

CIRCULAR RNA COMPOSITIONS AND METHODS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of, and priority to, U.S. Provisional Application No. 62/992,518, filed on March 20, 2020, the contents of which are hereby incorporated by reference in entirety for all purposes.

BACKGROUND

[0002] Conventional gene therapy involves the use of DNA for insertion of desired genetic information into host cells. The DNA introduced into the cell is usually integrated to a certain extent into the genome of one or more transfected cells, allowing for long-lasting action of the introduced genetic material in the host. While there may be substantial benefits to such sustained action, integration of exogenous DNA into a host genome may also have many deleterious effects. For example, it is possible that the introduced DNA will be inserted into an intact gene, resulting in a mutation which impedes or even totally eliminates the function of the endogenous gene. Thus, gene therapy with DNA may result in the impairment of a vital genetic function in the treated host, such as *e.g.*, elimination or deleteriously reduced production of an essential enzyme or interruption of a gene critical for the regulation of cell growth, resulting in unregulated or cancerous cell proliferation. In addition, with conventional DNA based gene therapy it is necessary for effective expression of the desired gene product to include a strong promoter sequence, which again may lead to undesirable changes in the regulation of normal gene expression in the cell. It is also possible that the DNA based genetic material will result in the induction of undesired anti-DNA antibodies, which in turn, may trigger a possibly fatal immune response. Gene therapy approaches using viral vectors can also result in an adverse immune response. In some circumstances, the viral vector may even integrate into the host genome. In addition, production of clinical grade viral vectors is also expensive and time consuming. Targeting delivery of the introduced genetic material using viral vectors can also be difficult to control. Thus, while DNA based gene therapy has been evaluated for delivery of secreted proteins using viral vectors (U.S. Patent No. 6,066,626; U.S. Publication No. US2004/0110709), these approaches may be limited for these various reasons.

[0003] In contrast to DNA, the use of RNA as a gene therapy agent is substantially safer because RNA does not involve the risk of being stably integrated into the genome of the

transfected cell, thus eliminating the concern that the introduced genetic material will disrupt the normal functioning of an essential gene, or cause a mutation that results in deleterious or oncogenic effects, and extraneous promoter sequences are not required for effective translation of the encoded protein, again avoiding possible deleterious side effects. In addition, it is not necessary for mRNA to enter the nucleus to perform its function, while DNA must overcome this major barrier.

[0004] Circular RNA is useful in the design and production of stable forms of RNA. The circularization of an RNA molecule provides an advantage to the study of RNA structure and function, especially in the case of molecules that are prone to folding in an inactive conformation (Wang and Ruffner, 1998). Circular RNA can also be particularly interesting and useful for *in vivo* applications, especially in the research area of RNA-based control of gene expression and therapeutics, including protein replacement therapy and vaccination.

[0005] Prior to this invention, there were three main techniques for making circularized RNA *in vitro*: the splint-mediated method, the permuted intron-exon method, and the RNA ligase-mediated method. However, the existing methodologies are limited by the size of RNA that can be circularized, thus limiting their therapeutic application.

SUMMARY

[0006] Circular RNA, along with related compositions and methods are described herein. In some embodiments, the inventive circular RNA comprises post splicing group I intron fragments, spacers, an IRES, optional duplex forming regions, and more than one expression sequence. In some embodiments, the inventive circular RNA comprises duplex forming regions. In some embodiments, the expression sequences are separated by one or more polynucleotide sequences encoding a cleavage site. In some embodiments, a cleavage site is a self-cleaving peptide. In some embodiments, the self-cleaving peptide is a 2A self-cleaving peptide. In some embodiments, the first and second expression sequences are separated by ribosomal skipping element. In some embodiments, each expression sequence encodes a therapeutic protein. In some embodiments, the first expression sequence encodes a cytokine or a functional fragment thereof. In some embodiments, the first expression sequence encodes a transcription factor. In some embodiments, the first expression encodes an immune checkpoint inhibitor. In some embodiments, the first expression sequence encodes for a chimeric antigen receptor. In some embodiments, the first expression sequence encodes a first T-cell receptor (TCR) chain, and the second expression encodes a second TCR chain. In some embodiments, circular RNA of the invention has improved expression, functional

stability, ease of manufacturing, and/or half-life when compared to linear RNA. In some embodiments, circular RNA of the invention has reduced immunogenicity. In some embodiments, inventive methods and constructs result in improved circularization efficiency, splicing efficiency, and/or purity when compared to existing RNA circularization approaches.

[0007] In some embodiments, the circular RNA polynucleotide comprises one or more microRNA binding sites. The microRNA binding site is recognized by a microRNA expressed in the liver. In some embodiments, the microRNA binding site is recognized by miR-122.

[0008] One aspect of the present application provides a circular RNA polynucleotide comprising, in the following order, a post splicing 3' group I intron fragment, an Internal Ribosome Entry Site (IRES), a first expression sequence, a second expression sequence, and a post splicing 5' group I intron fragment.

[0009] In some embodiments, the circular RNA polynucleotide comprises a polynucleotide sequence encoding a cleavage site between the first expression sequence and the second expression sequence. In some embodiments, the cleavage site is a self-cleaving spacer. In some embodiments, the self-cleaving spacer is a 2A self-cleaving peptide.

[0010] In some embodiments, the circular RNA polynucleotide comprises a second IRES between the first expression sequence and the second expression sequence. In some embodiments, the first IRES consists of or comprises a sequence according to any of SEQ ID NO: 1-72. In some embodiments, the second IRES consists of or comprises a sequence according to any of SEQ ID NO: 1-72.

[0011] In some embodiments, the first expression sequence encodes a first therapeutic protein, and the second expression sequence encodes a second therapeutic protein. In some embodiments, the first expression sequence or the second expression sequence encodes an antibody. In some embodiments, the first expression sequence or the second expression sequence encodes a chimeric antigen receptor. In some embodiments, the first expression sequence or the second expression sequence encodes a transcription factor. In some embodiments, wherein the first expression sequence or the second expression sequence encodes a cytokine. In some embodiments, wherein the first expression sequence or the second expression sequence encodes an immune inhibitory molecule. In some embodiments, the first expression sequence or the second expression sequence encodes an agonist of a costimulatory molecule. In some embodiments, wherein the first expression sequence or the second expression sequence encodes an inhibitor of an immune checkpoint molecule. In some embodiments, wherein the first expression sequence encodes the alpha chain of a T cell

receptor (TCR) and the second expression sequence encodes the beta chain of a T cell receptor (TCR).

[0012] In some embodiments, the first expression sequence encodes the beta chain of a T cell receptor (TCR) and the second expression sequence encodes the alpha chain of a T cell receptor (TCR). In some embodiments, the first expression sequence encodes the gamma chain of a T cell receptor (TCR) and the second expression sequence encodes the delta chain of a T cell receptor (TCR). In some embodiments, the first expression sequence encodes the delta chain of a T cell receptor (TCR) and the second expression sequence encodes the gamma chain of a T cell receptor (TCR). In some embodiments, the first expression sequence encodes a T cell receptor (TCR) and the second expression sequence encodes a chemokine. In some embodiments, the first expression sequence encodes for a chemokine and the second expression sequence encodes for a T cell receptor (TCR). In some embodiments, the first expression sequence encodes a chimeric antigen receptor (CAR) and the second expression sequence encodes a PD1 or PDL1 antagonist. In some embodiments, the first expression sequence encodes a PD1 or PDL1 antagonist and the second expression sequence encodes a chimeric antigen receptor (CAR). In some embodiments, the first expression sequence encodes for a chimeric antigen receptor (CAR), and the second expression sequence encodes a chemokine. In some embodiments, the first expression sequence encodes for a chemokine, and the second expression sequence encodes for a chimeric antigen receptor (CAR). In some embodiments, the first expression sequence encodes a transcription factor and the second expression sequence encodes a cytokine.

[0013] In some embodiments, the first expression sequence encodes a T cell receptor (TCR) and the second expression sequence encodes a cytokine. In some embodiments, the first expression sequence encodes a cytokine and the second expression sequence encodes a T cell receptor (TCR). In some embodiments, the cytokine is selected from IL-2, IL-7, IL-12, and IL-15.

[0014] In some embodiments, the first expression sequence encodes for a T cell receptor (TCR) and the second expression sequence encodes for a transcription factor. In some embodiments, the first expression sequence encodes for a transcription factor and the second expression sequence encodes for a T cell receptor (TCR). In some embodiments, the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt, and cd25.

[0015] In some embodiments, the first expression sequence encodes a chimeric antigen receptor (CAR) and the second expression sequence encodes a cytokine. In some

embodiments, the first expression sequence encodes a cytokine and the second expression sequence encodes a chimeric antigen receptor (CAR). In some embodiments, the cytokine is selected from IL-2, IL-7, IL-12, and IL-15.

[0016] In some embodiments, the first expression sequence encodes a cytokine and the second expression sequence encodes a transcription factor. In some embodiments, the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, ROR γ t, and cd25. In some embodiments, the cytokine is selected from IL-10, IL-12, and TGF β .

[0017] In some embodiments, the first expression sequence encodes a transcription factor and the second expression sequence encodes a chemokine. In some embodiments, the first expression sequence encodes a chemokine and the second expression sequence encodes a transcription factor. In some embodiments, the transcription factor is selected from FOXP3, STAT5B, and HELIOS. In some embodiments, the chemokine is a CC chemokine, CXC chemokine, C chemokine, or a CX3C chemokine. In some embodiments, the chemokine is selected from CCL1, CCL2, CCL3, CCL4, CCL5, CCL6, CCL7, CCL8, CCL9/CCL10, CCL11, CCL12, CCL13, CCL14, CCL15, CCL16, CCL17, CCL18, CCL19, CCL20, CCL21, CCL22, CCL23, CCL24, CCL25, CCL26, CCL27, CCL28, CXCL 1, CXCL2, CXCL3, CXCL4, CXCL5, CXCL6, CXCL7, CXCL8, CXCL9, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL15, CXCL16, CXCL17, XCL1, XCL2, and CX3CL1.

[0018] In some embodiments, the first expression sequence encodes a tumor antigen and the second expression sequence encodes a cytokine. In some embodiments, the first expression sequence encodes a cytokine and the second expression sequence encodes a tumor antigen. In some embodiments, the antigen is a neoantigen. In some embodiments, the cytokine is IFN γ .

[0019] In some embodiments, the first expression sequence encodes a CAR and the second expression sequence encodes a CAR.

[0020] In some embodiments, the first expression sequence encodes a cytokine and the second expression sequence encodes a cytokine. In some embodiments, the first or second expression sequence encodes a cytokine selected from IL-10, TGF β , and IL-35. In some embodiments, the first or second expression sequence encodes a cytokine selected from IFN γ , IL-2, IL-7, IL-15, and IL-18.

[0021] In some embodiments, the first expression sequence encodes a T cell receptor (TCR) and the second expression sequence encodes a T cell receptor (TCR). In some embodiments, the first expression sequence encodes for a chemokine and the second expression sequence encodes for a chemokine. In some embodiments, the first or second

expression sequence encodes for an immunosuppressive enzyme. In some embodiments, the first expression sequence encodes a rate limiting enzyme and the second expression sequence encodes a flux-limiting enzyme. In some embodiments, the first expression sequence encodes a flux-limiting enzyme and the second expression sequence encodes a rate limiting enzyme.

[0022] In some embodiments, the first expression sequence encodes a transcription factor and the second expression sequence encodes a survival factor. In some embodiments, the first expression sequence encodes a survival factor and the second expression sequence encodes a transcription factor. In some embodiments, the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt, and cd25. In some embodiments, the survival factor is selected from BCL-XL.

[0023] In some embodiments, the first or second expression sequence encodes for a chaperone protein or complex. In some embodiments, the first expression sequence encodes a transcription factor and the second expression sequence encodes a chaperone protein or complex. In some embodiments, the first expression sequence encodes a chaperone protein or complex and the second expression sequence encodes for a transcription factor. In some embodiments, the chaperone protein or complex is selected from Skp, Spy, FkpA, SurA, Hsp60, Hsp70, GroEL, GroES, Hsp90, HtpG, Hsp100, ClpA, ClpX, ClpP, and Hsp104. In some embodiments, the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt, and cd25.

[0024] In some embodiments, one or both expression sequences encode a signaling protein.

[0025] In some embodiments, the first expression sequence encodes for an enzyme and the second expression encodes for the first expression sequence's negative regulatory inhibitor. In some embodiments, the first expression sequence encodes for a negative regulatory inhibitor protein of an enzyme encoded in the second expression sequence. In some embodiments, the negative regulatory inhibitor is selected from a p57kip2, BAX inhibitor, and TIPE2.

[0026] In some embodiments, the first expression sequence encodes a dominant negative protein and the second expression sequence encodes an immune protein. In some embodiments, the first expression sequence encodes an immune protein and the second expression sequence encodes a dominant negative protein. In some embodiments, the first or second expression sequence encodes for an anti-inflammatory protein.

[0027] In some embodiments, the first expression sequence encodes a transcription factor

and the second expression sequence is capable of converting 5-fluorocytosine (5-FC) into 5-fluorouracil (5-FU). In some embodiments, the first expression sequence is capable of converting 5-fluorocytosine (5-FC) into 5-fluorouracil (5-FU) and the second expression sequence is a transcription factor. In some embodiments, wherein the expression sequence capable of converting 5-fluorocytosine (5-FC) into 5-fluorouracil (5-FU) is cytosine deaminase.

[0028] In some embodiments, the circular RNA polynucleotide comprises a first spacer between the 5' duplex forming region and the post splicing 3' group I intron fragment, and a second spacer between the post splicing 5' group I intron fragment and the 3' duplex forming region. In some embodiments, the first and second spacers each have a length of about 10 to about 60 nucleotides. In some embodiments, the first and second duplex forming regions each have a length of about 9 to about 19 nucleotides. In some embodiments, the first and second duplex forming regions each have a length of about 30 nucleotides. In some embodiments, the IRES has a sequence of an IRES from Taura syndrome virus, Triatoma virus, Theiler's encephalomyelitis virus, Simian Virus 40, Solenopsis invicta virus 1, Rhopalosiphum padi virus, Reticuloendotheliosis virus, Human poliovirus 1, Plautia stali intestine virus, Kashmir bee virus, Human rhinovirus 2, Homalodisca coagulata virus- 1, Human Immunodeficiency Virus type 1, Homalodisca coagulata virus- 1, Himetobi P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinitis virus, Ectropis obliqua picorna-like virus, Encephalomyocarditis virus, Drosophila C Virus, Human coxsackievirus B3, Crucifer tobamovirus, Cricket paralysis virus, Bovine viral diarrhea virus 1, Black Queen Cell Virus, Aphid lethal paralysis virus, Avian encephalomyelitis virus, Acute bee paralysis virus, Hibiscus chlorotic ringspot virus, Classical swine fever virus, Human FGF2, Human SFTPA1, Human AML1/RUNX1, Drosophila antennapedia, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 alpha, Human n.myc, Mouse Gtx, Human p27kip1, Human PDGF2/c-sis, Human p53, Human Pim-1, Mouse Rbm3, Drosophila reaper, Canine Scamper, Drosophila Ubx, Human UNR, Mouse UtrA, Human VEGF-A, Human XIAP, Drosophila hairless, S. cerevisiae TFIID, S. cerevisiae YAP1, tobacco etch virus, turnip crinkle virus, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, Picobirnavirus, HCV QC64, Human Cosavirus E/D, Human Cosavirus F, Human Cosavirus JMY, Rhinovirus NAT001, HRV14, HRV89, HRVC-02, HRV-A21, Salivirus A SH1, Salivirus FHB, Salivirus NG-J1, Human Parechovirus 1, Crohivirus B, Yc-3, Rosavirus M-7,

Shanbavirus A, Pasivirus A, Pasivirus A 2, Echovirus E14, Human Parechovirus 5, Aichi Virus, Hepatitis A Virus HA16, Phopivirus, CVA10, Enterovirus C, Enterovirus D, Enterovirus J, Human Pegivirus 2, GBV-C GT110, GBV-C K1737, GBV-C Iowa, Pegivirus A 1220, Pasivirus A 3, Sapelovirus, Rosavirus B, Bakunsa Virus, Tremovirus A, Swine Pasivirus 1, PLV-CHN, Pasivirus A, Sicinivirus, Hepacivirus K, Hepacivirus A, BVDV1, Border Disease Virus, BVDV2, CSFV-PK15C, SF573 Dicistrovirus, Hubei Picorna-like Virus, CRPV, Salivirus A BN5, Salivirus A BN2, Salivirus A 02394, Salivirus A GUT, Salivirus A CH, Salivirus A SZ1, Salivirus FHB, CVB3, CVB1, Echovirus 7, CVB5, EVA71, CVA3, CVA12, EV24 or an aptamer to eIF4G.

[0029] In some embodiments, the circular RNA polynucleotide comprises natural nucleotides. In some embodiments, the circular RNA polynucleotide consists of natural nucleotides. In some embodiments, the expression sequence is codon optimized.

[0030] In some embodiments, the circular RNA polynucleotide is optimized to lack at least one microRNA binding site present in an equivalent pre-optimized polynucleotide. In some embodiments, the circular RNA polynucleotide is optimized to lack at least one endonuclease susceptible site present in an equivalent pre-optimized polynucleotide. In some embodiments, the circular RNA polynucleotide is optimized to lack at least one RNA-editing susceptible site present in an equivalent pre-optimized polynucleotide.

[0031] In some embodiments, the circular RNA polynucleotide of any one of the preceding claims, wherein the circular RNA polynucleotide is from about 100 nucleotides to about 15 kilobases in length.

[0032] In some embodiments, the circular RNA polynucleotide of has an *in vivo* duration of therapeutic effect in humans of at least about 20 hours. In some embodiments, the circular RNA polynucleotide of has a functional half-life of at least about 20 hours. In some embodiments, the circular RNA polynucleotide of has a duration of therapeutic effect in a human cell greater than or equal to that of an equivalent linear RNA polynucleotide comprising the same expression sequence. In some embodiments, the circular RNA polynucleotide of has a functional half-life in a human cell greater than or equal to that of an equivalent linear RNA polynucleotide comprising the same expression sequence. In some embodiments, the circular RNA polynucleotide of has an *in vivo* duration of therapeutic effect in humans greater than that of an equivalent linear RNA polynucleotide having the same expression sequence. In some embodiments, the circular RNA polynucleotide of has an *in vivo* functional half-life in humans greater than that of an equivalent linear RNA polynucleotide having the same expression sequence.

[0033] In another aspect, the present application provides a pharmaceutical composition comprising a circular RNA polynucleotide as described herein, a nanoparticle, and optionally, a targeting moiety operably connected to the nanoparticle. In some embodiments, the nanoparticle is a lipid nanoparticle, a core-shell nanoparticle, a biodegradable nanoparticle, a biodegradable lipid nanoparticle, a polymer nanoparticle, or a biodegradable polymer nanoparticle.

[0034] In some embodiments, the pharmaceutical composition comprises a targeting moiety, wherein the targeting moiety mediates receptor-mediated endocytosis or direct fusion into selected cells of a selected cell population or tissue in the absence of cell isolation or purification. In some embodiments, the pharmaceutical composition comprises a targeting moiety operably connected to the nanoparticle. In some embodiments, the targeting moiety is a scFv, nanobody, peptide, minibody, polynucleotide aptamer, heavy chain variable region, light chain variable region or fragment thereof.

[0035] In some embodiments, the pharmaceutical composition comprises less than 1%, by weight, of the polynucleotides in the composition that are double stranded RNA, DNA splints, or triphosphorylated RNA. In some embodiments, the pharmaceutical composition comprises less than 1%, by weight, of the polynucleotides and proteins in the pharmaceutical composition that are double stranded RNA, DNA splints, triphosphorylated RNA, phosphatase proteins, protein ligases, and capping enzymes.

