tested on human subjects over 9 months. In the SAD, MAD, and OLE studies, cohorts of human subjects are tested with either a low dosage of the ReCode CFTR mRNA encapsulated LNPs, high dosage of the ReCode CFTR mRNA encapsulated LNPs, or a placebo. Further, the absolute change in percent of FEV1 of the human subjects are compared before the administration of the ReCode CFTR mRNA encapsulated LNPs and after treatment to evaluate therapeutic efficacy. An overview of the study is shown in **FIG. 19**.

[00470] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the invention be limited by the specific examples provided within the specification. While the invention has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. Furthermore, it shall be understood that all aspects of the invention are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is therefore contemplated that the invention shall also cover any such alternatives, modifications, variations or equivalents. It is intended that

the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A synthetic polynucleotide encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said synthetic polynucleotide comprises one or more nucleoside analogue(s).

2. The synthetic polynucleotide of claim 1, wherein said synthetic polynucleotide comprises 1-methylpseudouridine.

3. A synthetic polynucleotide encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said synthetic polynucleotide comprises a nucleic acid sequence (e.g., an open reading frame (ORF) sequence) having at least about 70%,75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 bases of a sequence selected from SEQ ID NOs: 1-4 and 23.

4. The synthetic polynucleotide of claim 1, wherein said nucleic acid sequence comprises fewer than about 115, 110, 105, 100, 95, or 90 UU or TT dinucleotide

5. The synthetic polynucleotide of claim 1, wherein said nucleic acid sequence comprises at least two synonymous codons encoding arginine.

6. The synthetic polynucleotide of claim 1, wherein said nucleic acid sequence comprises at least three synonymous codons encoding arginine.

7. The synthetic polynucleotide of claim 1, wherein no more than about 70%, 65%, 60%, 55%, or 50% of all arginine encoding codons of said nucleic acid sequence is AGA codon.

8. The synthetic polynucleotide of claim 1, wherein said nucleic acid sequence encodes a polypeptide that comprises an amino acid sequence having at least 70%, 75%, 80%,81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity over at least 100, 300, 500, 700, 900, or 1,000 contiguous amino acid residues to SEQ ID NO: 5.

9. The synthetic polynucleotide of claim 1, wherein said synthetic polynucleotide is a messenger ribonucleic acid (mRNA)

10. The synthetic polynucleotide of claim 1, wherein said synthetic polynucleotide further comprises a 3'- or 5'-noncoding region.

11. The synthetic polynucleotide of claim 10, wherein said 3'- or 5'-noncoding region enhances an expression or activity of said CFTR protein encoded by said synthetic polynucleotide within a cell.

12. The synthetic polynucleotide of claim 1, wherein said synthetic polynucleotide further comprises a 5' cap structure.

13. The synthetic polynucleotide of claim 1, wherein said 3' noncoding region comprises a poly adenosine tail.

14. The synthetic polynucleotide of claim 13, wherein said poly adenosine tail comprises at most 200 adenosines.

15. The synthetic polynucleotide of claim 13, wherein said poly adenosine tail improves a pharmacokinetic characteristic of said synthetic polynucleotide in a cell.

16. The synthetic polynucleotide of claim 15, wherein said poly adenosine tail improves a prolonged half-life of said synthetic polynucleotide in a cell.

17. A pharmaceutical composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid.

18. The pharmaceutical composition of claim 17, wherein said lipid composition comprises said ionizable cationic lipid at a molar percentage of about 5% to about 30%

19. The pharmaceutical composition of claim 17, wherein a mass or weight ratio of said ionizable cationic lipid to said synthetic polynucleotide is of no more than about 50:1, 40:1, 30:1, 20:1, 15:1 or 10:1

20. The pharmaceutical composition of claim 17, wherein said SORT lipid is a permanently cationic lipid

21. The pharmaceutical composition of claim 17, wherein said SORT lipid is a second ionizable cationic lipid

22. The pharmaceutical composition of claim 21, wherein said lipid composition comprises said SORT lipid at a molar percentage ofabout 5% to about 65%

23. The pharmaceutical composition of claim 21, wherein said lipid composition comprises said SORT lipid at a molar percentage of about 5% to about 30%

24. The pharmaceutical composition of claim 17, wherein said lipid composition further comprises a zwitterionic lipid (e.g., a phospholipid)

25. The pharmaceutical composition of claim 24, wherein said lipid composition comprises said zwitterionic lipid at a molar percentage of about 5% to about 25%

26. The pharmaceutical composition of claim 24, wherein a molar ratio of said zwitterionic lipid to said synthetic polynucleotide is of no more than about 50:1, 40:1, 30:1, or 20:1

27. The pharmaceutical composition of claim 17, wherein said lipid composition further comprises a steroid or steroid derivative

28. The pharmaceutical composition of claim 27, wherein said lipid composition comprises said steroid or steroid derivative at a molar percentage of about 15% to about 46%

29. The pharmaceutical composition of claim 17, wherein said lipid composition further comprises a polymer-conjugated lipid (e.g., poly(ethylene glycol) (PEG)-conjugated lipid).