[0036] In another aspect, the present application provides a method of treating a subject in need thereof comprising administering a therapeutically effective amount of a composition comprising the circular RNA polynucleotide described herein, a nanoparticle, and optionally, a targeting moiety operably connected to the nanoparticle.

[0037] In some embodiments, the composition comprises a targeting moiety, wherein the targeting moiety mediates receptor-mediated endocytosis selectively into cells of a selected cell population in the absence of cell selection or purification. In some embodiments, the targeting moiety is an scFv, nanobody, peptide, minibody, heavy chain variable region, light chain variable region or fragment thereof. In some embodiments, the nanoparticle is a lipid nanoparticle, a core-shell nanoparticle, or a biodegradable nanoparticle.

[0038] In some embodiments, the nanoparticle comprises one or more cationic lipids, ionizable lipids, or poly β -amino esters. In some embodiments, the nanoparticle comprises one or more non-cationic lipids. In some embodiments, the nanoparticle comprises one or more PEG-modified lipids, polyglutamic acid lipids, or hyaluronic acid lipids. In some embodiments, the nanoparticle comprises cholesterol. In some embodiments, the

nanoparticle comprises arachidonic acid or oleic acid. In some embodiments, the nanoparticle comprises more than one circular RNA polynucleotide.

[0039] In some embodiments, the subject has a cancer selected from the group consisting of acute lymphocytic leukemia; acute myeloid leukemia (AML); alveolar rhabdomyosarcoma; B cell malignancies; bladder cancer (e.g., bladder carcinoma); bone cancer; brain cancer (e.g., medulloblastoma); breast cancer; cancer of the anus, anal canal, or anorectum; cancer of the eye; cancer of the intrahepatic bile duct; cancer of the joints; cancer of the neck; gallbladder cancer; cancer of the pleura; cancer of the nose, nasal cavity, or middle ear; cancer of the oral cavity; cancer of the vulva; chronic lymphocytic leukemia; chronic myeloid cancer; colon cancer; esophageal cancer; cervical cancer; fibrosarcoma; gastrointestinal carcinoid tumor; head and neck cancer (e.g., head and neck squamous cell carcinoma); Hodgkin lymphoma; hypopharynx cancer; kidney cancer; larynx cancer; leukemia; liquid tumors; liver cancer; lung cancer (e.g., non-small cell lung carcinoma and lung adenocarcinoma); lymphoma; mesothelioma; mastocytoma; melanoma; multiple myeloma; nasopharynx cancer; non-Hodgkin lymphoma; B-chronic lymphocytic leukemia; hairy cell leukemia; acute lymphocytic leukemia (ALL); Burkitt's lymphoma; ovarian cancer; pancreatic cancer; cancer of the peritoneum; cancer of the omentum; mesentery cancer; pharynx cancer; prostate cancer; rectal cancer; renal cancer; skin cancer; small intestine cancer; soft tissue cancer; solid tumors; synovial sarcoma; gastric cancer; testicular cancer; thyroid cancer; and ureter cancer. In some embodiments, the subject has an autoimmune disorder selected from scleroderma, Grave's disease, Crohn's disease, Sjorgen's disease, multiple sclerosis, Hashimoto's disease, psoriasis, myasthenia gravis, autoimmune polyendocrinopathy syndromes, Type I diabetes mellitus (T1DM), autoimmune gastritis, autoimmune uveoretinitis, polymyositis, colitis, thyroiditis, and the generalized autoimmune diseases typified by human Lupus.

[0040] In another aspect, the present application provides a vector for making a circular RNA polynucleotide, comprising, in the following order, a 5' duplex forming region, a 3' Group I intron fragment, an Internal Ribosome Entry Site (IRES), a first expression sequence, a second expression sequence, a 5' Group I intron fragment, and a 3' duplex forming region.

[0041] In some embodiments, the vector of comprises a polynucleotide sequence encoding a cleavage site between the first expression sequence and the second expression sequence. In some embodiments, the cleavage site is a self-cleaving spacer. In some embodiments, the self-cleaving spacer is a 2A self-cleaving peptide.

[0042] In some embodiments, the vector of comprises a first spacer between the 5'

duplex forming region and the 3' group I intron fragment, and a second spacer between the 5' group I intron fragment and the 3' duplex forming region. In some embodiments, the first and second spacers each have a length of about 5 to about 60 nucleotides. In some embodiments, the first and second spacers each comprise an unstructured region at least 5 nucleotides long. In some embodiments, the first and second spacers each comprise a structured region at least 7 nucleotides long. In some embodiments, the first and second duplex forming regions each have a length of about 9 to 50 nucleotides.

[0043] In some embodiments, the vector is codon optimized. In some embodiments, the vector lacks at least one microRNA binding site present in an equivalent pre-optimization polynucleotide.

[0044] In another aspect, the present application provides a eukaryotic cell comprising a circular RNA polynucleotide as described herein. In some embodiments, the eukaryotic cell is a human cell. In some embodiments, the eukaryotic cell is an immune cell. In some embodiments, the eukaryotic cell is a T cell.

[0045] In an aspect, provided herein is a circular RNA polynucleotide comprising, in the following order, a post splicing 3' group I intron fragment, an Internal Ribosome Entry Site (IRES), a first expression sequence, a polynucleotide sequence encoding a cleavage site, a second expression sequence, and a post splicing 5' group I intron fragment. In an aspect, provided herein is a circular RNA polynucleotide comprising, in the following order, a post splicing 3' group I intron fragment, a first Internal Ribosome Entry Site (IRES), a first expression sequence, a second IRES, a second expression sequence, and a post splicing 5' group I intron fragment. In some embodiments, the first expression sequence and the second expression sequence encode different therapeutic proteins. In some embodiments, the first expression sequence and the second expression sequence encodes for the same therapeutic protein.

[0046] In an aspect, provided herein is a circular RNA polynucleotide produced from transcription of a vector comprising, in the following order, an optional 5' duplex forming region, a post splicing 3' group I intron fragment, an Internal Ribosome Entry Site (IRES), an expression sequence, a polynucleotide sequence encoding a cleavage site, a second expression sequence, a 5' group I intron fragment, and an optional 3' duplex forming region. In an aspect, provided herein is a circular RNA polynucleotide produced from transcription of a vector comprising, in the following order, an optional 5' duplex forming region, a 3' group I intron fragment, a first Internal Ribosome Entry Site (IRES), an expression sequence, a second IRES, a second expression sequence, a 5' group I intron fragment, and an optional 3'

duplex forming region. In some embodiments, a circular RNA polynucleotide or vector provided herein comprises 3' and 5' duplex forming regions.

[0047] In some embodiments, a circular RNA polynucleotide comprises a first spacer between the 5' duplex forming region and the post splicing 3' group I intron fragment, and a second spacer between the post splicing 5' group I intron fragment and the 3' duplex forming region. In some embodiments, the first and second spacers each have a length of about 10 to about 60 nucleotides. In certain embodiments, the first and second duplex forming regions each have a length of about 9 to about 19 nucleotides. In certain other embodiments, the first and second duplex forming regions each have a length of about 30 nucleotides.

[0048] In certain embodiments, the IRES is selected from Table 17 has a sequence of an IRES, or is a functional fragment or variant thereof. In some embodiments, the IRES has a sequence of an IRES from Taura syndrome virus, Triatoma virus, Theiler's encephalomyelitis virus, Simian Virus 40, Solenopsis invicta virus 1, Rhopalosiphum padi virus, Reticuloendotheliosis virus, Human poliovirus 1, Plautia stali intestine virus, Kashmir bee virus, Human rhinovirus 2, Homalodisca coagulata virus- 1, Human Immunodeficiency Virus type 1, Homalodisca coagulata virus- 1, Himetobi P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinitis virus, Ectropis obliqua picorna-like virus, Encephalomyocarditis virus, Drosophila C Virus, Human coxsackievirus B3, Crucifer tobamovirus, Cricket paralysis virus, Bovine viral diarrhea virus 1, Black Queen Cell Virus, Aphid lethal paralysis virus, Avian encephalomyelitis virus, Acute bee paralysis virus, Hibiscus chlorotic ringspot virus, Classical swine fever virus, Human FGF2, Human SFTPA1, Human AML1/RUNX1, Drosophila antennapedia, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 alpha, Human n.myc, Mouse Gtx, Human p27kipl, Human PDGF2/c-sis, Human p53, Human Pim-1, Mouse Rbm3, Drosophila reaper, Canine Scamper, Drosophila Ubx, Human UNR, Mouse UtrA, Human VEGF-A, Human XIAP, Drosophila hairless, S. cerevisiae TFIID, S. cerevisiae YAP1, tobacco etch virus, turnip crinkle virus, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, Picobirnavirus, HCV QC64, Human Cosavirus E/D, Human Cosavirus F, Human Cosavirus JMY, Rhinovirus NAT001, HRV14, HRV89, HRVC-02, HRV-A21, Salivirus A SH1, Salivirus FHB, Salivirus NG-J1, Human Parechovirus 1, Crohivirus B, Yc-3, Rosavirus M-7, Shanbavirus A, Pasivirus A, Pasivirus A 2, Echovirus E14, Human Parechovirus 5, Aichi Virus, Hepatitis A Virus HA16, Phopivirus, CVA10, Enterovirus C, Enterovirus D, Enterovirus J, Human Pegivirus 2, GBV-C GT110,

GBV-C K1737, GBV-C Iowa, Pegivirus A 1220, Pasivirus A 3, Sapelovirus, Rosavirus B, Bakunsa Virus, Tremovirus A, Swine Pasivirus 1, PLV-CHN, Pasivirus A, Sicinivirus, Hepacivirus K, Hepacivirus A, BVDV1, Border Disease Virus, BVDV2, CSFV-PK15C, SF573 Dicistrovirus, Hubei Picorna-like Virus, CRPV, Salivirus A BN5, Salivirus A BN2, Salivirus A 02394, Salivirus A GUT, Salivirus A CH, Salivirus A SZ1, Salivirus FHB, CVB3, CVB1, Echovirus 7, CVB5, EVA71, CVA3, CVA12, EV24 or an aptamer to eIF4G.

[0049] In some embodiments, the first and second polyA sequences each have a length of 15-50nt. In some embodiments, the first and second polyA sequences each have a length of about 20-25nt.

[0050] In certain embodiments, the circular RNA polynucleotide consists of naturally occurring nucleotides. In some embodiments, the circular RNA contains at least about 80%, at least 90%, at least about 95%, or at least about 99% naturally occurring nucleotides. In certain embodiments, the expression sequence is codon-optimized. In certain embodiments, the circular RNA polynucleotide is optimized to lack at least one microRNA binding site present in an equivalent pre-optimized polynucleotide. In certain embodiments, the circular RNA polynucleotide is optimized to lack at least one endonuclease susceptible site present in an equivalent pre-optimized polynucleotide. In certain embodiments, the circular RNA polynucleotide is optimized to lack at least one RNA-editing susceptible site present in an equivalent pre-optimized polynucleotide.

[0051] In some embodiments, the circular RNA polynucleotide is from about 100 nucleotides to about 15 kilobases in length. In certain embodiments, the circular RNA polynucleotide of the present disclosure, has an *in vivo* duration of therapeutic effect in humans of at least about 20 hours. In certain embodiments, the circular RNA polynucleotide has a functional half-life of at least about 20 hours. In certain embodiments, the circular RNA polynucleotide has a duration of therapeutic effect in a human cell greater than or equal to that of an equivalent linear RNA polynucleotide comprising the same expression sequences. In certain embodiments, the circular RNA polynucleotide has a functional half-life in a human cell greater than or equal to that of an equivalent linear RNA polynucleotide comprising the same expression sequences. In certain embodiments, the circular RNA polynucleotide has an *in vivo* functional half-life in humans greater than that of an equivalent linear RNA polynucleotide having the same expression sequences. In some embodiments, the reference linear RNA polynucleotide is a linear, unmodified or nucleoside-modified, fully-processed mRNA comprising a cap1 structure and a polyA tail at least 80nt in length.

[0052] In some embodiments, the pharmaceutical composition has a functional half-life

in a human cell greater than or equal that of a pre-determined threshold value. In some embodiments, the pharmaceutical composition has a functional half-life *in vivo* in humans greater than that of a pre-determined threshold value. In some embodiments, the functional protein assay is an *in vitro* luciferase assay. In some embodiments, the functional protein assay comprises measuring levels of protein encoded by the expression sequence of the circular RNA polynucleotide in a patient serum or tissue sample. In some embodiments, wherein the pre-determined threshold value is the functional half-life of a reference linear RNA polynucleotide comprising the same expression sequence as the circular RNA polynucleotide. In some embodiments, the pharmaceutical composition has a functional half-life of at least about 20 hours.

[0053] In an aspect, provided herein is a pharmaceutical composition comprising a circular RNA polynucleotide, a nanoparticle, and optionally, a targeting moiety operably connected to the nanoparticle. In some embodiments, the nanoparticle is a lipid nanoparticle, a core-shell nanoparticle, a biodegradable nanoparticle, a biodegradable lipid nanoparticle, a polymer nanoparticle, or a biodegradable polymer nanoparticle. In certain embodiments, the nanoparticle comprises one or more cationic lipids selected from the group C12-200, MC3, DLinDMA, DLinkC2DMA, cKK-E12, ICE (Imidazol- based), HGT5000, HGT5001, DODAC, DDAB, DMRIE, DOSPA, DOGS, DODAP, DODMA and DMDMA, DODAC, DLenDMA, DMRIE, CLinDMA, CpLinDMA, DMOBA, DOcarbDAP, DLinDAP, DLincarbDAP, DLinCDAP, KLin-K-DMA, DLin-K-XTC2-DMA, HGT4003, and combinations thereof.

[0054] In certain embodiments, the pharmaceutical composition comprises a targeting moiety, wherein the targeting moiety mediates receptor-mediated endocytosis or direct fusion into selected cells of a selected cell population or tissue in the absence of cell isolation or purification. In some embodiments, the targeting moiety is a scFv, nanobody, peptide, minibody, polynucleotide aptamer, heavy chain variable region, light chain variable region or fragment thereof. In certain embodiments, the circular RNA polynucleotide is in an amount effective to treat an autoimmune disorder or cancer in a human subject in need thereof. In certain embodiments, the pharmaceutical composition has an enhanced safety profile when compared to a pharmaceutical composition comprising vectors comprising exogenous DNA encoding the same expression sequences.

[0055] In certain embodiments, less than 1%, by weight, of the polynucleotides in the composition are double stranded RNA, DNA splints, or triphosphorylated RNA. In certain embodiments, less than 1%, by weight, of the polynucleotides and proteins in the

pharmaceutical composition are double stranded RNA, DNA splints, triphosphorylated RNA, phosphatase proteins, protein ligases, and capping enzymes.

[0056] In an aspect, provided herein is a method of treating a subject in need thereof comprising administering a therapeutically effective amount of a composition comprising the circular RNA polynucleotide, a nanoparticle, and optionally, a targeting moiety operably connected to the nanoparticle. In some embodiments, the subject has an autoimmune disorder or cancer.

[0057] In some embodiments, the targeting moiety is an scFv, nanobody, peptide, minibody, heavy chain variable region, light chain variable region or fragment thereof. In some embodiments, the composition comprises a targeting moiety, wherein the targeting moiety mediates receptor-mediated endocytosis into selected cells of a selected cell population in the absence of cell isolation or purification.

[0058] In some embodiments, the nanoparticle is a lipid nanoparticle, a core-shell nanoparticle, or a biodegradable nanoparticle. In some embodiments, the nanoparticle comprises one or more cationic lipids, ionizable lipids, or poly β -amino esters. In some embodiments, the nanoparticle comprises one or more non-cationic lipids. In some embodiments, the nanoparticle comprises one or more PEG-modified lipids, polyglutamic acid lipids, structural lipid or hyaluronic acid lipids. In certain embodiments, the nanoparticle comprises cholesterol. In some embodiments, the structural lipid is a beta-sitosterol. In some embodiments, the structural lipid is not a beta sitosterol. In some embodiments, the nanoparticle comprises arachidonic acid or oleic acid. In some embodiments, the nanoparticle comprises more than one circular RNA polynucleotide.

[0059] In some embodiments, the structural lipid binds to C1q and/or promotes the binding of the transfer vehicle comprising said lipid to C1q compared to a control transfer vehicle lacking the structural lipid and/or increases uptake of C1q-bound transfer vehicle into an immune cell compared to a control transfer vehicle lacking the structural lipid.

[0060] In some embodiments, the PEG-modified lipid is DSPE-PEG, DMG-PEG, or PEG-1. In some embodiments, the PEG-modified lipid is DSPE-PEG(2000).

[0061] In some embodiments, the pharmaceutical composition further comprises a helper lipid. In some embodiments, the helper lipid is DSPC or DOPE.

[0062] In some embodiments, the pharmaceutical composition comprises DOPE, cholesterol, and DSPE-PEG.

[0063] In some embodiments, the transfer vehicle comprises about 0.5% to about 4% PEG-modified lipids by molar ratio. In some embodiments, the transfer vehicle comprises

about 1% to about 2% PEG-modified lipids by molar ratio.

[0064] In some embodiments, the molar ration of ionizable lipid:DSPC:cholesterol:DSPE-PEG(2000) is 62:4:33:1.

[0065] In some embodiments, the transfer vehicle comprises an ionizable lipid, a DOPE, cholesterol, and a DSPE-PEG(2000).

[0066] In some embodiments, the molar ration of ionizable lipid:DSPC:cholesterol:DSPE-PEG(2000) is 50:10:38.5:1.5.

[0067] In some embodiments, the transfer vehicle has a nitrogen:phosphate (N:P) ratio of about 3 to about 6.

[0068] In some embodiments, the transfer vehicle is formulated for endosomal release of the circular RNA polynucleotide.

[0069] In some embodiments, the transfer vehicle is capable of binding to APOE. In some embodiments, the transfer vehicle interacts with apolipoprotein E (APOE) less than an equivalent transfer vehicle loaded with a reference linear RNA having the same expression sequence as the circular RNA polynucleotide. In some embodiments, the exterior surface of the transfer vehicle is substantially free of APOE binding sites.

[0070] In some embodiments, the transfer vehicle has a diameter of less than about 120nm. In some embodiments, the transfer vehicle does not form aggregates with a diameter of more than 300nm.

[0071] In some embodiments, the transfer vehicle has a diameter of less than about 120nm. In some embodiments, the transfer vehicle does not form aggregates with a diameter of more than 300nm.

[0072] In some embodiments, the transfer vehicle has an in vivo half-life of less than about 30 hours.

[0073] In some embodiments, the transfer vehicle is capable of low density lipoprotein receptor (LDLR) dependent uptake into a cell. In some embodiments, the transfer vehicle is capable of LDLR independent uptake into a cell.

[0074] In some embodiments, the pharmaceutical composition is substantially free of linear RNA.

[0075] In some embodiments, the pharmaceutical composition further comprises a targeting moiety operably connected to the transfer vehicle. In some embodiments, the targeting moiety specifically binds an immune cell antigen or indirectly. In some embodiments, the immune cell antigen is a T cell antigen. In some embodiments, the T cell antigen is selected from the group consisting of CD2, CD3, CD5, CD7, CD8, CD4, beta7

integrin, beta2 integrin, and C1q.

[0076] In some embodiments, the pharmaceutical composition further comprises an adapter molecule comprising a transfer vehicle binding moiety and a cell binding moiety, wherein the targeting moiety specifically binds the transfer vehicle binding moiety and the cell binding moiety specifically binds a target cell antigen. In some embodiments, the target cell antigen is an immune cell antigen. In some embodiments, the immune cell antigen is a T cell antigen, an NK cell, an NKT cell, a macrophage, or a neutrophil. In some embodiments, the T cell antigen is selected from the group consisting of CD2, CD3, CD5, CD7, CD8, CD4, beta7 integrin, beta2 integrin, CD25, CD39, CD73, A2a Receptor, A2b Receptor, and C1q. In some embodiments, the immune cell antigen is a macrophage antigen. In some embodiments, the macrophage antigen is selected from the group consisting of mannose receptor, CD206, and C1q.

[0077] In some embodiments, the targeting moiety is a small molecule. In some embodiments, the small molecule binds to an ectoenzyme on an immune cell, wherein the ectoenzyme is selected from the group consisting of CD38, CD73, adenosine 2a receptor, and adenosine 2b receptor. In some embodiments, the small molecule is mannose, a lectin, acivicin, biotin, or digoxigenin.

[0078] In some embodiments, the targeting moiety is a single chain Fv (scFv) fragment, nanobody, peptide, peptide-based macrocycle, minibody, small molecule ligand such as folate, arginylglycylaspartic acid (RGD), or phenol-soluble modulin alpha 1 peptide (PSMA1), heavy chain variable region, light chain variable region or fragment thereof.

[0079] In some embodiments, the ionizable lipid has a half-life in a cell membrane less than about 2 weeks. In some embodiments, the ionizable lipid has a half-life in a cell membrane less than about 1 week. In some embodiments, the ionizable lipid has a half-life in a cell membrane less than about 30 hours. In some embodiments, the ionizable lipid has a half-life in a cell membrane less than the functional half-life of the circular RNA polynucleotide.

[0080] In another aspect, the present application provides a method of treating or preventing a disease, disorder, or condition, comprising administering an effective amount of a pharmaceutical composition disclosed herein. In some embodiments, the disease, disorder, or condition is associated with aberrant expression, activity, or localization of a polypeptide selected from Tables 27 or 28. In some embodiments, the circular RNA polynucleotide encodes a therapeutic protein. In some embodiments, therapeutic protein expression in the spleen is higher than therapeutic protein expression in the liver. In some embodiments,

therapeutic protein expression in the spleen is at least about 2.9x therapeutic protein expression in the liver. In some embodiments, the therapeutic protein is not expressed at functional levels in the liver. In some embodiments, the therapeutic protein is not expressed at detectable levels in the liver. In some embodiments, therapeutic protein expression in the spleen is at least about 63% of total therapeutic protein expression.

[0081] In some embodiments, the linear RNA polynucleotide comprises a 3' anabaena group I intron fragment and a 5' anabaena group I intron fragment. In some embodiments, the reference RNA polynucleotide comprises a reference 3' anabaena group I intron fragment and a reference 5' anabaena group I intron fragment. In some embodiments, the reference 3' anabaena group I intron fragment and reference 5' anabaena group I intron fragment were generated using the L6-5 permutation site. In some embodiments, the 3' anabaena group I intron fragment and 5' anabaena group I intron fragment were not generated using the L6-5 permutation site. In some embodiments, the 3' anabaena group I intron fragment comprises or consists of a sequence selected from SEQ ID NO: 112-123 and 125-150. In some embodiments, the 5' anabaena group I intron fragment comprises a corresponding sequence selected from SEQ ID NO: 73-84 and 86-111. In some embodiments, the 5' anabaena group I intron fragment comprises or consists of a sequence selected from SEQ ID NO: 73-84 and 86-111. In some embodiments, the 3' anabaena group I intron fragment comprises or consists of a corresponding sequence selected from SEQ ID NO: 112-124 and 125-150.

[0082] In some embodiments, the IRES comprises a nucleotide sequence selected from SEQ ID NOs: 348-351. In some embodiments, the reference IRES is CVB3. In some embodiments, the IRES is not CVB3. In some embodiments, the IRES comprises a sequence selected from SEQ ID NOs: 1-64 and 66-72.

[0083] In another aspect, the present application discloses a circular RNA polynucleotide produced from the linear RNA disclosed herein.

[0084] In another aspect, the present application discloses a circular RNA comprising, from 5' to 3', a 3' group I intron fragment, an IRES, an expression sequence, and a 5' group I intron fragment, wherein the IRES comprises a nucleotide sequence selected from SEQ ID NOs: 348-351.

[0085] In some embodiments, the circular RNA polynucleotide further comprises a spacer between the 3' group I intron fragment and the IRES.

[0086] In some embodiments, the circular RNA polynucleotide further comprises a first and a second duplex forming regions capable of forming a duplex. In some embodiments, the first and second duplex forming regions each have a length of about 9 to 19 nucleotides. In

some embodiments, the first and second duplex forming regions each have a length of about 30 nucleotides.

[0087] In an aspect, provided herein is a vector for making a circular RNA polynucleotide, comprising, in the following order, an optional 5' duplex forming region, a 3' Group I intron fragment, an Internal Ribosome Entry Site (IRES), a first expression sequence, a polynucleotide sequence encoding a cleavage site, a second expression sequence, a 5' Group I intron fragment, and an optional 3' duplex forming region. In an aspect, provided herein is a vector for making a circular RNA polynucleotide, comprising, in the following order, an optional 5' duplex forming region, a 3' Group I intron fragment, a first Internal Ribosome Entry Site (IRES), a first expression sequence, a second IRES, a second expression sequence, a 5' Group I intron fragment, and an optional 3' duplex forming region. In some embodiments, a polynucleotide contains a 3' duplex forming region and a 5' duplex forming region. In certain embodiments, vector comprises a first spacer between the 5' duplex forming region and the 3' group I intron fragment, and a second spacer between the 5' group I intron fragment and the 3' duplex forming region. In certain embodiments, the first and second spacers each have a length of about 5 to about 60 nucleotides. In certain embodiments, the first and second spacers each have a length of about 8 to about 60 nucleotides. In certain embodiments, the first and second spacers each comprise an unstructured region at least 5 nucleotides long. In certain embodiments, the first and second spacers each comprise a structured region at least 7 nucleotides long. In certain embodiments, the first and second duplex forming regions each have a length of about 9 to 50 nucleotides. In certain embodiments, the vector is codon optimized. In certain embodiments, the vector is lacking at least one microRNA binding site present in an equivalent pre-optimization polynucleotide.

[0088] In an aspect, provided herein is a prokaryotic cell comprising a vector for making a circular RNA polynucleotide, comprising, in the following order, an optional 5' duplex forming region, a 3' Group I intron fragment, an Internal Ribosome Entry Site (IRES), a first expression sequence, a polynucleotide sequence encoding a cleavage site, a second expression sequence, a 5' Group I intron fragment, and an optional 3' duplex forming region. In some embodiments, a polynucleotide contains a 3' duplex forming region and a 5' duplex forming region. In an aspect, provided herein is a prokaryotic cell comprising a vector for making a circular RNA polynucleotide, comprising, in the following order, an optional 5' duplex forming region, a 3' Group I intron fragment, a first Internal Ribosome Entry Site (IRES), a first expression sequence, a second IRES, a second expression sequence, a 5' Group I intron fragment, and an

optional 3' duplex forming region. In some embodiments, a polynucleotide contains a 3' duplex forming region and a 5' duplex forming region.

[0089] In an aspect, provided herein is a eukaryotic cell comprising a circular RNA polynucleotide of the present disclosure. In some embodiments, the eukaryotic cell is a human cell. In some embodiments, the eukaryotic cell is an immune cell. In some embodiments, the eukaryotic cell is a T cell, NK cell, an NKT cell, a macrophage, or a neutrophil.

BRIEF DESCRIPTION OF THE DRAWINGS

[0090] FIG. 1 depicts luminescence in supernatants of HEK293 (FIGs. 1A, 1D, and 1E), HepG2 (FIG. 1B), or 1C1C7 (FIG. 1C) cells 24 hours after transfection with circular RNA comprising a *Gaussia* luciferase expression sequence and various IRES sequences.

[0091] FIG. 2 depicts luminescence in supernatants of HEK293 (FIG. 2A), HepG2 (FIG. 2B), or 1C1C7 (FIG. 2C) cells 24 hours after transfection with circular RNA comprising a *Gaussia* luciferase expression sequence and various IRES sequences having different lengths.

[0092] FIG. 3 depicts stability of select IRES constructs in HepG2 (FIG. 3A) or 1C1C7 (FIG. 3B) cells over 3 days as measured by luminescence.

[0093] FIG. 4A and FIG. 4B depict protein expression from select IRES constructs in Jurkat cells, as measured by luminescence from secreted *Gaussia* luciferase in cell supernatants.

[0094] FIG. 5A and FIG. 5B depict stability of select IRES constructs in Jurkat cells over 3 days as measured by luminescence.

[0095] FIG. 6 depicts comparisons of 24 hour luminescence (FIG. 6A) or relative luminescence over 3 days (FIG. 6B) of modified linear, unpurified circular, or purified circular RNA encoding *Gaussia* luciferase.

[0096] FIG. 7 depicts transcript induction of IFN γ (FIG. 7A), IL-6 (FIG. 7B), IL-2 (FIG. 7C), RIG-I (FIG. 7D), IFN- β 1 (FIG. 7E), and TNF α (FIG. 7F) after electroporation of Jurkat cells with modified linear, unpurified circular, or purified circular RNA.

[0097] FIG. 8 depicts a comparison of luminescence of circular RNA and modified linear RNA encoding *Gaussia* luciferase in human primary monocytes (FIG. 8A) and macrophages (FIG. 8B and FIG. 8C).

[0098] FIG. 9 depicts relative luminescence over 3 days (FIG. 9A) in supernatant of primary T cells after transduction with circular RNA comprising a *Gaussia* luciferase expression sequence and varying IRES sequences or 24 hour luminescence (FIG. 9B).

[0099] FIG. 10 depicts 24 hour luminescence in supernatant of primary T cells (FIG.

10A) after transduction with circular RNA or modified linear RNA comprising a *Gaussia* luciferase expression sequence, or relative luminescence over 3 days (FIG. 10B), and 24 hour luminescence in PBMCs (FIG. 10C).

[0100] FIG. 11 depicts HPLC chromatograms (FIG. 11A) and circularization efficiencies (FIG. 11B) of RNA constructs having different permutation sites.

[0101] FIG. 12 depicts HPLC chromatograms (FIG. 12A) and circularization efficiencies (FIG. 12B) of RNA constructs having different introns and/or permutation sites.

[0102] FIG. 13 depicts HPLC chromatograms (FIG. 13A) and circularization efficiencies (FIG. 13B) of 3 RNA constructs with or without homology arms.

[0103] FIG. 14 depicts circularization efficiencies of 3 RNA constructs without homology arms or with homology arms having various lengths and GC content.

[0104] FIG. 15A and 15B depict HPLC chromatograms showing the contribution of strong homology arms to improved splicing efficiency, the relationship between circularization efficiency and nicking in select constructs, and combinations of permutations sites and homology arms hypothesized to demonstrate improved circularization efficiency.

[0105] FIG. 16 shows fluorescent images of T cells mock electroporated (left) or electroporated with circular RNA encoding a CAR (right) and co-cultured with Raji cells expressing GFP and firefly luciferase.

[0106] FIG. 17 shows bright field (left), fluorescent (center), and overlay (right) images of T cells mock electroporated (top) or electroporated with circular RNA encoding a CAR (bottom) and co-cultured with Raji cells expressing GFP and firefly luciferase.

[0107] FIG. 18 depicts specific lysis of Raji target cells by T cells mock electroporated or electroporated with circular RNA encoding different CAR sequences.

[0108] FIG. 19 depicts luminescence in supernatants of Jurkat cells (left) or resting primary human CD3+ T cells (right) 24 hours after transduction with linear or circular RNA comprising a *Gaussia* luciferase expression sequence and varying IRES sequences (FIG. 19A), and relative luminescence over 3 days (FIG. 19B).

[0109] FIG. 20 depicts transcript induction of IFN- β 1 (Fig. 20A), RIG-I (Fig. 20B), IL-2 (Fig. 20C), IL-6 (Fig. 20D), IFN γ (Fig. 20E), and TNF α (Fig. 20F) after electroporation of human CD3+ T cells with modified linear, unpurified circular, or purified circular RNA.

[0110] FIG. 21 depicts specific lysis of Raji target cells by human primary CD3+ T cells electroporated with circRNA encoding a CAR as determined by detection of firefly luminescence (FIG. 21A), and IFN γ transcript induction 24 hours after electroporation with different quantities of circular or linear RNA encoding a CAR sequence (FIG. 21B).

- [0111] FIG. 22 depicts specific lysis of target or non-target cells by human primary CD3+ T cells electroporated with circular or linear RNA encoding a CAR at different E:T ratios (FIG. 22A and FIG. 22B) as determined by detection of firefly luminescence.
- [0112] FIG. 23 depicts specific lysis of target cells by human CD3+ T cells electroporated with RNA encoding a CAR at 1, 3, 5, and 7 days post electroporation.
- [0113] FIG. 24 depicts specific lysis of target cells by human CD3+ T cells electroporated with circular RNA encoding a CD19 or BCMA targeted CAR.
- [0114] FIG. 25 depicts total Flux of organs harvested from CD-1 mice dosed with circular RNA encoding FLuc and formulated with 50% Lipid 15 (Table 10b), 10% DSPC, 1.5% PEG-DMG, and 38.5% cholesterol.
- [0115] FIG. 26 shows images highlighting the luminescence of organs harvested from CD-1 mice dosed with circular RNA encoding FLuc and formulated with 50% Lipid 15 (Table 10b), 10% DSPC, 1.5% PEG-DMG, and 38.5% cholesterol.
- [0116] FIG. 27 depicts molecular characterization of Lipids 26 and 27 from Table 10a. FIG. 27A shows the proton nuclear magnetic resonance (NMR) spectrum of Lipid 26. FIG. 27B shows the retention time of Lipid 26 measured by liquid chromatography-mass spectrometry (LC-MS). FIG. 27C shows the mass spectrum of Lipid 26. FIG. 27D shows the proton NMR spectrum of Lipid 27. FIG. 27E shows the retention time of Lipid 27 measured by LC-MS. FIG. 27F shows the mass spectrum of Lipid 27.
- [0117] FIG. 28 depicts molecular characterization of Lipid 22-S14 and its synthetic intermediates. FIG. 28A depicts the NMR spectrum of 2-(tetradecylthio)ethan-1-ol. FIG. 28B depicts the NMR spectrum of 2-(tetradecylthio)ethyl acrylate. FIG. 28C depicts the NMR spectrum of bis(2-(tetradecylthio)ethyl) 3,3'-((3-(2-methyl-1H-imidazol-1-yl)propyl)azanediyl)dipropionate (Lipid 22-S14).
- [0118] FIG. 29 depicts the NMR spectrum of bis(2-(tetradecylthio)ethyl) 3,3'-((3-(1H-imidazol-1-yl)propyl)azanediyl)dipropionate (Lipid 93-S14).
- [0119] FIG. 30 depicts molecular characterization of heptadecan-9-yl 8-((3-(2-methyl-1H-imidazol-1-yl)propyl)(8-(nonyloxy)-8-oxooctyl)amino)octanoate (Lipid 54 from Table 10a). FIG. 30A shows the proton NMR spectrum of Lipid 54. FIG. 30B shows the retention time of Lipid 54 measured by LC-MS. FIG. 30C shows the mass spectrum of Lipid 54.
- [0120] FIG. 31 depicts molecular characterization of heptadecan-9-yl 8-((3-(1H-imidazol-1-yl)propyl)(8-(nonyloxy)-8-oxooctyl)amino)octanoate (Lipid 53 from Table 10a). FIG. 31A shows the proton NMR spectrum of Lipid 53. FIG. 31B shows the retention time of Lipid 53 measured by LC-MS. FIG. 31C shows the mass spectrum of Lipid 53.

[0121] FIG. 32A depicts total flux of spleen and liver harvested from CD-1 mice dosed with circular RNA encoding firefly luciferase (FLuc) and formulated with ionizable lipid of interest, DSPC, cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio. FIG. 32B depicts average radiance for biodistribution of protein expression.

[0122] FIG. 33A depicts images highlighting the luminescence of organs harvested from CD-1 mice dosed with circular RNA encoding FLuc and formulated with ionizable Lipid 22-S14, DSPC, cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio. FIG. 33B depicts whole body IVIS images of CD-1 mice dosed with circular RNA encoding FLuc and formulated with ionizable Lipid 22-S14, DSPC, cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio.

[0123] FIG. 34A depicts images highlighting the luminescence of organs harvested from CD-1 mice dosed with circular RNA encoding FLuc and formulated with ionizable Lipid 93-S14, DSPC, cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio. FIG. 34B depicts whole body IVIS images of CD-1 mice dosed with circular RNA encoding FLuc and formulated with ionizable Lipid 93-S14, DSPC, cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio.

[0124] FIG. 35A depicts images highlighting the luminescence of organs harvested from CD-1 mice dosed with circular RNA encoding FLuc and formulated with ionizable Lipid 26 from Table 10a, DSPC, cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio. FIG. 35B depicts whole body IVIS images of CD-1 mice dosed with circular RNA encoding FLuc and formulated with ionizable Lipid 26, DSPC, cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio.

[0125] FIG. 36 depicts images highlighting the luminescence of organs harvested from c57BL/6J mice dosed with circular RNA encoding FLuc and encapsulated in lipid nanoparticles formed with Lipid 15 from Table 10b (FIG. 36A), Lipid 53 from Table 10a (FIG. 36B), or Lipid 54 from Table 10a (FIG. 36C). PBS was used as control (FIG. 36D).

[0126] FIG. 37A and FIG. 37B depict relative luminescence in the lysates of human PBMCs after 24-hour incubation with testing lipid nanoparticles containing circular RNA encoding firefly luciferase.

[0127] FIG. 38 shows the expression of GFP (FIG. 38A) and CD19 CAR (FIG. 38B) in

human PBMCs after incubating with testing lipid nanoparticle containing circular RNA encoding either GFP or CD19 CAR.

[0128] FIG. 39 depicts the expression of an anti-murine CD19 CAR in 1C1C7 cells lipotransfected with circular RNA comprising an anti-murine CD19 CAR expression sequence and varying IRES sequences.

[0129] FIG. 40 shows the cytotoxicity of an anti-murine CD19 CAR to murine T cells. The CD19 CAR is encoded by and expressed from a circular RNA, which is electroporated into the murine T cells.

[0130] FIG. 41 depicts the B cell counts in peripheral blood (FIG. 40A and FIG. 40B) or spleen (FIG. 40C) in C57BL/6J mice injected every other day with testing lipid nanoparticles encapsulating a circular RNA encoding an anti-murine CD19 CAR.

[0131] FIGs. 42A and 42B compares the expression level of an anti-human CD19 CAR expressed from a circular RNA with that expressed from a linear mRNA.

[0132] FIGs. 43A and 43B compares the cytotoxic effect of an anti-human CD19 CAR expressed from a circular RNA with that expressed from a linear mRNA

[0133] FIG. 44 depicts the cytotoxicity of two CARs (anti-human CD19 CAR and anti-human BCMA CAR) expressed from a single circular RNA in T cells.

[0134] FIG. 45A shows representative FACS plots with frequencies of tdTomato expression in various spleen immune cell subsets following treatment with LNPs formed with Lipid 27 or 26 from Table 10a or Lipid 15 from Table 10b. FIG. 45B shows the quantification of the proportion of myeloid cells, B cells, and T cells expressing tdTomato (mean + std. dev., n = 3), equivalent to the proportion of each cell population successfully transfected with Cre circular RNA. FIG. 45C illustrates the proportion of additional splenic immune cell populations, including NK cells, classical monocytes, nonclassical monocytes, neutrophils, and dendritic cells, expressing tdTomato after treatment with Lipids 27 and 26 (mean + std. dev., n = 3).