30. The pharmaceutical composition of claim 29, wherein said lipid composition comprises said polymer-conjugated lipid at a molar percentage of about 0.5% to about 10%

31. The pharmaceutical composition of claim 17, wherein a molar ratio of nitrogen in said lipid composition to phosphate in said synthetic polynucleotide (N/P ratio) is of no more than about 50:1, 40:1, 30:1, or 20:1

32. The pharmaceutical composition of claim 31, wherein said N/P ratio is from about 5:1 to about 30:1.

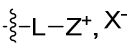
33. The pharmaceutical composition of claim 17, wherein a mass or weight ratio of said synthetic polynucleotide to total lipids of said lipid composition is no more than about 1:20, 1:50, or 1:100.

34. The pharmaceutical composition of claim 17, wherein said SORT lipid comprises a permanently positively charged moiety (e.g., a quaternary ammonium ion).

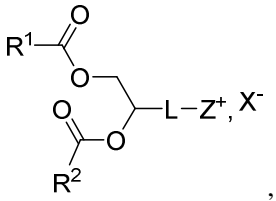
35. The pharmaceutical composition of claim 34, wherein said SORT lipid comprises a counterion.

36. The pharmaceutical composition of claim 17, wherein said SORT lipid is a phosphocholine lipid (e.g., saturated or unsaturated).

37. The pharmaceutical composition of claim 36, wherein said SORT lipid is an ethylphosphocholine.

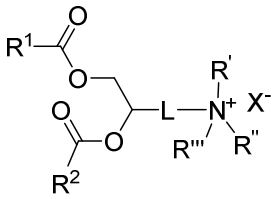
38. The pharmaceutical composition of claim 17, wherein said SORT lipid comprises a headgroup having a structural formula: , wherein L is a (e.g., biodegradable) linker; Z^+ is positively charged moiety (e.g., a quaternary ammonium ion); and X^- is a counterion.

39. The pharmaceutical composition of claim 38, wherein said SORT lipid has a structural formula:

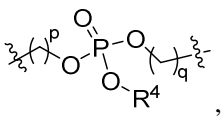


, wherein R^1 and R^2 are each independently an optionally substituted C_6-C_{24} alkyl, or an optionally substituted C_6-C_{24} alkenyl.

40. The pharmaceutical composition of claim 38, wherein said SORT lipid has a structural formula:



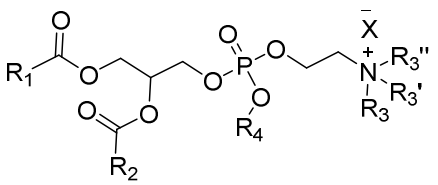
.

41. The pharmaceutical composition of claim 40, wherein L is , wherein:

p and q are each independently 1, 2, or 3; and

R^4 is an optionally substituted C_1-C_6 alkyl.

42. The pharmaceutical composition of claim 38, wherein said SORT lipid has a structural formula:



(IA), wherein:

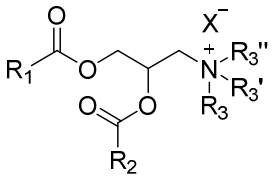
R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$;

R_4 is $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$; and

X^- is a monovalent anion.

43. The pharmaceutical composition of claim 17, wherein said SORT lipid has a structural formula:



(S-I')

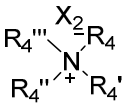
wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$;

X^- is a monovalent anion.

44. The pharmaceutical composition of claim 17, wherein said SORT lipid has a structural formula:



(S-II')

wherein:

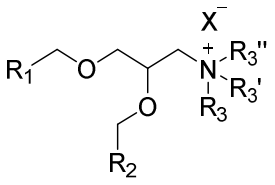
R_4 and R_4' are each independently $alkyl_{(C6\text{-}C24)}$, $alkenyl_{(C6\text{-}C24)}$, or a substituted version of either group;

R_4'' is $alkyl_{(C\leq24)}$, $alkenyl_{(C\leq24)}$, or a substituted version of either group;

R_4''' is $alkyl_{(C1\text{-}C8)}$, $alkenyl_{(C2\text{-}C8)}$, or a substituted version of either group; and

X_2 is a monovalent anion.