[0135] FIG. 46A depicts an exemplary RNA construct design with built-in polyA sequences in the introns. FIG. 46B shows the chromatography trace of unpurified circular RNA. FIG. 46C shows the chromatography trace of affinity-purified circular RNA. FIG. 46D shows the immunogenicity of the circular RNAs prepared with varying IVT conditions and purification methods. (Commercial = commercial IVT mix; Custom = customized IVT mix; Aff = affinity purification; Enz = enzyme purification; GMP:GTP ratio = 8, 12.5, or 13.75).

[0136] FIG. 47A depicts an exemplary RNA construct design with a dedicated binding sequence as an alternative to polyA for hybridization purification. FIG. 47B shows the

chromatography trace of unpurified circular RNA. **FIG. 46C** shows the chromatography trace of affinity-purified circular RNA.

[0137] **FIG. 48A** shows the chromatography trace of unpurified circular RNA encoding dystrophin. **FIG. 48B** shows the chromatography trace of enzyme-purified circular RNA encoding dystrophin.

[0138] **FIG. 49** compares the expression (**FIG. 49A**) and stability (**FIG. 49B**) of purified circRNAs with different 5' spacers between the 3' intron fragment/5' internal duplex region and the IRES in Jurkat cells. (AC = only A and C were used in the spacer sequence; UC = only U and C were used in the spacer sequence.)

[0139] **FIG. 50** shows luminescence expression levels and stability of expression in primary T cells from circular RNAs containing the original or modified IRES elements indicated.

[0140] **FIG. 51** shows luminescence expression levels and stability of expression in HepG2 cells from circular RNAs containing the original or modified IRES elements indicated.

[0141] **FIG. 52** shows luminescence expression levels and stability of expression in 1C1C7 cells from circular RNAs containing the original or modified IRES elements indicated.

[0142] **FIG. 53** shows luminescence expression levels and stability of expression in HepG2 cells from circular RNAs containing IRES elements with untranslated regions (UTRs) inserted or hybrid IRES elements. "Scr" means Scrambled, which was used as a control.

[0143] **FIG. 54** shows luminescence expression levels and stability of expression in 1C1C7 cells from circular RNAs containing an IRES and variable stop codon cassettes operably linked to a *Gaussia* luciferase coding sequence.

[0144] **FIG. 55** shows luminescence expression levels and stability of expression in 1C1C7 cells from circular RNAs containing an IRES and variable untranslated regions (UTRs) inserted before the start codon of a gaussian luciferase coding sequence.

[0145] **FIG. 56** shows expression levels of human erythropoietin (hEPO) in Huh7 cells from circular RNAs containing two miR-122 target sites downstream from the hEPO coding sequence.

[0146] **FIG. 57** shows CAR expression levels in the peripheral blood (**FIG. 57A**) and spleen (**FIG. 57B**) when treated with LNP encapsulating circular RNA that expresses anti-CD19 CAR. Anti-CD20 (aCD20) and circular RNA encoding luciferase (oLuc) were used for comparison.

[0147] **FIG. 58** shows the overall frequency of anti-CD19 CAR expression, the

frequency of anti-CD19 CAR expression on the surface of cells and effect on anti-tumor response of IRES specific circular RNA encoding anti-CD19 CARs on T-cells. **FIG. 58A** shows anti-CD19 CAR geometric mean florescence intensity, **FIG. 58B** shows percentage of anti-CD19 CAR expression, and **FIG. 58C** shows the percentage target cell lysis performed by the anti-CD19 CAR. (CK = Caprine Kobuvirus; AP = Apodemus Picornavirus; CK* = Caprine Kobuvirus with codon optimization; PV = Parabovirus; SV = Salivirus.)

[0148] **FIG. 59** shows CAR expression levels of A20 FLuc target cells when treated with IRES specific circular RNA constructs.

[0149] **FIG. 60** shows luminescence expression levels for cytosolic (**FIG. 60A**) and surface (**FIG. 60B**) proteins from circular RNA in primary human T-cells.

[0150] **FIG. 61** shows luminescence expression in human T-cells when treated with IRES specific circular constructs. Expression in circular RNA constructs were compared to linear mRNA. **FIG. 61A**, **FIG. 61B**, and **FIG. 61G** provide *Gaussia* luciferase expression in multiple donor cells. **FIG. 61C**, **FIG. 61D**, **FIG. 61E**, and **FIG. 61F** provides firefly luciferase expression in multiple donor cells.

[0151] **FIG. 62** shows anti-CD19 CAR (**FIG. 62A** and **FIG. 62B**) and anti-BCMA CAR (**FIG. 62B**) expression in human T-cells following treatment of a lipid nanoparticle encompassing a circular RNA that encodes either an anti-CD19 or anti-BCMA CAR to a firefly luciferase expressing K562 cell.

[0152] **FIG. 63** shows anti-CD19 CAR expression levels resulting from delivery via electroporation *in vitro* of a circular RNA encoding an anti-CD19 CAR in a specific antigen-dependent manner. **FIG. 63A** shows Nalm6 cell lysing with an anti-CD19 CAR. **FIG. 63B** shows K562 cell lysing with an anti-CD19 CAR.

[0153] **FIG. 64** shows transfection of LNP mediated by use of ApoE3 in solutions containing LNP and circular RNA expressing green fluorescence protein (GFP). **FIG. 64A** showed the live-dead results. **FIG. 64B**, **FIG. 61C**, **FIG. 61D**, and **FIG. 64E** provide the frequency of expression for multiple donors.

DETAILED DESCRIPTION

[0154] The present invention provides, among other things, methods and compositions for treating an autoimmune disorder or cancer based on circular RNA therapy. In particular, the present invention provides methods for treating an autoimmune disorder or cancer by administering to a subject in need of treatment a composition comprising an RNA encoding 2 therapeutic proteins at an effective dose and an administration interval such that at least one

symptom or feature of an autoimmune disorder or cancer is reduced in intensity, severity, or frequency or is delayed in onset.

[0155] In certain embodiments, provided herein is a vector for making circular RNA, the vector comprising an optional 5' duplex forming region, a 3' group I intron fragment, optionally a first spacer, an Internal Ribosome Entry Site (IRES), a first expression sequence, a polynucleotide sequence encoding a cleavage site, a second expression sequence, optionally a second spacer, a 5' group I intron fragment, and an optional 3' duplex forming region. In certain embodiments, provided herein is a vector for making circular RNA, the vector comprising an optional 5' duplex forming region, a 3' group I intron fragment, optionally a first spacer, a first Internal Ribosome Entry Site (IRES), a first expression sequence, a second IRES, a second expression sequence, optionally a second spacer, a 5' group I intron fragment, and an optional 3' duplex forming region. In some embodiments, these elements are positioned in the vector in the above order. In some embodiments, a polynucleotide contains a 3' duplex forming region and a 5' duplex forming region. In some embodiments, the vector further comprises an internal 5' duplex forming region between the 3' group I intron fragment and the IRES and an internal 3' duplex forming region between the expression sequences and the 5' group I intron fragment. In some embodiments, the internal duplex forming regions are capable of forming a duplex between each other but not with the external duplex forming regions. In some embodiments, the internal duplex forming regions are part of the first and second spacers. Additional embodiments include circular RNA polynucleotides, including circular RNA polynucleotides made using the vectors provided herein, compositions comprising such circular RNA, cells comprising such circular RNA, methods of using and making such vectors, circular RNA, compositions and cells.

[0156] In some embodiments, provided herein are methods comprising administration of circular RNA polynucleotides provided herein into cells for therapy or production of useful proteins. In some embodiments, the method is advantageous in providing the production of a desired polypeptide inside eukaryotic cells with a longer half-life than linear RNA, due to the resistance of the circular RNA to ribonucleases.

[0157] Circular RNA polynucleotides lack the free ends necessary for exonuclease-mediated degradation, causing them to be resistant to several mechanisms of RNA degradation and granting extended half-lives when compared to an equivalent linear RNA. Circularization may allow for the stabilization of RNA polynucleotides that generally suffer from short half-lives and may improve the overall efficacy of exogenous mRNA in a variety of applications. In an embodiment, the functional half-life of the circular RNA

polynucleotides provided herein in eukaryotic cells (*e.g.*, mammalian cells, such as human cells) as assessed by protein synthesis is at least 20 hours (*e.g.*, at least 80 hours).

1. Definitions

[0158] As used herein, the terms “circRNA” or “circular polyribonucleotide” or “circular RNA” are used interchangeably and refers to a polyribonucleotide that forms a circular structure through covalent bonds.

[0159] As used herein, the term “3’ group I intron fragment” refers to a sequence with 75% or higher similarity to the 3’-proximal end of a natural group I intron including the splice site dinucleotide and optionally a stretch of natural exon sequence.

[0160] As used herein, the term “5’ group I intron fragment” refers to a sequence with 75% or higher similarity to the 5’-proximal end of a natural group I intron including the splice site dinucleotide and optionally a stretch of natural exon sequence.

[0161] As used herein, the term “permutation site” refers to the site in a group I intron where a cut is made prior to permutation of the intron. This cut generates 3’ and 5’ group I intron fragments that are permuted to be on either side of a stretch of precursor RNA to be circularized.

[0162] As used herein, the term “splice site” refers to a dinucleotide that is partially or fully included in a group I intron and between which a phosphodiester bond is cleaved during RNA circularization.

[0163] The expression sequences in the polynucleotide construct may be separated by a “cleavage site” sequence which enables polypeptides encoded by the expression sequences, once translated, to be expressed separately by the cell.

[0164] A “self-cleaving peptide” refers to a peptide which is translated without a peptide bond between two adjacent amino acids, or functions such that when the polypeptide comprising the proteins and the self-cleaving peptide is produced, it is immediately cleaved or separated into distinct and discrete first and second polypeptides without the need for any external cleavage activity.

[0165] As used herein, the term “therapeutic protein” refers to any protein that, when administered to a subject directly or indirectly in the form of a translated nucleic acid, has a therapeutic, diagnostic, and/or prophylactic effect and/or elicits a desired biological and/or pharmacological effect.

[0166] The α and β chains of $\alpha\beta$ TCR's are generally regarded as each having two

domains or regions, namely variable and constant domains/regions. The variable domain consists of a concatenation of variable regions and joining regions. In the present specification and claims, the term “TCR alpha variable domain” therefore refers to the concatenation of TRAV and TRAJ regions, and the term TCR alpha constant domain refers to the extracellular TRAC region, or to a C-terminal truncated TRAC sequence. Likewise the term “TCR beta variable domain” refers to the concatenation of TRBV and TRBD/TRBJ regions, and the term TCR beta constant domain refers to the extracellular TRBC region, or to a C-terminal truncated TRBC sequence.

[0167] As used herein, the term “immunogenic” refers to a potential to induce an immune response to a substance. An immune response may be induced when an immune system of an organism or a certain type of immune cell is exposed to an immunogenic substance. The term “non-immunogenic” refers to a lack of or absence of an immune response above a detectable threshold to a substance. No immune response is detected when an immune system of an organism or a certain type of immune cell is exposed to a non-immunogenic substance. In some embodiments, a non-immunogenic circular polyribonucleotide as provided herein, does not induce an immune response above a pre-determined threshold when measured by an immunogenicity assay. In some embodiments, no innate immune response is detected when an immune system of an organism or a certain type of immune cell is exposed to a non-immunogenic circular polyribonucleotide as provided herein. In some embodiments, no adaptive immune response is detected when an immune system of an organism or a certain type of immune cell is exposed to a non-immunogenic circular polyribonucleotide as provided herein.

[0168] As used herein, the term “circularization efficiency” refers to a measurement of resultant circular polyribonucleotide as compared to its linear starting material.

[0169] As used herein, the term “translation efficiency” refers to a rate or amount of protein or peptide production from a ribonucleotide transcript. In some embodiments, translation efficiency can be expressed as amount of protein or peptide produced per given amount of transcript that codes for the protein or peptide.

[0170] The term “nucleotide” refers to a ribonucleotide, a deoxyribonucleotide, a modified form thereof, or an analog thereof. Nucleotides include species that comprise purines, *e.g.*, adenine, hypoxanthine, guanine, and their derivatives and analogs, as well as pyrimidines, *e.g.*, cytosine, uracil, thymine, and their derivatives and analogs. Nucleotide analogs include nucleotides having modifications in the chemical structure of the base, sugar and/or phosphate, including, but not limited to, 5'-position pyrimidine modifications, 8'-

position purine modifications, modifications at cytosine exocyclic amines, and substitution of 5-bromo-uracil; and 2'-position sugar modifications, including but not limited to, sugar-modified ribonucleotides in which the 2'-OH is replaced by a group such as an H, OR, R, halo, SH, SR, NH₂, NHR, NR₂, or CN, wherein R is an alkyl moiety as defined herein. Nucleotide analogs are also meant to include nucleotides with bases such as inosine, queuosine, xanthine; sugars such as 2'-methyl ribose; non-natural phosphodiester linkages such as methylphosphonate, phosphorothioate and peptide linkages. Nucleotide analogs include 5-methoxyuridine, 1-methylpseudouridine, and 6-methyladenosine.

[0171] The term “nucleic acid” and “polynucleotide” are used interchangeably herein to describe a polymer of any length, *e.g.*, greater than about 2 bases, greater than about 10 bases, greater than about 100 bases, greater than about 500 bases, greater than 1000 bases, or up to about 10,000 or more bases, composed of nucleotides, *e.g.*, deoxyribonucleotides or ribonucleotides, and may be produced enzymatically or synthetically (*e.g.*, as described in U.S. Pat. No. 5,948,902 and the references cited therein), which can hybridize with naturally occurring nucleic acids in a sequence specific manner analogous to that of two naturally occurring nucleic acids, *e.g.*, can participate in Watson-Crick base pairing interactions. Naturally occurring nucleic acids are comprised of nucleotides including guanine, cytosine, adenine, thymine, and uracil (G, C, A, T, and U respectively).

[0172] The terms “ribonucleic acid” and “RNA” as used herein mean a polymer composed of ribonucleotides.

[0173] The terms “deoxyribonucleic acid” and “DNA” as used herein mean a polymer composed of deoxyribonucleotides.

[0174] “Isolated” or “purified” generally refers to isolation of a substance (for example, in some embodiments, a compound, a polynucleotide, a protein, a polypeptide, a polynucleotide composition, or a polypeptide composition) such that the substance comprises a significant percent (*e.g.*, greater than 1%, greater than 2%, greater than 5%, greater than 10%, greater than 20%, greater than 50%, or more, usually up to about 90%-100%) of the sample in which it resides. In certain embodiments, a substantially purified component comprises at least 50%, 80%-85%, or 90%-95% of the sample. Techniques for purifying polynucleotides and polypeptides of interest are well-known in the art and include, for example, ion-exchange chromatography, affinity chromatography and sedimentation according to density. Generally, a substance is purified when it exists in a sample in an amount, relative to other components of the sample, that is more than as it is found naturally.

[0175] The terms “duplexed,” “double-stranded,” or “hybridized” as used herein refer to

nucleic acids formed by hybridization of two single strands of nucleic acids containing complementary sequences. In most cases, genomic DNA is double-stranded. Sequences can be fully complementary or partially complementary.

[0176] As used herein, “unstructured” with regard to RNA refers to an RNA sequence that is not predicted by the RNAFold software or similar predictive tools to form a structure (*e.g.*, a hairpin loop) with itself or other sequences in the same RNA molecule. In some embodiments, unstructured RNA can be functionally characterized using nuclease protection assays.

[0177] As used herein, “structured” with regard to RNA refers to an RNA sequence that is predicted by the RNAFold software or similar predictive tools to form a structure (*e.g.*, a hairpin loop) with itself or other sequences in the same RNA molecule.

[0178] As used herein, two “duplex forming regions,” “homology arms,” or “homology regions,” complement, or are complementary, to one another when the two regions share a sufficient level of sequence identity to one another’s reverse complement to act as substrates for a hybridization reaction. As used herein, polynucleotide sequences have “homology” when they are either identical or share sequence identity to a reverse complement or “complementary” sequence. The percent sequence identity between a duplex forming region and a counterpart duplex forming region’s reverse complement can be any percent of sequence identity that allows for hybridization to occur. In some embodiments, an internal duplex forming region of an inventive polynucleotide is capable of forming a duplex with another internal duplex forming region and does not form a duplex with an external duplex forming region.

[0179] Linear nucleic acid molecules are said to have a “5’-terminus” (5’ end) and a “3’-terminus” (3’ end) because nucleic acid phosphodiester linkages occur at the 5’ carbon and 3’ carbon of the sugar moieties of the substituent mononucleotides. The end nucleotide of a polynucleotide at which a new linkage would be to a 5’ carbon is its 5’ terminal nucleotide. The end nucleotide of a polynucleotide at which a new linkage would be to a 3’ carbon is its 3’ terminal nucleotide. A terminal nucleotide, as used herein, is the nucleotide at the end position of the 3’- or 5’-terminus.

[0180] “Transcription” means the formation or synthesis of an RNA molecule by an RNA polymerase using a DNA molecule as a template. The invention is not limited with respect to the RNA polymerase that is used for transcription. For example, in some embodiments, a T7-type RNA polymerase can be used.

[0181] “Translation” means the formation of a polypeptide molecule by a ribosome based

upon an RNA template.

[0182] It is to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting. As used in this specification and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to “a cell” includes combinations of two or more cells, or entire cultures of cells; reference to “a polynucleotide” includes, as a practical matter, many copies of that polynucleotide. Unless specifically stated or obvious from context, as used herein, the term “or” is understood to be inclusive. Unless defined herein and below in the remainder of the specification, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains.

[0183] Unless specifically stated or obvious from context, as used herein, the term “about,” is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. “About” can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 0.1%, 0.09%, 0.08%, 0.07%, 0.06%, 0.05%, 0.04%, 0.03%, 0.02%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term “about.”

[0184] As used herein, the term “encode” refers broadly to any process whereby the information in a polymeric macromolecule is used to direct the production of a second molecule that is different from the first. The second molecule may have a chemical structure that is different from the chemical nature of the first molecule.

[0185] By “co-administering” is meant administering a therapeutic agent provided herein in conjunction with one or more additional therapeutic agents sufficiently close in time such that the therapeutic agent provided herein can enhance the effect of the one or more additional therapeutic agents, or vice versa.

[0186] The terms “treat,” and “prevent” as well as words stemming therefrom, as used herein, do not necessarily imply 100% or complete treatment or prevention. Rather, there are varying degrees of treatment or prevention of which one of ordinary skill in the art recognizes as having a potential benefit or therapeutic effect. The treatment or prevention provided by the method disclosed herein can include treatment or prevention of one or more conditions or symptoms of the disease. Also, for purposes herein, “prevention” can encompass delaying the onset of the disease, or a symptom or condition thereof.

[0187] As used herein, “autoimmunity” is defined as persistent and progressive immune

reactions to non-infectious self-antigens, as distinct from infectious non self-antigens from bacterial, viral, fungal, or parasitic organisms which invade and persist within mammals and humans. Autoimmune conditions include scleroderma, Grave's disease, Crohn's disease, Sjorgen's disease, multiple sclerosis, Hashimoto's disease, psoriasis, myasthenia gravis, autoimmune polyendocrinopathy syndromes, Type I diabetes mellitus (T1DM), autoimmune gastritis, autoimmune uveoretinitis, polymyositis, colitis, and thyroiditis, as well as in the generalized autoimmune diseases typified by human Lupus. "Autoantigen" or "self-antigen" as used herein refers to an antigen or epitope which is native to the mammal and which is immunogenic in said mammal.

[0188] As used herein, the term "expression sequence" can refer to a nucleic acid sequence that encodes a product, *e.g.*, a peptide or polypeptide, regulatory nucleic acid, or non-coding nucleic acid. An exemplary expression sequence that codes for a peptide or polypeptide can comprise a plurality of nucleotide triads, each of which can code for an amino acid and is termed as a "codon".

[0189] As used herein, a "spacer" refers to a region of a polynucleotide sequence ranging from 1 nucleotide to hundreds or thousands of nucleotides separating two other elements along a polynucleotide sequence. The sequences can be defined or can be random. A spacer is typically non-coding. In some embodiments, spacers include duplex forming regions.

[0190] As used herein, an "internal ribosome entry site" or "IRES" refers to an RNA sequence or structural element ranging in size from 10 nt to 1000 nt or more, capable of initiating translation of a polypeptide in the absence of a typical RNA cap structure. An IRES is typically about 500 nt to about 700 nt in length.

[0191] As used herein, an "miRNA site" refers to a stretch of nucleotides within a polynucleotide that is capable of forming a duplex with at least 8 nucleotides of a natural miRNA sequence.

[0192] As used herein, an "endonuclease site" refers to a stretch of nucleotides within a polynucleotide that is capable of being recognized and cleaved by an endonuclease protein.

[0193] As used herein, "bicistronic RNA" refers to a polynucleotide that includes two expression sequences coding for two distinct proteins. These expression sequences are often separated by a cleavable peptide such as a 2A site or an IRES sequence. They can also be separated by a ribosomal skipping element or a protease cleavage.

[0194] As used herein, the term "ribosomal skipping element" refers to a nucleotide sequence encoding a short peptide sequence capable of causing generation of two peptide chains from translation of one RNA molecule. While wishing not to be bound by theory, it is

hypothesized that the ribosomal skipping elements function by: (1) terminating translation of the first peptide chain and re-initiating translation of the second peptide chain; or (2) cleavage of a peptide bond in the peptide sequence encoded by the ribosomal skipping element by an intrinsic protease activity of the encoded peptide, or by another protease in the environment (*e.g.*, cytosol).

[0195] As used herein, the term “co-formulate” refers to a nanoparticle formulation comprising two or more nucleic acids or a nucleic acid and other active drug substance. Typically, the ratios are equimolar or defined in the ratiometric amount of the two or more nucleic acids or the nucleic acid and other active drug substance.

[0196] As used herein, “transfer vehicle” includes any of the standard pharmaceutical carriers, diluents, excipients, and the like, which are generally intended for use in connection with the administration of biologically active agents, including nucleic acids.

[0197] As used herein, the phrase “lipid nanoparticle” refers to a transfer vehicle comprising one or more lipids (*e.g.*, in some embodiments, cationic lipids, non-cationic lipids, and PEG-modified lipids).

[0198] As used herein, the phrase “cationic lipid” refers to any of a number of lipid species that carry a net positive charge at a selected pH, such as physiological pH.

[0199] As used herein, the phrase “non-cationic lipid” refers to any neutral, zwitterionic or anionic lipid.

[0200] As used herein, the phrase “anionic lipid” refers to any of a number of lipid species that carry a net negative charge at a selected pH, such as physiological pH.

[0201] As used herein, the phrase “ionizable lipid” refers to any of a number of lipid species that carry a net positive charge at a selected pH, such as physiological pH 4 and a neutral charge at other pHs such as physiological pH 7.

[0202] In some embodiments, a lipid, *e.g.*, an ionizable lipid, disclosed herein comprises one or more cleavable groups. The terms “cleave” and “cleavable” are used herein to mean that one or more chemical bonds (*e.g.*, one or more of covalent bonds, hydrogen-bonds, van der Waals' forces and/or ionic interactions) between atoms in or adjacent to the subject functional group are broken (*e.g.*, hydrolyzed) or are capable of being broken upon exposure to selected conditions (*e.g.*, upon exposure to enzymatic conditions). In certain embodiments, the cleavable group is a disulfide functional group, and in particular embodiments is a disulfide group that is capable of being cleaved upon exposure to selected biological conditions (*e.g.*, intracellular conditions). In certain embodiments, the cleavable group is an ester functional group that is capable of being cleaved upon exposure to selected biological

conditions. For example, the disulfide groups may be cleaved enzymatically or by a hydrolysis, oxidation or reduction reaction. Upon cleavage of such disulfide functional group, the one or more functional moieties or groups (*e.g.*, one or more of a head-group and/or a tail-group) that are bound thereto may be liberated. Exemplary cleavable groups may include, but are not limited to, disulfide groups, ester groups, ether groups, and any derivatives thereof (*e.g.*, alkyl and aryl esters). In certain embodiments, the cleavable group is not an ester group or an ether group. In some embodiments, a cleavable group is bound (*e.g.*, bound by one or more of hydrogen-bonds, van der Waals' forces, ionic interactions and covalent bonds) to one or more functional moieties or groups (*e.g.*, at least one head-group and at least one tail-group). In certain embodiments, at least one of the functional moieties or groups is hydrophilic (*e.g.*, a hydrophilic head-group comprising one or more of imidazole, guanidinium, amino, imine, enamine, optionally-substituted alkyl amino and pyridyl).

[0203] As used herein, the term “hydrophilic” is used to indicate in qualitative terms that a functional group is water-preferring, and typically such groups are water-soluble. For example, disclosed herein are compounds that comprise a cleavable disulfide (S—S) functional group bound to one or more hydrophilic groups (*e.g.*, a hydrophilic head-group), wherein such hydrophilic groups comprise or are selected from the group consisting of imidazole, guanidinium, amino, imine, enamine, an optionally-substituted alkyl amino (*e.g.*, an alkyl amino such as dimethylamino) and pyridyl.

[0204] In certain embodiments, at least one of the functional groups of moieties that comprise the compounds disclosed herein is hydrophobic in nature (*e.g.*, a hydrophobic tail-group comprising a naturally occurring lipid such as cholesterol). As used herein, the term “hydrophobic” is used to indicate in qualitative terms that a functional group is water-avoiding, and typically such groups are not water soluble. For example, disclosed herein are compounds that comprise a cleavable functional group (*e.g.*, a disulfide (S—S) group) bound to one or more hydrophobic groups, wherein such hydrophobic groups comprise one or more naturally occurring lipids such as cholesterol, and/or an optionally substituted, variably saturated or unsaturated C₆-C₂₀ alkyl and/or an optionally substituted, variably saturated or unsaturated C₆-C₂₀ acyl.

[0205] Compound described herein may also comprise one or more isotopic substitutions. For example, H may be in any isotopic form, including ¹H, ²H (D or deuterium), and ³H (T or tritium); C may be in any isotopic form, including ¹²C, ¹³C, and ¹⁴C; O may be in any isotopic form, including ¹⁶O and ¹⁸O; F may be in any isotopic form, including ¹⁸F and ¹⁹F; and the like.

[0206] When describing the invention, which may include compounds and pharmaceutically acceptable salts thereof, pharmaceutical compositions containing such compounds and methods of using such compounds and compositions, the following terms, if present, have the following meanings unless otherwise indicated. It should also be understood that when described herein any of the moieties defined forth below may be substituted with a variety of substituents, and that the respective definitions are intended to include such substituted moieties within their scope as set out below. Unless otherwise stated, the term “substituted” is to be defined as set out below. It should be further understood that the terms “groups” and “radicals” can be considered interchangeable when used herein.

[0207] When a range of values is listed, it is intended to encompass each value and sub-range within the range. For example, “C₁₋₆ alkyl” is intended to encompass, C₁, C₂, C₃, C₄, C₅, C₆, C₁₋₆, C₁₋₅, C₁₋₄, C₁₋₃, C₁₋₂, C₂₋₆, C₂₋₅, C₂₋₄, C₂₋₃, C₃₋₆, C₃₋₅, C₃₋₄, C₄₋₆, C₄₋₅, and C₅₋₆ alkyl.

[0208] In certain embodiments, the compounds disclosed herein comprise, for example, at least one hydrophilic head-group and at least one hydrophobic tail-group, each bound to at least one cleavable group, thereby rendering such compounds amphiphilic. As used herein to describe a compound or composition, the term “amphiphilic” means the ability to dissolve in both polar (e.g., water) and non-polar (e.g., lipid) environments. For example, in certain embodiments, the compounds disclosed herein comprise at least one lipophilic tail-group (e.g., cholesterol or a C₆-C₂₀ alkyl) and at least one hydrophilic head-group (e.g., imidazole), each bound to a cleavable group (e.g., disulfide).

[0209] It should be noted that the terms “head-group” and “tail-group” as used describe the compounds of the present invention, and in particular functional groups that comprise such compounds, are used for ease of reference to describe the orientation of one or more functional groups relative to other functional groups. For example, in certain embodiments a hydrophilic head-group (e.g., guanidinium) is bound (e.g., by one or more of hydrogen-bonds, van der Waals' forces, ionic interactions and covalent bonds) to a cleavable functional group (e.g., a disulfide group), which in turn is bound to a hydrophobic tail-group (e.g., cholesterol).

[0210] As used herein, the term “alkyl” refers to both straight and branched chain C₁-C₄₀ hydrocarbons (e.g., C₆-C₂₀ hydrocarbons), and include both saturated and unsaturated hydrocarbons. In certain embodiments, the alkyl may comprise one or more cyclic alkyls and/or one or more heteroatoms such as oxygen, nitrogen, or sulfur and may optionally be substituted with substituents (e.g., one or more of alkyl, halo, alkoxyl, hydroxy, amino, aryl,

ether, ester or amide). In certain embodiments, a contemplated alkyl includes (9Z,12Z)-octadeca-9,12-dien. The use of designations such as, for example, “C₆-C₂₀” is intended to refer to an alkyl (e.g., straight or branched chain and inclusive of alkenes and alkyls) having the recited range carbon atoms. In some embodiments, an alkyl group has 1 to 10 carbon atoms (“C₁₋₁₀ alkyl”). In some embodiments, an alkyl group has 1 to 9 carbon atoms (“C₁₋₉ alkyl”). In some embodiments, an alkyl group has 1 to 8 carbon atoms (“C₁₋₈ alkyl”). In some embodiments, an alkyl group has 1 to 7 carbon atoms (“C₁₋₇ alkyl”). In some embodiments, an alkyl group has 1 to 6 carbon atoms (“C₁₋₆ alkyl”). In some embodiments, an alkyl group has 1 to 5 carbon atoms (“C₁₋₅ alkyl”). In some embodiments, an alkyl group has 1 to 4 carbon atoms (“C₁₋₄ alkyl”). In some embodiments, an alkyl group has 1 to 3 carbon atoms (“C₁₋₃ alkyl”). In some embodiments, an alkyl group has 1 to 2 carbon atoms (“C₁₋₂ alkyl”). In some embodiments, an alkyl group has 1 carbon atom (“C₁ alkyl”). Examples of C₁₋₆ alkyl groups include methyl, ethyl, propyl, isopropyl, butyl, isobutyl, pentyl, hexyl, and the like.

[0211] As used herein, “alkenyl” refers to a radical of a straight-chain or branched hydrocarbon group having from 2 to 20 carbon atoms, one or more carbon-carbon double bonds (e.g., 1, 2, 3, or 4 carbon-carbon double bonds), and optionally one or more carbon-carbon triple bonds (e.g., 1, 2, 3, or 4 carbon-carbon triple bonds) (“C₂₋₂₀ alkenyl”). In certain embodiments, alkenyl does not contain any triple bonds. In some embodiments, an alkenyl group has 2 to 10 carbon atoms (“C₂₋₁₀ alkenyl”). In some embodiments, an alkenyl group has 2 to 9 carbon atoms (“C₂₋₉ alkenyl”). In some embodiments, an alkenyl group has 2 to 8 carbon atoms (“C₂₋₈ alkenyl”). In some embodiments, an alkenyl group has 2 to 7 carbon atoms (“C₂₋₇ alkenyl”). In some embodiments, an alkenyl group has 2 to 6 carbon atoms (“C₂₋₆ alkenyl”). In some embodiments, an alkenyl group has 2 to 5 carbon atoms (“C₂₋₅ alkenyl”). In some embodiments, an alkenyl group has 2 to 4 carbon atoms (“C₂₋₄ alkenyl”). In some embodiments, an alkenyl group has 2 to 3 carbon atoms (“C₂₋₃ alkenyl”). In some embodiments, an alkenyl group has 2 carbon atoms (“C₂ alkenyl”). The one or more carbon-carbon double bonds can be internal (such as in 2-butenyl) or terminal (such as in 1-butenyl). Examples of C₂₋₄ alkenyl groups include ethenyl (C₂), 1-propenyl (C₃), 2-propenyl (C₃), 1-butenyl (C₄), 2-butenyl (C₄), butadienyl (C₄), and the like. Examples of C₂₋₆ alkenyl groups include the aforementioned C₂₋₄ alkenyl groups as well as pentenyl (C₅), pentadienyl (C₅), hexenyl (C₆), and the like. Additional examples of alkenyl include heptenyl (C₇), octenyl (C₈), octatrienyl (C₈), and the like.

[0212] As used herein, “alkynyl” refers to a radical of a straight-chain or branched

hydrocarbon group having from 2 to 20 carbon atoms, one or more carbon-carbon triple bonds (e.g., 1, 2, 3, or 4 carbon-carbon triple bonds), and optionally one or more carbon-carbon double bonds (e.g., 1, 2, 3, or 4 carbon-carbon double bonds) (“C₂₋₂₀ alkynyl”). In certain embodiments, alkynyl does not contain any double bonds. In some embodiments, an alkynyl group has 2 to 10 carbon atoms (“C₂₋₁₀ alkynyl”). In some embodiments, an alkynyl group has 2 to 9 carbon atoms (“C₂₋₉ alkynyl”). In some embodiments, an alkynyl group has 2 to 8 carbon atoms (“C₂₋₈ alkynyl”). In some embodiments, an alkynyl group has 2 to 7 carbon atoms (“C₂₋₇ alkynyl”). In some embodiments, an alkynyl group has 2 to 6 carbon atoms (“C₂₋₆ alkynyl”). In some embodiments, an alkynyl group has 2 to 5 carbon atoms (“C₂₋₅ alkynyl”). In some embodiments, an alkynyl group has 2 to 4 carbon atoms (“C₂₋₄ alkynyl”). In some embodiments, an alkynyl group has 2 to 3 carbon atoms (“C₂₋₃ alkynyl”). In some embodiments, an alkynyl group has 2 carbon atoms (“C₂ alkynyl”). The one or more carbon-carbon triple bonds can be internal (such as in 2-butyne) or terminal (such as in 1-butyne). Examples of C₂₋₄ alkynyl groups include, without limitation, ethynyl (C₂), 1-propynyl (C₃), 2-propynyl (C₃), 1-butyne (C₄), 2-butyne (C₄), and the like. Examples of C₂₋₆ alkenyl groups include the aforementioned C₂₋₄ alkynyl groups as well as pentynyl (C₅), hexynyl (C₆), and the like. Additional examples of alkynyl include heptynyl (C₇), octynyl (C₈), and the like.

[0213] As used herein, “alkylene,” “alkenylene,” and “alkynylene,” refer to a divalent radical of an alkyl, alkenyl, and alkynyl group respectively. When a range or number of carbons is provided for a particular “alkylene,” “alkenylene,” or “alkynylene,” group, it is understood that the range or number refers to the range or number of carbons in the linear carbon divalent chain. “Alkylene,” “alkenylene,” and “alkynylene,” groups may be substituted or unsubstituted with one or more substituents as described herein.

[0214] As used herein, the term “aryl” refers to aromatic groups (e.g., monocyclic, bicyclic and tricyclic structures) containing six to ten carbons in the ring portion. The aryl groups may be optionally substituted through available carbon atoms and in certain embodiments may include one or more heteroatoms such as oxygen, nitrogen or sulfur. In some embodiments, an aryl group has six ring carbon atoms (“C₆ aryl”; e.g., phenyl). In some embodiments, an aryl group has ten ring carbon atoms (“C₁₀ aryl”; e.g., naphthyl such as 1-naphthyl and 2-naphthyl).

[0215] As used herein, “heteroaryl” refers to a radical of a 5–10 membered monocyclic or bicyclic 4n+2 aromatic ring system (e.g., having 6 or 10 electrons shared in a cyclic array) having ring carbon atoms and 1–4 ring heteroatoms provided in the aromatic ring system,

wherein each heteroatom is independently selected from nitrogen, oxygen and sulfur ("5–10 membered heteroaryl"). In heteroaryl groups that contain one or more nitrogen atoms, the point of attachment can be a carbon or nitrogen atom, as valency permits. Heteroaryl bicyclic ring systems can include one or more heteroatoms in one or both rings. "Heteroaryl" includes ring systems wherein the heteroaryl ring, as defined above, is fused with one or more carbocyclyl or heterocyclyl groups wherein the point of attachment is on the heteroaryl ring, and in such instances, the number of ring members continue to designate the number of ring members in the heteroaryl ring system. "Heteroaryl" also includes ring systems wherein the heteroaryl ring, as defined above, is fused with one or more aryl groups wherein the point of attachment is either on the aryl or heteroaryl ring, and in such instances, the number of ring members designates the number of ring members in the fused (aryl/heteroaryl) ring system. Bicyclic heteroaryl groups wherein one ring does not contain a heteroatom (e.g., indolyl, quinolinyl, carbazolyl, and the like) the point of attachment can be on either ring, i.e., either the ring bearing a heteroatom (e.g., 2-indolyl) or the ring that does not contain a heteroatom (e.g., 5-indolyl).

[0216] The term "cycloalkyl" refers to a monovalent saturated cyclic, bicyclic, or bridged cyclic (e.g., adamantyl) hydrocarbon group of 3-12, 3-8, 4-8, or 4-6 carbons, referred to herein, e.g., as "C₄₋₈cycloalkyl," derived from a cycloalkane. Exemplary cycloalkyl groups include, but are not limited to, cyclohexanes, cyclopentanes, cyclobutanes and cyclopropanes.

[0217] As used herein, "heterocyclyl" or "heterocyclic" refers to a radical of a 3– to 10–membered non-aromatic ring system having ring carbon atoms and 1 to 4 ring heteroatoms, wherein each heteroatom is independently selected from nitrogen, oxygen, sulfur, boron, phosphorus, and silicon ("3–10 membered heterocyclyl"). In heterocyclyl groups that contain one or more nitrogen atoms, the point of attachment can be a carbon or nitrogen atom, as valency permits. A heterocyclyl group can either be monocyclic ("monocyclic heterocyclyl") or a fused, bridged or spiro ring system such as a bicyclic system ("bicyclic heterocyclyl"), and can be saturated or can be partially unsaturated. Heterocyclyl bicyclic ring systems can include one or more heteroatoms in one or both rings. "Heterocyclyl" also includes ring systems wherein the heterocyclyl ring, as defined above, is fused with one or more carbocyclyl groups wherein the point of attachment is either on the carbocyclyl or heterocyclyl ring, or ring systems wherein the heterocyclyl ring, as defined above, is fused with one or more aryl or heteroaryl groups, wherein the point of attachment is on the heterocyclyl ring, and in such instances, the number of ring members continue to designate the number of ring members in the heterocyclyl ring system. The terms "heterocycle,"

“heterocyclyl,” “heterocyclyl ring,” “heterocyclic group,” “heterocyclic moiety,” and “heterocyclic radical,” may be used interchangeably.

[0218] As used herein, “cyano” refers to -CN.

[0219] The terms “halo” and “halogen” as used herein refer to an atom selected from fluorine (fluoro, F), chlorine (chloro, Cl), bromine (bromo, Br), and iodine (iodo, I). In certain embodiments, the halo group is either fluoro or chloro.