45. The pharmaceutical composition of claim 17, wherein said SORT lipid has a structural formula:



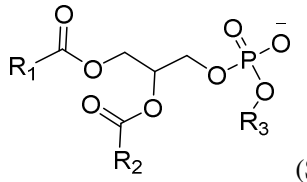
(S-III), wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3, R_3', and R_3'' are each independently $alkyl_{(C\leq6)}$ or substituted $alkyl_{(C\leq6)}$; and

X^- is a monovalent anion.

46. The pharmaceutical composition of claim 17, wherein said SORT lipid has a structural formula:



(S-IV), wherein:

R_1 and R_2 are each independently $alkyl_{(C8\text{-}C24)}$, $alkenyl_{(C8\text{-}C24)}$, or a substituted version of either group;

R_3 is hydrogen, $alkyl_{(C\leq6)}$, or substituted $alkyl_{(C\leq6)}$, or $-Y_1-R_4$, wherein:

Y_1 is $alkanediyl_{(C\leq6)}$ or substituted $alkanediyl_{(C\leq6)}$; and

R_4 is $acyloxy_{(C\leq8\text{-}24)}$ or substituted $acyloxy_{(C\leq8\text{-}24)}$.

47. The pharmaceutical composition of claim 17, wherein said pharmaceutical composition is an aerosol composition.

48. The pharmaceutical composition of claim 45, wherein said aerosol composition has a droplet size from 0.5 micron (μm) to 10 μm.

49. The pharmaceutical composition of claim 45, wherein said aerosol composition has a median droplet size from 0.5 μm to 10 μm.

50. The pharmaceutical composition of claim 45, wherein said aerosol composition has an average droplet size from 0.5 μm to 10 μm.

51. The pharmaceutical composition of claim 17, wherein said pharmaceutical composition is formulated for aerosol administration

52. The pharmaceutical composition of claim 17, wherein said pharmaceutical composition is formulated for apical delivery.

53. The pharmaceutical composition of claim 17, wherein said pharmaceutical composition is formulated for nebulization.

54. A method for enhancing an expression or activity of cystic fibrosis transmembrane conductance regulator (CFTR) protein in a cell, the method comprising:

contacting said cell with a composition comprising a synthetic polynucleotide assembled with a lipid composition, wherein said synthetic polynucleotide encodes a CFTR protein; and wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after contacting, optionally wherein said therapeutically effective activity of said functional variant of CFTR protein is determined by measuring a change in a transepithelial ion transport characteristic of a plurality of cells comprising said cell as compared to that of a reference plurality of cells in absence of said contacting.

55. The method of claim 54, wherein said contacting is repeated

56. The method of claim 55, wherein said contacting is at least once a week

57. The method of claim 55, wherein said contacting is at least twice a week

58. The method of claim 54, wherein the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after each contacting

59. The method of claim 54, wherein said contacting is a first contacting, and wherein the method comprises a second contacting, optionally, performed at least 1, 2, or 3 day(s) after said first contacting

60. The method of claim 59, further comprising a third contacting, optionally wherein said third contacting is performed at least 1, 2, or 3 day(s) after said second contacting

61. The method of claim 59, wherein the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after said second contacting

62. The method of claim 60, wherein the method yields a therapeutically effective amount or activity of a functional variant of CFTR protein in said cell at least 24 hours after said third contacting

63. The method of claim 54, wherein said contacting comprises administering to a subject said composition comprising said synthetic polynucleotide assembled with said lipid composition

64. The method of claim 63, wherein said subject is a mammal.

65. The method of claim 63, wherein said subject is a human.

66. The method of claim 63, wherein said administering comprises inhalation by nebulization.

67. The method of claim 54, wherein said composition in each contacting is identical

68. The method of claim 54, wherein said cell is a lung airway cell.

69. The method of claim 68, wherein said cell is a lung secretory cell

70. The method of claim 68, wherein said cell is a bronchial epithelial cell

71. The method of claim 54, wherein said cell is undifferentiated.

72. The method of claim 54, wherein said cell is differentiated.

73. The method of claim 54, wherein said cell is derived from said subject.

74. The method of claim 54, wherein said contacting is *in vivo*.

75. The method of claim 54, wherein said contacting is *in vitro*.

76. The method of claim 54-73, wherein said contacting is *ex vivo*.

77. The method of claim 54, wherein said functional variant of CFTR protein is a wild-type CFTR protein

78. The method of claim 54-, wherein said functional variant of CFTR protein is a full-length CFTR protein

79. The method of claim 54-, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least about 5 micro-Ampere (μA), e.g., as determined in an *in vitro* assay.