[0220] The term “alkoxy,” as used herein, refers to an alkyl group which is attached to another moiety via an oxygen atom (-O(alkyl)). Non-limiting examples include e.g., methoxy, ethoxy, propoxy, and butoxy.

[0221] As used herein, “oxo” refers to -C=O.

[0222] In general, the term “substituted”, whether preceded by the term “optionally” or not, means that at least one hydrogen present on a group (e.g., a carbon or nitrogen atom) is replaced with a permissible substituent, e.g., a substituent which upon substitution results in a stable compound, e.g., a compound which does not spontaneously undergo transformation such as by rearrangement, cyclization, elimination, or other reaction. Unless otherwise indicated, a “substituted” group has a substituent at one or more substitutable positions of the group, and when more than one position in any given structure is substituted, the substituent is either the same or different at each position.

[0223] As used herein, “pharmaceutically acceptable salt” refers to those salts which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of humans and lower animals without undue toxicity, irritation, allergic response and the like, and are commensurate with a reasonable benefit/risk ratio. Pharmaceutically acceptable salts are well known in the art. For example, Berge *et al.*, describes pharmaceutically acceptable salts in detail in *J. Pharmaceutical Sciences* (1977) 66:1–19. Pharmaceutically acceptable salts of the compounds of this invention include those derived from suitable inorganic and organic acids and bases. Examples of pharmaceutically acceptable, nontoxic acid addition salts are salts of an amino group formed with inorganic acids such as hydrochloric acid, hydrobromic acid, phosphoric acid, sulfuric acid and perchloric acid or with organic acids such as acetic acid, oxalic acid, maleic acid, tartaric acid, citric acid, succinic acid or malonic acid or by using other methods used in the art such as ion exchange. Other pharmaceutically acceptable salts include adipate, alginate, ascorbate, aspartate, benzenesulfonate, benzoate, bisulfate, borate, butyrate, camphorate, camphorsulfonate, citrate, cyclopentanepropionate, digluconate, dodecylsulfate, ethanesulfonate, formate, fumarate, glucoheptonate, glycerophosphate, gluconate, hemisulfate, heptanoate, hexanoate, hydroiodide, 2-hydroxy-

ethanesulfonate, lactobionate, lactate, laurate, lauryl sulfate, malate, maleate, malonate, methanesulfonate, 2-naphthalenesulfonate, nicotinate, nitrate, oleate, oxalate, palmitate, pamoate, pectinate, persulfate, 3-phenylpropionate, phosphate, picrate, pivalate, propionate, stearate, succinate, sulfate, tartrate, thiocyanate, p-toluenesulfonate, undecanoate, valerate salts, and the like. Pharmaceutically acceptable salts derived from appropriate bases include alkali metal, alkaline earth metal, ammonium and $N^+(C_{1-4}alkyl)_4$ salts. Representative alkali or alkaline earth metal salts include sodium, lithium, potassium, calcium, magnesium, and the like. Further pharmaceutically acceptable salts include, when appropriate, nontoxic ammonium, quaternary ammonium, and amine cations formed using counterions such as halide, hydroxide, carboxylate, sulfate, phosphate, nitrate, lower alkyl sulfonate, and aryl sulfonate.

[0224] In typical embodiments, the present invention is intended to encompass the compounds disclosed herein, and the pharmaceutically acceptable salts, pharmaceutically acceptable esters, tautomeric forms, polymorphs, and prodrugs of such compounds. In some embodiments, the present invention includes a pharmaceutically acceptable addition salt, a pharmaceutically acceptable ester, a solvate (e.g., hydrate) of an addition salt, a tautomeric form, a polymorph, an enantiomer, a mixture of enantiomers, a stereoisomer or mixture of stereoisomers (pure or as a racemic or non-racemic mixture) of a compound described herein.

[0225] Compounds described herein can comprise one or more asymmetric centers, and thus can exist in various isomeric forms, e.g., enantiomers and/or diastereomers. For example, the compounds described herein can be in the form of an individual enantiomer, diastereomer or geometric isomer, or can be in the form of a mixture of stereoisomers, including racemic mixtures and mixtures enriched in one or more stereoisomer. Isomers can be isolated from mixtures by methods known to those skilled in the art, including chiral high pressure liquid chromatography (HPLC) and the formation and crystallization of chiral salts; or preferred isomers can be prepared by asymmetric syntheses. See, for example, Jacques *et al.*, *Enantiomers, Racemates and Resolutions* (Wiley Interscience, New York, 1981); Wilen *et al.*, *Tetrahedron* 33:2725 (1977); Eliel, *Stereochemistry of Carbon Compounds* (McGraw-Hill, NY, 1962); and Wilen, *Tables of Resolving Agents and Optical Resolutions* p. 268 (E.L. Eliel, Ed., Univ. of Notre Dame Press, Notre Dame, IN 1972). The invention additionally encompasses compounds described herein as individual isomers substantially free of other isomers, and alternatively, as mixtures of various isomers.

[0226] In certain embodiments the compounds and the transfer vehicles of which such compounds are a component (e.g., lipid nanoparticles) exhibit an enhanced (e.g., increased)

ability to transfect one or more target cells. Accordingly, also provided herein are methods of transfecting one or more target cells. Such methods generally comprise the step of contacting the one or more target cells with the compounds and/or pharmaceutical compositions disclosed herein such that the one or more target cells are transfected with the circular RNA encapsulated therein. As used herein, the terms “transfect” or “transfection” refer to the intracellular introduction of one or more encapsulated materials (e.g., nucleic acids and/or polynucleotides) into a cell, or preferably into a target cell. The term “transfection efficiency” refers to the relative amount of such encapsulated material (e.g., polynucleotides) up-taken by, introduced into and/or expressed by the target cell which is subject to transfection. In some embodiments, transfection efficiency may be estimated by the amount of a reporter polynucleotide product produced by the target cells following transfection. In some embodiments, a transfer vehicle has high transfection efficiency. In some embodiments, a transfer vehicle has at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% transfection efficiency.

[0227] As used herein, the term “liposome” generally refers to a vesicle composed of lipids (e.g., amphiphilic lipids) arranged in one or more spherical bilayer or bilayers. In certain embodiments, the liposome is a lipid nanoparticle (e.g., a lipid nanoparticle comprising one or more of the ionizable lipid compounds disclosed herein). Such liposomes may be unilamellar or multilamellar vesicles which have a membrane formed from a lipophilic material and an aqueous interior that contains the encapsulated circRNA to be delivered to one or more target cells, tissues and organs. In certain embodiments, the compositions described herein comprise one or more lipid nanoparticles. Examples of suitable lipids (e.g., ionizable lipids) that may be used to form the liposomes and lipid nanoparticles contemplated include one or more of the compounds disclosed herein (e.g., HGT4001, HGT4002, HGT4003, HGT4004 and/or HGT4005). Such liposomes and lipid nanoparticles may also comprise additional ionizable lipids such as C12-200, DLin-KC2-DMA, and/or HGT5001, helper lipids, structural lipids, PEG-modified lipids, MC3, DLinDMA, DLinkC2DMA, cKK-E12, ICE, HGT5000, DODAC, DDAB, DMRIE, DOSPA, DOGS, DODAP, DODMA, DMDMA, DODAC, DLenDMA, DMRIE, CLinDMA, CpLinDMA, DMOBA, DOcarbDAP, DLinDAP, DLincarbDAP, DLinCDAP, KLin-K-DMA, DLin-K-XTC2-DMA, HGT4003, and combinations thereof.

[0228] As used herein, the phrases “non-cationic lipid”, “non-cationic helper lipid”, and “helper lipid” are used interchangeably and refer to any neutral, zwitterionic or anionic lipid.

[0229] As used herein, the phrase “anionic lipid” refers to any of a number of lipid

species that carry a net negative charge at a selected pH, such as physiological pH.

[0230] As used herein, the phrase “biodegradable lipid” or “degradable lipid” refers to any of a number of lipid species that are broken down in a host environment on the order of minutes, hours, or days ideally making them less toxic and unlikely to accumulate in a host over time. Common modifications to lipids include ester bonds, and disulfide bonds among others to increase the biodegradability of a lipid.

[0231] As used herein, the phrase “biodegradable PEG lipid” or “degradable PEG lipid” refers to any of a number of lipid species where the PEG molecules are cleaved from the lipid in a host environment on the order of minutes, hours, or days ideally making them less immunogenic. Common modifications to PEG lipids include ester bonds, and disulfide bonds among others to increase the biodegradability of a lipid.

[0232] In certain embodiments of the present invention, the transfer vehicles (e.g., lipid nanoparticles) are prepared to encapsulate one or more materials or therapeutic agents (e.g., circRNA). The process of incorporating a desired therapeutic agent (e.g., circRNA) into a transfer vehicle is referred to herein as or “loading” or “encapsulating” (Lasic, *et al.*, FEBS Lett., 312: 255-258, 1992). The transfer vehicle-loaded or -encapsulated materials (e.g., circRNA) may be completely or partially located in the interior space of the transfer vehicle, within a bilayer membrane of the transfer vehicle, or associated with the exterior surface of the transfer vehicle.

[0233] As used herein, the term “structural lipid” refers to sterols and also to lipids containing sterol moieties.

[0234] As defined herein, “sterols” are a subgroup of steroids consisting of steroid alcohols.

[0235] As used herein, the term “structural lipid” refers to sterols and also to lipids containing sterol moieties.

[0236] As used herein, the term “PEG” means any polyethylene glycol or other polyalkylene ether polymer.

[0237] As generally defined herein, a “PEG-OH lipid” (also referred to herein as “hydroxy-PEGylated lipid”) is a PEGylated lipid having one or more hydroxyl (–OH) groups on the lipid.

[0238] As used herein, a “phospholipid” is a lipid that includes a phosphate moiety and one or more carbon chains, such as unsaturated fatty acid chains.

[0239] All nucleotide sequences disclosed herein can represent an RNA sequence or a corresponding DNA sequence. It is understood that deoxythymidine (dT or T) in a DNA is

transcribed into a uridine (U) in an RNA. As such, “T” and “U” are used interchangeably herein in nucleotide sequences.

[0240] The recitations “sequence identity” or, for example, comprising a “sequence 50% identical to,” as used herein, refer to the extent that sequences are identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a “percentage of sequence identity” may be calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, I) or the identical amino acid residue (*e.g.*, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. Included are nucleotides and polypeptides having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to any of the reference sequences described herein, typically where the polypeptide variant maintains at least one biological activity of the reference polypeptide.

[0241] The term “antibody” (Ab) includes, without limitation, a glycoprotein immunoglobulin which binds specifically to an antigen. In general, an antibody may comprise at least two heavy (H) chains and two light (L) chains interconnected by disulfide bonds, or an antigen-binding molecule thereof. Each H chain may comprise a heavy chain variable region (abbreviated herein as VH) and a heavy chain constant region. The heavy chain constant region can comprise three constant domains, CH1, CH2 and CH3. Each light chain can comprise a light chain variable region (abbreviated herein as VL) and a light chain constant region. The light chain constant region can comprise one constant domain, CL. The VH and VL regions may be further subdivided into regions of hypervariability, termed complementarity determining regions (CDRs), interspersed with regions that are more conserved, termed framework regions (FR). Each VH and VL may comprise three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, and FR4. The variable regions of the heavy and light chains contain a binding domain that interacts with an antigen. The constant regions of the Abs may mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system (*e.g.*, effector cells) and the first component of the classical complement system. Antibodies may include, for example, monoclonal antibodies, recombinantly produced antibodies, monospecific antibodies, multispecific antibodies

(including bispecific antibodies), human antibodies, engineered antibodies, humanized antibodies, chimeric antibodies, immunoglobulins, synthetic antibodies, tetrameric antibodies comprising two heavy chain and two light chain molecules, an antibody light chain monomer, an antibody heavy chain monomer, an antibody light chain dimer, an antibody heavy chain dimer, an antibody light chain- antibody heavy chain pair, intrabodies, antibody fusions (sometimes referred to herein as “antibody conjugates”), heteroconjugate antibodies, single domain antibodies, monovalent antibodies, single chain antibodies or single-chain variable fragments (scFv), camelized antibodies, affibodies, Fab fragments, F(ab')₂ fragments, disulfide-linked variable fragments (sdFv), anti-idiotypic (anti-id) antibodies (including, *e.g.*, anti-anti-Id antibodies), minibodies, domain antibodies, synthetic antibodies (sometimes referred to herein as “antibody mimetics”), and antigen-binding fragments of any of the above. In some embodiments, antibodies described herein refer to polyclonal antibody populations.

[0242] An immunoglobulin may derive from any of the commonly known isotypes, including but not limited to IgA, secretory IgA, IgG and IgM. IgG subclasses are also well known to those in the art and include but are not limited to human IgG1, IgG2, IgG3 and IgG4. “Isotype” refers to the Ab class or subclass (*e.g.*, IgM or IgG1) that is encoded by the heavy chain constant region genes. The term “antibody” includes, by way of example, both naturally occurring and non-naturally occurring Abs; monoclonal and polyclonal Abs; chimeric and humanized Abs; human or nonhuman Abs; wholly synthetic Abs; and single chain Abs. A nonhuman Ab may be humanized by recombinant methods to reduce its immunogenicity in humans. Where not expressly stated, and unless the context indicates otherwise, the term “antibody” also includes an antigen-binding fragment or an antigen-binding portion of any of the aforementioned immunoglobulins, and includes a monovalent and a divalent fragment or portion, and a single chain Ab.

[0243] An “antigen binding molecule,” “antigen binding portion,” or “antibody fragment” refers to any molecule that comprises the antigen binding parts (*e.g.*, CDRs) of the antibody from which the molecule is derived. An antigen binding molecule may include the antigenic complementarity determining regions (CDRs). Examples of antibody fragments include, but are not limited to, Fab, Fab', F(ab')₂, Fv fragments, dAb, linear antibodies, scFv antibodies, and multispecific antibodies formed from antigen binding molecules. Peptibodies (*i.e.* Fc fusion molecules comprising peptide binding domains) are another example of suitable antigen binding molecules. In some embodiments, the antigen binding molecule binds to an antigen on a tumor cell. In some embodiments, the antigen binding molecule

binds to an antigen on a cell involved in a hyperproliferative disease or to a viral or bacterial antigen. In some embodiments, the antigen binding molecule binds to BCMA. In further embodiments, the antigen binding molecule is an antibody fragment, including one or more of the complementarity determining regions (CDRs) thereof, that specifically binds to the antigen. In further embodiments, the antigen binding molecule is a single chain variable fragment (scFv). In some embodiments, the antigen binding molecule comprises or consists of avimers.

[0244] As used herein, the term “variable region” or “variable domain” is used interchangeably and are common in the art. The variable region typically refers to a portion of an antibody, generally, a portion of a light or heavy chain, typically about the amino-terminal 110 to 120 amino acids in the mature heavy chain and about 90 to 115 amino acids in the mature light chain, which differ extensively in sequence among antibodies and are used in the binding and specificity of a particular antibody for its particular antigen. The variability in sequence is concentrated in those regions called complementarity determining regions (CDRs) while the more highly conserved regions in the variable domain are called framework regions (FR). Without wishing to be bound by any particular mechanism or theory, it is believed that the CDRs of the light and heavy chains are primarily responsible for the interaction and specificity of the antibody with antigen. In some embodiments, the variable region is a human variable region. In some embodiments, the variable region comprises rodent or murine CDRs and human framework regions (FRs). In particular embodiments, the variable region is a primate (*e.g.*, non-human primate) variable region. In some embodiments, the variable region comprises rodent or murine CDRs and primate (*e.g.*, non-human primate) framework regions (FRs).

[0245] The terms “VL” and “VL domain” are used interchangeably to refer to the light chain variable region of an antibody or an antigen-binding molecule thereof.

[0246] The terms “VH” and “VH domain” are used interchangeably to refer to the heavy chain variable region of an antibody or an antigen-binding molecule thereof.

[0247] A number of definitions of the CDRs are commonly in use: Kabat numbering, Chothia numbering, AbM numbering, or contact numbering. The AbM definition is a compromise between the two used by Oxford Molecular’s AbM antibody modelling software. The contact definition is based on an analysis of the available complex crystal structures. The term “Kabat numbering” and like terms are recognized in the art and refer to a system of numbering amino acid residues in the heavy and light chain variable regions of an antibody, or an antigen-binding molecule thereof. In certain aspects, the CDRs of an antibody

may be determined according to the Kabat numbering system (*see, e.g.*, Kabat EA & Wu TT (1971) Ann NY Acad Sci 190: 382-391 and Kabat EA *et al.*, (1991) Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242). Using the Kabat numbering system, CDRs within an antibody heavy chain molecule are typically present at amino acid positions 31 to 35, which optionally may include one or two additional amino acids, following 35 (referred to in the Kabat numbering scheme as 35A and 35B) (CDR1), amino acid positions 50 to 65 (CDR2), and amino acid positions 95 to 102 (CDR3). Using the Kabat numbering system, CDRs within an antibody light chain molecule are typically present at amino acid positions 24 to 34 (CDR1), amino acid positions 50 to 56 (CDR2), and amino acid positions 89 to 97 (CDR3). In a specific embodiment, the CDRs of the antibodies described herein have been determined according to the Kabat numbering scheme. In certain aspects, the CDRs of an antibody may be determined according to the Chothia numbering scheme, which refers to the location of immunoglobulin structural loops (*see, e.g.*, Chothia C & Lesk AM, (1987), J Mol Biol 196: 901-917; Al-Lazikani B *et al.*, (1997) J Mol Biol 273: 927-948; Chothia C *et al.*, (1992) J Mol Biol 227: 799-817; Tramontano A *et al.*, (1990) J Mol Biol 215(1): 175- 82; and U.S. Patent No. 7,709,226). Typically, when using the Kabat numbering convention, the Chothia CDR-H1 loop is present at heavy chain amino acids 26 to 32, 33, or 34, the Chothia CDR-H2 loop is present at heavy chain amino acids 52 to 56, and the Chothia CDR-H3 loop is present at heavy chain amino acids 95 to 102, while the Chothia CDR-L1 loop is present at light chain amino acids 24 to 34, the Chothia CDR-L2 loop is present at light chain amino acids 50 to 56, and the Chothia CDR-L3 loop is present at light chain amino acids 89 to 97. The end of the Chothia CDR-H1 loop when numbered using the Kabat numbering convention varies between H32 and H34 depending on the length of the loop (this is because the Kabat numbering scheme places the insertions at H35A and H35B; if neither 35A nor 35B is present, the loop ends at 32; if only 35A is present, the loop ends at 33; if both 35A and 35B are present, the loop ends at 34). In a specific embodiment, the CDRs of the antibodies described herein have been determined according to the Chothia numbering scheme.

[0248] As used herein, the terms “constant region” and “constant domain” are interchangeable and have a meaning common in the art. The constant region is an antibody portion, *e.g.*, a carboxyl terminal portion of a light and/or heavy chain which is not directly involved in binding of an antibody to antigen but which may exhibit various effector functions, such as interaction with the Fc receptor. The constant region of an immunoglobulin molecule generally has a more conserved amino acid sequence relative to an immunoglobulin

variable domain.

[0249] “Binding affinity” generally refers to the strength of the sum total of non-covalent interactions between a single binding site of a molecule (*e.g.*, an antibody) and its binding partner (*e.g.*, an antigen). Unless indicated otherwise, as used herein, “binding affinity” refers to intrinsic binding affinity which reflects a 1 : 1 interaction between members of a binding pair (*e.g.*, antibody and antigen). The affinity of a molecule X for its partner Y may generally be represented by the dissociation constant (KD or Kd). Affinity may be measured and/or expressed in a number of ways known in the art, including, but not limited to, equilibrium dissociation constant (KD), and equilibrium association constant (KA or Ka). The KD is calculated from the quotient of koff/kon, whereas KA is calculated from the quotient of kon/koff. kon refers to the association rate constant of, *e.g.*, an antibody to an antigen, and koff refers to the dissociation of, *e.g.*, an antibody to an antigen. The kon and koff may be determined by techniques known to one of ordinary skill in the art, such as BIAcore® or KinExA.

[0250] As used herein, a “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine, tryptophan), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). In some embodiments, one or more amino acid residues within a CDR(s) or within a framework region(s) of an antibody or antigen-binding molecule thereof may be replaced with an amino acid residue with a similar side chain.

[0251] As, used herein, the term “heterologous” means from any source other than naturally occurring sequences.

[0252] As used herein, an “epitope” is a term in the art and refers to a localized region of an antigen to which an antibody may specifically bind. An epitope may be, for example, contiguous amino acids of a polypeptide (linear or contiguous epitope) or an epitope can, for example, come together from two or more non-contiguous regions of a polypeptide or polypeptides (conformational, non-linear, discontinuous, or non-contiguous epitope). In some embodiments, the epitope to which an antibody binds may be determined by, *e.g.*, NMR spectroscopy, X-ray diffraction crystallography studies, ELISA assays, hydrogen/deuterium

exchange coupled with mass spectrometry (*e.g.*, liquid chromatography electrospray mass spectrometry), array -based oligo-peptide scanning assays, and/or mutagenesis mapping (*e.g.*, site- directed mutagenesis mapping). For X-ray crystallography, crystallization may be accomplished using any of the known methods in the art (*e.g.*, Giege R *et al.*, (1994) *Acta Crystallogr D Biol Crystallogr* 50(Pt 4): 339-350; McPherson A (1990) *Eur J Biochem* 189: 1-23; Chayen NE (1997) *Structure* 5: 1269- 1274; McPherson A (1976) *J Biol Chem* 251: 6300-6303). Antibody: antigen crystals may be studied using well known X-ray diffraction techniques and may be refined using computer software such as X- PLOR (Yale University, 1992, distributed by Molecular Simulations, Inc.; see *e.g.* *Meth Enzymol* (1985) volumes 114 & 115, eds Wyckoff HW *et al.*; U.S. Patent Publication No. 2004/0014194), and BUSTER (Bricogne G (1993) *Acta Crystallogr D Biol Crystallogr* 49(Pt 1): 37-60; Bricogne G (1997) *Meth Enzymol* 276A: 361-423, ed Carter CW; Roversi P *et al.*, (2000) *Acta Crystallogr D Biol Crystallogr* 56(Pt 10): 1316-1323).

[0253] As used herein, an antigen binding molecule, an antibody, or an antigen binding molecule thereof “cross-competes” with a reference antibody or an antigen binding molecule thereof if the interaction between an antigen and the first binding molecule, an antibody, or an antigen binding molecule thereof blocks, limits, inhibits, or otherwise reduces the ability of the reference binding molecule, reference antibody, or an antigen binding molecule thereof to interact with the antigen. Cross competition may be complete, *e.g.*, binding of the binding molecule to the antigen completely blocks the ability of the reference binding molecule to bind the antigen, or it may be partial, *e.g.*, binding of the binding molecule to the antigen reduces the ability of the reference binding molecule to bind the antigen. In some embodiments, an antigen binding molecule that cross-competes with a reference antigen binding molecule binds the same or an overlapping epitope as the reference antigen binding molecule. In other embodiments, the antigen binding molecule that cross-competes with a reference antigen binding molecule binds a different epitope as the reference antigen binding molecule. Numerous types of competitive binding assays may be used to determine if one antigen binding molecule competes with another, for example: solid phase direct or indirect radioimmunoassay (RIA); solid phase direct or indirect enzyme immunoassay (EIA), sandwich competition assay (Stahli *et al.*, 1983, *Methods in Enzymology* 9:242-253); solid phase direct biotin-avidin EIA (Kirkland *et al.*, 1986, *J. Immunol.* 137:3614-3619); solid phase direct labeled assay, solid phase direct labeled sandwich assay (Harlow and Lane, 1988, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Press); solid phase direct label RIA using 1-125 label (Morel *et al.*, 1988, *Molec. Immunol.* 25:7-15); solid phase direct

biotin-avidin EIA (Cheung, *et al.*, 1990, Virology 176:546-552); and direct labeled RIA (Moldenhauer *et al.*, 1990, Scand. J. Immunol. 32:77-82).

[0254] As used herein, the terms “immunospecifically binds,” “immunospecifically recognizes,” “specifically binds,” and “specifically recognizes” are analogous terms in the context of antibodies and refer to molecules that bind to an antigen (*e.g.*, epitope or immune complex) as such binding is understood by one skilled in the art. For example, a molecule that specifically binds to an antigen may bind to other peptides or polypeptides, generally with lower affinity as determined by, *e.g.*, immunoassays, BIAcore®, KinExA 3000 instrument (Sapidyne Instruments, Boise, ID), or other assays known in the art. In a specific embodiment, molecules that specifically bind to an antigen bind to the antigen with a K_A that is at least 2 logs, 2.5 logs, 3 logs, 4 logs or greater than the K_A when the molecules bind to another antigen.

[0255] An “antigen” refers to any molecule that provokes an immune response or is capable of being bound by an antibody or an antigen binding molecule. The immune response may involve either antibody production, or the activation of specific immunologically - competent cells, or both. A person of skill in the art would readily understand that any macromolecule, including virtually all proteins or peptides, may serve as an antigen. An antigen may be endogenously expressed, *i.e.* expressed by genomic DNA, or may be recombinantly expressed. An antigen may be specific to a certain tissue, such as a cancer cell, or it may be broadly expressed. In addition, fragments of larger molecules may act as antigens. In some embodiments, antigens are tumor antigens.

[0256] The term “autologous” refers to any material derived from the same individual to which it is later to be re-introduced. For example, the engineered autologous cell therapy (eACT™) method described herein involves collection of lymphocytes from a patient, which are then engineered to express, *e.g.*, a CAR construct, and then administered back to the same patient.

[0257] The term “allogeneic” refers to any material derived from one individual which is then introduced to another individual of the same species, *e.g.*, allogeneic T cell transplantation.

[0258] A “cancer” refers to a broad group of various diseases characterized by the uncontrolled growth of abnormal cells in the body. Unregulated cell division and growth results in the formation of malignant tumors that invade neighboring tissues and may also metastasize to distant parts of the body through the lymphatic system or bloodstream. A “cancer” or “cancer tissue” may include a tumor. Examples of cancers that may be treated by

the methods disclosed herein include, but are not limited to, cancers of the immune system including lymphoma, leukemia, myeloma, and other leukocyte malignancies. In some embodiments, the methods disclosed herein may be used to reduce the tumor size of a tumor derived from, for example, bone cancer, pancreatic cancer, skin cancer, cancer of the head or neck, cutaneous or intraocular malignant melanoma, uterine cancer, ovarian cancer, rectal cancer, cancer of the anal region, stomach cancer, testicular cancer, uterine cancer, multiple myeloma, Hodgkin's Disease, non-Hodgkin's lymphoma (NHL), primary mediastinal large B cell lymphoma (PMBC), diffuse large B cell lymphoma (DLBCL), follicular lymphoma (FL), transformed follicular lymphoma, splenic marginal zone lymphoma (SMZL), cancer of the esophagus, cancer of the small intestine, cancer of the endocrine system, cancer of the thyroid gland, cancer of the parathyroid gland, cancer of the adrenal gland, cancer of the urethra, cancer of the penis, chronic or acute leukemia, acute myeloid leukemia, chronic myeloid leukemia, acute lymphoblastic leukemia (ALL) (including non T cell ALL), chronic lymphocytic leukemia (CLL), solid tumors of childhood, lymphocytic lymphoma, cancer of the bladder, cancer of the kidney or ureter, neoplasm of the central nervous system (CNS), primary CNS lymphoma, tumor angiogenesis, spinal axis tumor, brain stem glioma, pituitary adenoma, epidermoid cancer, squamous cell cancer, T cell lymphoma, environmentally induced cancers including those induced by asbestos, other B cell malignancies, and combinations of said cancers. In some embodiments, the methods disclosed herein may be used to reduce the tumor size of a tumor derived from, for example, sarcomas and carcinomas, fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, Kaposi's sarcoma, sarcoma of soft tissue, and other sarcomas, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, hepatocellular carcinoma, lung cancer, colorectal cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma (for example adenocarcinoma of the pancreas, colon, ovary, lung, breast, stomach, prostate, cervix, or esophagus), sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, Wilms' tumor, cervical cancer, testicular tumor, bladder carcinoma, carcinoma of the fallopian tubes, carcinoma of the endometrium, carcinoma of the cervix, carcinoma of the vagina, carcinoma of the vulva, carcinoma of the renal pelvis, CNS tumors (such as a glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, menangioma, melanoma, neuroblastoma and retinoblastoma). The

particular cancer may be responsive to chemo- or radiation therapy or the cancer may be refractory. A refractory cancer refers to a cancer that is not amenable to surgical intervention and the cancer is either initially unresponsive to chemo- or radiation therapy or the cancer becomes unresponsive over time.

[0259] An “anti-tumor effect” as used herein, refers to a biological effect that may present as a decrease in tumor volume, a decrease in the number of tumor cells, a decrease in tumor cell proliferation, a decrease in the number of metastases, an increase in overall or progression-free survival, an increase in life expectancy, or amelioration of various physiological symptoms associated with the tumor. An anti-tumor effect may also refer to the prevention of the occurrence of a tumor, *e.g.*, a vaccine.

[0260] A “cytokine,” as used herein, refers to a non-antibody protein that is released by one cell in response to contact with a specific antigen, wherein the cytokine interacts with a second cell to mediate a response in the second cell. “Cytokine” as used herein is meant to refer to proteins released by one cell population that act on another cell as intercellular mediators. A cytokine may be endogenously expressed by a cell or administered to a subject. Cytokines may be released by immune cells, including macrophages, B cells, T cells, neutrophils, dendritic cells, eosinophils and mast cells to propagate an immune response. Cytokines may induce various responses in the recipient cell. Cytokines may include homeostatic cytokines, chemokines, pro- inflammatory cytokines, effectors, and acute-phase proteins. For example, homeostatic cytokines, including interleukin (IL) 7 and IL-15, promote immune cell survival and proliferation, and pro- inflammatory cytokines may promote an inflammatory response. Examples of homeostatic cytokines include, but are not limited to, IL-2, IL-4, IL-5, IL-7, IL-10, IL-12p40, IL-12p70, IL-15, and interferon (IFN) gamma. Examples of pro-inflammatory cytokines include, but are not limited to, IL-1a, IL-1b, IL- 6, IL-13, IL-17a, IL-23, IL-27, tumor necrosis factor (TNF)-alpha, TNF-beta, fibroblast growth factor (FGF) 2, granulocyte macrophage colony-stimulating factor (GM-CSF), soluble intercellular adhesion molecule 1 (sICAM-1), soluble vascular adhesion molecule 1 (sVCAM-1), vascular endothelial growth factor (VEGF), VEGF-C, VEGF-D, and placental growth factor (PLGF). Examples of effectors include, but are not limited to, granzyme A, granzyme B, soluble Fas ligand (sFasL), TGF-beta, IL-35, and perforin. Examples of acute phase-proteins include, but are not limited to, C-reactive protein (CRP) and serum amyloid A (SAA).

[0261] The term “lymphocyte” as used herein includes natural killer (NK) cells, T cells, or B cells. NK cells are a type of cytotoxic (cell toxic) lymphocyte that represent a major

component of the innate immune system. NK cells reject tumors and cells infected by viruses. It works through the process of apoptosis or programmed cell death. They were termed “natural killers” because they do not require activation in order to kill cells. T cells play a major role in cell-mediated-immunity (no antibody involvement). T cell receptors (TCR) differentiate T cells from other lymphocyte types. The thymus, a specialized organ of the immune system, is the primary site for T cell maturation. There are numerous types of T cells, including: helper T cells (*e.g.*, CD4⁺ cells), cytotoxic T cells (also known as TC, cytotoxic T lymphocytes, CTL, T-killer cells, cytolytic T cells, CD8⁺ T cells or killer T cells), memory T cells ((i) stem memory cells (TSCM), like naive cells, are CD45RO⁻, CCR7⁺, CD45RA⁺, CD62L⁺ (L-selectin), CD27⁺, CD28⁺ and IL-7Ra⁺, but also express large amounts of CD95, IL-2R, CXCR3, and LFA-1, and show numerous functional attributes distinctive of memory cells); (ii) central memory cells (TCM) express L-selectin and CCR7, they secrete IL-2, but not IFN γ or IL-4, and (iii) effector memory cells (TEM), however, do not express L-selectin or CCR7 but produce effector cytokines like IFN γ and IL-4), regulatory T cells (Tregs, suppressor T cells, or CD4⁺CD25⁺ or CD4⁺ FoxP3⁺ regulatory T cells), natural killer T cells (NKT) and gamma delta T cells. B-cells, on the other hand, play a principal role in humoral immunity (with antibody involvement). B-cells make antibodies, are capable of acting as antigen-presenting cells (APCs) and turn into memory B-cells and plasma cells, both short-lived and long-lived, after activation by antigen interaction. In mammals, immature B-cells are formed in the bone marrow.

[0262] The term “genetically engineered” or “engineered” refers to a method of modifying the genome of a cell, including, but not limited to, deleting a coding or non-coding region or a portion thereof or inserting a coding region or a portion thereof. In some embodiments, the cell that is modified is a lymphocyte, *e.g.*, a T cell, which may either be obtained from a patient or a donor. The cell may be modified to express an exogenous construct, such as, *e.g.*, a chimeric antigen receptor (CAR) or a T cell receptor (TCR), which is incorporated into the cell's genome.

[0263] An “immune response” refers to the action of a cell of the immune system (for example, T lymphocytes, B lymphocytes, natural killer (NK) cells, macrophages, eosinophils, mast cells, dendritic cells and neutrophils) and soluble macromolecules produced by any of these cells or the liver (including Abs, cytokines, and complement) that results in selective targeting, binding to, damage to, destruction of, and/or elimination from a vertebrate's body of invading pathogens, cells or tissues infected with pathogens, cancerous or other abnormal cells, or, in cases of autoimmunity or pathological inflammation, normal human cells or

tissues.

[0264] A “costimulatory signal,” as used herein, refers to a signal, which in combination with a primary signal, such as TCR/CD3 ligation, leads to a T cell response, such as, but not limited to, proliferation and/or upregulation or down regulation of key molecules.

[0265] A “costimulatory ligand,” as used herein, includes a molecule on an antigen presenting cell that specifically binds a cognate co-stimulatory molecule on a T cell. Binding of the costimulatory ligand provides a signal that mediates a T cell response, including, but not limited to, proliferation, activation, differentiation, and the like. A costimulatory ligand induces a signal that is in addition to the primary signal provided by a stimulatory molecule, for instance, by binding of a T cell receptor (TCR)/CD3 complex with a major histocompatibility complex (MHC) molecule loaded with peptide. A co-stimulatory ligand may include, but is not limited to, 3/TR6, 4-1BB ligand, agonist or antibody that binds Toll-like receptor, B7-1 (CD80), B7-2 (CD86), CD30 ligand, CD40, CD7, CD70, CD83, herpes virus entry mediator (HVEM), human leukocyte antigen G (HLA-G), ILT4, immunoglobulin-like transcript (ILT) 3, inducible costimulatory ligand (ICOS-L), intercellular adhesion molecule (ICAM), ligand that specifically binds with B7-H3, lymphotoxin beta receptor, MHC class I chain-related protein A (MICA), MHC class I chain-related protein B (MICB), OX40 ligand, PD-L2, or programmed death (PD) L1. A co-stimulatory ligand includes, without limitation, an antibody that specifically binds with a co-stimulatory molecule present on a T cell, such as, but not limited to, 4-1BB, B7-H3, CD2, CD27, CD28, CD30, CD40, CD7, ICOS, ligand that specifically binds with CD83, lymphocyte function-associated antigen-1 (LFA-1), natural killer cell receptor C (NKG2C), OX40, PD-1, or tumor necrosis factor superfamily member 14 (TNFSF14 or LIGHT).

[0266] A “costimulatory molecule” is a cognate binding partner on a T cell that specifically binds with a costimulatory ligand, thereby mediating a costimulatory response by the T cell, such as, but not limited to, proliferation. Costimulatory molecules include, but are not limited to, 4-1BB/CD137, B7-H3, BAFFR, BLAME (SLAMF8), BTLA, CD 33, CD 45, CD100 (SEMA4D), CD103, CD134, CD137, CD154, CD16, CD160 (BY55), CD 18, CD19, CD19a, CD2, CD22, CD247, CD27, CD276 (B7-H3), CD28, CD29, CD3 (alpha; beta; delta; epsilon; gamma; zeta), CD30, CD37, CD4, CD40, CD49a, CD49D, CD49f, CD5, CD64, CD69, CD7, CD80, CD83 ligand, CD84, CD86, CD8alpha, CD8beta, CD9, CD96 (Tactile), CD1-1a, CD1-1b, CD1-1c, CD1-1d, CDS, CEACAM1, CRT AM, DAP-10, DNAM1 (CD226), Fc gamma receptor, GADS, GITR, HVEM (LIGHTR), IA4, ICAM-1, ICAM-1, ICOS, Ig alpha (CD79a), IL2R beta, IL2R gamma, IL7R alpha, integrin, ITGA4, ITGA4,

ITGA6, ITGAD, ITGAE, ITGAL, ITGAM, ITGAX, ITGB2, ITGB7, ITGB1, KIRDS2, LAT, LFA-1, LFA-1, LIGHT, LIGHT (tumor necrosis factor superfamily member 14; TNFSF14), LTBR, Ly9 (CD229), lymphocyte function-associated antigen-1 (LFA-1 (CD11a/CD18)), MHC class I molecule, NKG2C, NKG2D, NKp30, NKp44, NKp46, NKp80 (KLRF1), OX40, PAG/Cbp, PD-1, PSGL1, SELPLG (CD162), signaling lymphocytic activation molecule, SLAM (SLAMF1; CD150; IPO-3), SLAMF4 (CD244; 2B4), SLAMF6 (NTB-A; Lyl08), SLAMF7, SLP-76, TNF, TNF α , TNFR2, Toll ligand receptor, TRANCE/RANKL, VLA1, or VLA-6, or fragments, truncations, or combinations thereof.

[0267] The recitations “sequence identity” or, for example, comprising a “sequence 50% identical to,” as used herein, refer to the extent that sequences are identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a “percentage of sequence identity” may be calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U) or the identical amino acid residue (*e.g.*, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.* the window size), and multiplying the result by 100 to yield the percentage of sequence identity. Included are nucleotides and polypeptides having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to any of the reference sequences described herein, typically where, in the case of polypeptides, the polypeptide variant maintains at least one biological activity of the reference polypeptide.

[0268] As used herein, a “vaccine” refers to a composition for generating immunity for the prophylaxis and/or treatment of diseases. Accordingly, vaccines are medicaments which comprise antigens and are intended to be used in humans or animals for generating specific defense and protective substances upon administration to the human or animal.

[0269] As used herein, a “neoantigen” refers to a class of tumor antigens which arises from tumor-specific mutations in an expressed protein.

2. Vectors, precursor RNA, and circular RNA

[0270] Also provided herein are circular RNAs, precursor RNAs that can circularize into the circular RNAs, and vectors (*e.g.*, DNA vectors) that can be transcribed into the precursor RNAs or the circular RNAs.