80. The method of claim 79, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from about 5 micro-Ampere (μA) to about 30 μA.

81. The method of claim 54, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current of at least about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$), e.g., as determined in an *in vitro* assay.

82. The method of claim 81, wherein said therapeutically effective activity of said functional variant of CFTR protein corresponds to a transepithelial current from about 2 micro-Ampere (μA) per squared centimeter per minute ($\mu A \cdot cm^{-2} \cdot min^{-1}$) to about 20 $\mu A \cdot cm^{-2} \cdot min^{-1}$.

83. The method of claim 54, wherein the method increases an amount or activity of said functional variant of CFTR protein in said cell (e.g., by at least about 1.1-fold) relative to a corresponding control (e.g., that of a corresponding cell absent said contacting).

84. The method of claim 54, wherein the method enhances (e.g., chloride) ion transport in said cell (e.g., by at least about 1.1-fold) relative to a corresponding control (e.g., that of a corresponding cell absent said contacting).

85. The method of claim 54, wherein said subject exhibits or is determined to exhibit a mutation in a cystic fibrosis transmembrane conductance regulator (CFTR) gene.

86. The method of claim 85, wherein said mutation is a loss-of-function mutation.

87. The method of claim 85 or, wherein said mutation is a nonsense or frameshift mutation.

88. The method of claim 85, wherein said mutation is in one or more of exons 11-27 of CFTR gene.

89. The method of claim 85, wherein said mutation is R553X, G542X, F508del, or R1162X, or a combination thereof.

90. A method for lung secretory cell or lung basal cell delivery, comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a therapeutically effective amount or activity of said synthetic polynucleotide in a lung secretory cell of said subject, optionally wherein said therapeutically effective activity of said synthetic polynucleotide is determined by measuring a change in a transepithelial ion transport characteristic of a lung comprising said lung secretory cell or lung basal cell as compared to that of a reference lung in absence of said contacting.

91. A method for lung secretory cell or lung basal cell delivery, comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a greater therapeutic amount or activity of said synthetic polynucleotide in a lung secretory cell or lung basal cell of said subject as compared to that in a lung non-secretory cell or non basal cell of said subject.

92. The method of claim 91, wherein the method is characterized in that: (i) at least about 50%, 55%, or 60% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected in lung secretory cells, lung basal cells, or a combination thereof; or (ii) no more than about 50%, 45%, or 40% of (e.g., pulmonary) expression of said synthetic polynucleotide is detected in lung non-secretory cells, lung non-basal cells, or a combination thereof.

93. The method of claim 91, wherein said lung non-secretory cell is a lung ciliated cell

94. The method of claim 91, wherein said lung non-secretory cell is a lung basal cell

95. A method for lung secretory cell delivery, comprising administering to a subject a composition comprising a synthetic polynucleotide assembled with a lipid composition, which synthetic polynucleotide encodes a cystic fibrosis transmembrane conductance regulator (CFTR) protein, wherein said lipid composition comprises:

an ionizable cationic lipid; and

a selective organ targeting (SORT) lipid separate from said ionizable cationic lipid,

thereby yielding a therapeutic amount or activity of said synthetic polynucleotide in at least 5% of lung secretory cells of said subject.

96. The method of claim 95, wherein said administering comprises administering to a lung of said subject said composition comprising said synthetic polynucleotide assembled with said lipid composition.

97. The method of claim 95, wherein said lung secretory cell is a club cell or a goblet cell.

98. A method for treating a subject having or suspected of having a cystic fibrosis transmembrane conductance regulator (CFTR)-associated condition, the method comprising administering to said subject a pharmaceutical composition of any one of claim 17.

99. The method of claim 98, wherein said CFTR-associated condition is cystic fibrosis, hereditary emphysema, or chronic obstructive pulmonary disease (COPD).

100. The method of claim 98, wherein said administering comprises local administration

101. The method of claim 98, wherein said administering comprises nebulization