[0271] In certain aspects, provided herein are circular RNA polynucleotides comprising a post splicing 3' group I intron fragment, optionally a first spacer, an Internal Ribosome Entry Site (IRES), an expression sequence, optionally a second spacer, and a post splicing 5' group I intron fragment. In some embodiments, these regions are in that order. In some embodiments, the circular RNA is made by a method provided herein or from a vector provided herein.

[0272] In certain embodiments, transcription of a vector provided herein (*e.g.*, comprising a 5' duplex forming region, a 3' group I intron fragment, optionally a first spacer, an Internal Ribosome Entry Site (IRES), a first expression sequence, a polynucleotide sequence encoding a cleavage site, a second expression sequence, optionally a second spacer, a 5' group I intron fragment, and a 3' duplex forming region) results in the formation of a precursor linear RNA polynucleotide capable of circularizing. In some embodiments, this precursor linear RNA polynucleotide circularizes when incubated in the presence of guanosine nucleotide or nucleoside (*e.g.*, GTP) and divalent cation (*e.g.*, Mg²⁺).

[0273] In some embodiments, the vectors and precursor RNA polynucleotides provided herein comprise a first (5') duplex forming region and a second (3') duplex forming region. In certain embodiments, the first and second duplex forming regions may form perfect or imperfect duplexes. Thus, in certain embodiments at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the first and second duplex forming regions may be base paired with one another. In some embodiments, the duplex forming regions are predicted to have less than 50% (*e.g.*, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%) base pairing with unintended sequences in the RNA (*e.g.*, non-duplex forming region sequences). In some embodiments, including such duplex forming regions on the ends of the precursor RNA strand, and adjacent or very close to the group I intron fragment, bring the group I intron fragments in close proximity to each other, increasing splicing efficiency. In some embodiments, the duplex forming regions are 3 to 100 nucleotides in length (*e.g.*, 3-75 nucleotides in length, 3-50 nucleotides in length, 20-50 nucleotides in length, 35-50 nucleotides in length, 5-25 nucleotides in length, 9-19 nucleotides in length). In some embodiments, the duplex forming regions are about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 nucleotides in length. In some embodiments, the duplex forming regions have a length of about 9 to about 50 nucleotides. In one embodiment, the duplex forming regions have a length of about 9 to about 19 nucleotides. In some embodiments, the duplex forming regions have a length of about 20

to about 40 nucleotides. In certain embodiments, the duplex forming regions have a length of about 30 nucleotides.

[0274] Two types of spacers have been designed for improving precursor RNA circularization and/or gene expression from circular RNA. The first type of spacer is external spacer, *i.e.*, present in a precursor RNA but removed upon circularization. While not wishing to be bound by theory, it is contemplated that an external spacer may improve ribozyme-mediated circularization by maintaining the structure of the ribozyme itself and preventing other neighboring sequence elements from interfering with its folding and function. The second type of spacer is internal spacer, *i.e.*, present in a precursor RNA and retained in a resulting circular RNA. While not wishing to be bound by theory, it is contemplated that an internal spacer may improve ribozyme-mediated circularization by maintaining the structure of the ribozyme itself and preventing other neighboring sequence elements, particularly the neighboring IRES and coding region, from interfering with its folding and function. It is also contemplated that an internal spacer may improve protein expression from the IRES by preventing neighboring sequence elements, particularly the intron elements, from hybridizing with sequences within the IRES and inhibiting its ability to fold into its most preferred and active conformation.

[0275] In certain embodiments, the vectors, precursor RNA and circular RNA provided herein comprise a first (5') and/or a second (3') spacer. In some embodiments, including a spacer between the 3' group I intron fragment and the IRES may conserve secondary structures in those regions by preventing them from interacting, thus increasing splicing efficiency. In some embodiments, the first (between 3' group I intron fragment and IRES) and second (between the expression sequences and 5' group I intron fragment) spacers comprise additional base pairing regions that are predicted to base pair with each other and not to the first and second duplex forming regions. In some embodiments, such spacer base pairing brings the group I intron fragments in close proximity to each other, further increasing splicing efficiency. Additionally, in some embodiments, the combination of base pairing between the first and second duplex forming regions, and separately, base pairing between the first and second spacers, promotes the formation of a splicing bubble containing the group I intron fragments flanked by adjacent regions of base pairing. Typical spacers are contiguous sequences with one or more of the following qualities: 1) predicted to avoid interfering with proximal structures, for example, the IRES, expression sequences, or intron; 2) is at least 7 nt long and no longer than 100 nt; 3) is located after and adjacent to the 3' intron fragment and/or before and adjacent to the 5' intron fragment; and 4) contains one or

more of the following: a) an unstructured region at least 5 nt long, b) a region of base pairing at least 5 nt long to a distal sequence, including another spacer, and c) a structured region at least 7 nt long limited in scope to the sequence of the spacer. Spacers may have several regions, including an unstructured region, a base pairing region, a hairpin/structured region, and combinations thereof. In an embodiment, the spacer has a structured region with high GC content. In an embodiment, a region within a spacer base pairs with another region within the same spacer. In an embodiment, a region within a spacer base pairs with a region within another spacer. In an embodiment, a spacer comprises one or more hairpin structures. In an embodiment, a spacer comprises one or more hairpin structures with a stem of 4 to 12 nucleotides and a loop of 2 to 10 nucleotides. In an embodiment, there is an additional spacer between the 3' group I intron fragment and the IRES. In an embodiment, this additional spacer prevents the structured regions of the IRES from interfering with the folding of the 3' group I intron fragment or reduces the extent to which this occurs. In some embodiments, the 5' spacer sequence is at least 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25 or 30 nucleotides in length. In some embodiments, the 5' spacer sequence is no more than 100, 90, 80, 70, 60, 50, 45, 40, 35 or 30 nucleotides in length. In some embodiments the 5' spacer sequence is between 5 and 50, 10 and 50, 20 and 50, 20 and 40, and/or 25 and 35 nucleotides in length. In certain embodiments, the 5' spacer sequence is 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 nucleotides in length. In one embodiment, the 5' spacer sequence is a polyA sequence. In another embodiment, the 5' spacer sequence is a polyAC sequence. In one embodiment, a spacer comprises about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% polyAC content. In one embodiment, a spacer comprises about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% polypyrimidine (C/T or C/U) content.

[0276] In certain embodiments, a 3' group I intron fragment is a contiguous sequence at least 75% homologous (*e.g.*, at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% homologous) to a 3' proximal fragment of a natural group I intron including the 3' splice site dinucleotide and optionally the adjacent exon sequence at least 1 nt in length (*e.g.*, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 or 30 nt in length) and at most the length of the exon. Typically, a 5' group I intron fragment is a contiguous sequence at least 75% homologous (*e.g.*, at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% homologous) to a 5' proximal fragment of a natural group I intron including the 5' splice site dinucleotide and optionally the adjacent exon sequence at least 1 nt in length (*e.g.*, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 or 30 nt in length) and at most the

length of the exon. As described by Umekage *et al.* (2012), external portions of the 3' group I intron fragment and 5' group I intron fragment are removed in circularization, causing the circular RNA provided herein to comprise only the portion of the 3' group I intron fragment formed by the optional exon sequence of at least 1 nt in length and 5' group I intron fragment formed by the optional exon sequence of at least 1 nt in length, if such sequences were present on the non-circularized precursor RNA. The part of the 3' group I intron fragment that is retained by a circular RNA is referred to herein as the "post splicing 3' group I intron fragment". The part of the 5' group I intron fragment that is retained by a circular RNA is referred to herein as the "post splicing 5' group I intron fragment".

[0277] In certain embodiments, the vectors, precursor RNA and circular RNA provided herein comprise an internal ribosome entry site (IRES). Inclusion of an IRES permits the translation of one or more open reading frames from a circular RNA (*e.g.*, open reading frames that form the expression sequences). The IRES element attracts a eukaryotic ribosomal translation initiation complex and promotes translation initiation. See, *e.g.*, Kaufman *et al.*, *Nuc. Acids Res.* (1991) 19:4485-4490; Gurtu *et al.*, *Biochem. Biophys. Res. Comm.* (1996) 229:295-298; Rees *et al.*, *BioTechniques* (1996) 20: 102-110; Kobayashi *et al.*, *BioTechniques* (1996) 21 :399-402; and Mosser *et al.*, *BioTechniques* 1997 22 150-161.

[0278] A multitude of IRES sequences are available and include sequences derived from a wide variety of viruses, such as from leader sequences of picornaviruses such as the encephalomyocarditis virus (EMCV) UTR (Jang *et al.*, *J. Virol.* (1989) 63: 1651-1660), the polio leader sequence, the hepatitis A virus leader, the hepatitis C virus IRES, human rhinovirus type 2 IRES (Dobrikova *et al.*, *Proc. Natl. Acad. Sci.* (2003) 100(25): 15125-15130), an IRES element from the foot and mouth disease virus (Ramesh *et al.*, *Nucl. Acid Res.* (1996) 24:2697-2700), a giardiavirus IRES (Garlapati *et al.*, *J. Biol. Chem.* (2004) 279(5):3389-3397), and the like.

[0279] In some embodiments, an IRES is an IRES sequence of Taura syndrome virus, Triatoma virus, Theiler's encephalomyelitis virus, Simian Virus 40, Solenopsis invicta virus 1, Rhopalosiphum padi virus, Reticuloendotheliosis virus, Human poliovirus 1, Plautia stali intestine virus, Kashmir bee virus, Human rhinovirus 2, Homalodisca coagulata virus- 1, Human Immunodeficiency Virus type 1, , Himetobi P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus , Foot and mouth disease virus, Human enterovirus 71, Equine rhinitis virus, Ectropis obliqua picorna-like virus, Encephalomyocarditis virus, Drosophila C Virus, Human coxsackievirus B3, Crucifer tobamovirus, Cricket paralysis virus, Bovine viral diarrhea virus 1, Black Queen Cell Virus, Aphid lethal paralysis virus, Avian

encephalomyelitis virus, Acute bee paralysis virus, Hibiscus chlorotic ringspot virus, Classical swine fever virus, Human FGF2, Human SFTPA1, Human AML1/RUNX1, *Drosophila antennapedia*, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 alpha, Human n.myc, Mouse Gtx, Human p27kipl, Human PDGF2/c-sis, Human p53, Human Pim-1, Mouse Rbm3, *Drosophila reaper*, Canine Scamper, *Drosophila Ubx*, Human UNR, Mouse UtrA, Human VEGF-A, Human XIAP, *Drosophila hairless*, *S. cerevisiae* TFIID, *S. cerevisiae* YAP1, tobacco etch virus, turnip crinkle virus, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, Picobirnavirus, HCV QC64, Human Cosavirus E/D, Human Cosavirus F, Human Cosavirus JMY, Rhinovirus NAT001, HRV14, HRV89, HRVC-02, HRV-A21, Salivirus A SH1, Salivirus FHB, Salivirus NG-J1, Human Parechovirus 1, Crohivirus B, Yc-3, Rosavirus M-7, Shanbavirus A, Pasivirus A, Pasivirus A 2, Echovirus E14, Human Parechovirus 5, Aichi Virus, Hepatitis A Virus HA16, Phopivirus, CVA10, Enterovirus C, Enterovirus D, Enterovirus J, Human Pegivirus 2, GBV-C GT110, GBV-C K1737, GBV-C Iowa, Pegivirus A 1220, Pasivirus A 3, Sapelovirus, Rosavirus B, Bakunsa Virus, Tremovirus A, Swine Pasivirus 1, PLV-CHN, Pasivirus A, Sicinivirus, Hepacivirus K, Hepacivirus A, BVDV1, Border Disease Virus, BVDV2, CSFV-PK15C, SF573 Dicistrovirus, Hubei Picorna-like Virus, CRPV, Salivirus A BN5, Salivirus A BN2, Salivirus A 02394, Salivirus A GUT, Salivirus A CH, Salivirus A SZ1, Salivirus FHB, CVB3, CVB1, Echovirus 7, CVB5, EVA71, CVA3, CVA12, EV24 or an aptamer to eIF4G.

[0280] For driving protein expression, the circular RNA comprises an IRES operably linked to a protein coding sequence. Exemplary IRES sequences are provided in Table 17 below. In some embodiments, the circular RNA disclosed herein comprises an IRES sequence at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to an IRES sequence in Table 17. In some embodiments, the circular RNA disclosed herein comprises an IRES sequence in Table 17. Modifications of IRES and accessory sequences are disclosed herein to increase or reduce IRES activities, for example, by truncating the 5' and/or 3' ends of the IRES, adding a spacer 5' to the IRES, modifying the 6 nucleotides 5' to the translation initiation site (Kozak sequence), modification of alternative translation initiation sites, and creating chimeric/hybrid IRES sequences. In some embodiments, the IRES sequence in the circular RNA disclosed herein comprises one or more of these modifications relative to a native IRES (*e.g.*, a native IRES disclosed in Table 17).

[0281] A multitude of IRES sequences are available and include sequences derived from

a wide variety of viruses, such as from leader sequences of picornaviruses such as the encephalomyocarditis virus (EMCV) UTR (Jang *et al.* J. Virol. (1989) 63: 1651-1660), the polio leader sequence, the hepatitis A virus leader, the hepatitis C virus IRES, human rhinovirus type 2 IRES (Dobrikova *et al.*, Proc. Natl. Acad. Sci. (2003) 100(25): 15125-15130), an IRES element from the foot and mouth disease virus (Ramesh *et al.*, Nucl. Acid Res. (1996) 24:2697-2700), a giardavirus IRES (Garlapati *et al.*, J. Biol. Chem. (2004) 279(5):3389-3397), and the like.

[0282] In some embodiments, the IRES is an IRES sequence of Taura syndrome virus, Triatoma virus, Theiler's encephalomyelitis virus, Simian Virus 40, Solenopsis invicta virus 1, Rhopalosiphum padi virus, Reticuloendotheliosis virus, Human poliovirus 1, Plautia stali intestine virus, Kashmir bee virus, Human rhinovirus 2, Homalodisca coagulata virus- 1, Human Immunodeficiency Virus type 1, , Himetobi P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus , Foot and mouth disease virus, Human enterovirus 71, Equine rhinitis virus, Ectropis obliqua picorna-like virus, Encephalomyocarditis virus, Drosophila C Virus, Human coxsackievirus B3, Crucifer tobamovirus, Cricket paralysis virus, Bovine viral diarrhea virus 1, Black Queen Cell Virus, Aphid lethal paralysis virus, Avian encephalomyelitis virus, Acute bee paralysis virus, Hibiscus chlorotic ringspot virus, Classical swine fever virus, Human FGF2, Human SFTPA1, Human AML1/RUNX1, Drosophila antennapedia, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 alpha, Human n.myc, Mouse Gtx, Human p27kipl, Human PDGF2/c-sis, Human p53, Human Pim-1, Mouse Rbm3, Drosophila reaper, Canine Scamper, Drosophila Ubx, Human UNR, Mouse UtrA, Human VEGF-A, Human XIAP, Drosophila hairless, S. cerevisiae TFIID, S. cerevisiae YAP1, tobacco etch virus, turnip crinkle virus, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, Picobirnavirus, HCV QC64, Human Cosavirus E/D, Human Cosavirus F, Human Cosavirus JMY, Rhinovirus NAT001, HRV14, HRV89, HRVC-02, HRV-A21, Salivirus A SH1, Salivirus FHB, Salivirus NG-J1, Human Parechovirus 1, Crohivirus B, Yc-3, Rosavirus M-7, Shanbavirus A, Pasivirus A, Pasivirus A 2, Echovirus E14, Human Parechovirus 5, Aichi Virus, Hepatitis A Virus HA16, Phopivirus, CVA10, Enterovirus C, Enterovirus D, Enterovirus J, Human Pegivirus 2, GBV-C GT110, GBV-C K1737, GBV-C Iowa, Pegivirus A 1220, Pasivirus A 3, Sapelovirus, Rosavirus B, Bakunsa Virus, Tremovirus A, Swine Pasivirus 1, PLV-CHN, Pasivirus A, Sicinivirus, Hepacivirus K, Hepacivirus A, BVDV1, Border Disease Virus, BVDV2, CSFV-PK15C, SF573 Dicistrovirus, Hubei Picorna-like Virus, CRPV, Salivirus A BN5, Salivirus A BN2,

Salivirus A 02394, Salivirus A GUT, Salivirus A CH, Salivirus A SZ1, Salivirus FHB, CVB3, CVB1, Echovirus 7, CVB5, EVA71, CVA3, CVA12, EV24 or an aptamer to eIF4G.

[0283] In some embodiments, the polynucleotides herein comprise more than one expression sequence. In some embodiments, the circular RNA is a bicistronic RNA. The sequences encoding the two or more polypeptides can be separated by a ribosomal skipping element or a nucleotide sequence encoding a protease cleavage site. In certain embodiments, the ribosomal skipping element encodes thosea-asigna virus 2A peptide (T2A), porcine teschovirus-1 2 A peptide (P2A), foot-and-mouth disease virus 2 A peptide (F2A), equine rhinitis A virus 2A peptide (E2A), cytoplasmic polyhedrosis virus 2A peptide (BmCPV 2A), or flacherie virus of B. mori 2A peptide (BmIFV 2A).

[0284] In certain embodiments, the vectors provided herein comprise a 3' UTR. In some embodiments, the 3' UTR is from human beta globin, human alpha globin, xenopus beta globin, xenopus alpha globin, human prolactin, human GAP-43, human eEF1a, human Tau, human TNF α , dengue virus, hantavirus small mRNA, bunyavirus small mRNA, turnip yellow mosaic virus, hepatitis C virus, rubella virus, tobacco mosaic virus, human IL-8, human actin, human GAPDH, human tubulin, hibiscus chlorotic ringspot virus, woodchuck hepatitis virus post translationally regulated element, sindbis virus, turnip crinkle virus, tobacco etch virus, or Venezuelan equine encephalitis virus.

[0285] In some embodiments, the vectors provided herein comprise a 5' UTR. In some embodiments, the 5' UTR is from human beta globin, *Xenopus laevis* beta globin, human alpha globin, *Xenopus laevis* alpha globin, rubella virus, tobacco mosaic virus, mouse Gtx, dengue virus, heat shock protein 70kDa protein 1A, tobacco alcohol dehydrogenase, tobacco etch virus, turnip crinkle virus, or the adenovirus tripartite leader.

[0286] In some embodiments, a vector provided herein comprises a polyA region external of the 3' and/or 5' group I intron fragments. In some embodiments the polyA region is at least 15, 30, or 60 nucleotides long. In some embodiments, one or both polyA regions is 15-50 nucleotides long. In some embodiments, one or both polyA regions is 20-25 nucleotides long. The polyA sequence is removed upon circularization. Thus, an oligonucleotide hybridizing with the polyA sequence, such as a deoxythymine oligonucleotide (oligo(dT)) conjugated to a solid surface (e.g., a resin), can be used to separate circular RNA from its precursor RNA. Other sequences can also be disposed 5' to the 3' group I intron fragment or 3' to the 5' group I intron fragment and a complementary sequence can similarly be used for circular RNA purification.

[0287] In some embodiments, the DNA (e.g., vector), linear RNA (e.g., precursor RNA),

and/or circular RNA polynucleotide provided herein is between 300 and 15000, 300 and 14000, 300 and 13000, 300 and 12000, 300 and 11000, 300 and 10000, 400 and 9000, 500 and 8000, 600 and 7000, 700 and 6000, 800 and 5000, 900 and 5000, 1000 and 5000, 1100 and 5000, 1200 and 5000, 1300 and 5000, 1400 and 5000, and/or 1500 and 5000 nucleotides in length. In some embodiments, the polynucleotide is at least 300 nt, 400 nt, 500 nt, 600 nt, 700 nt, 800 nt, 900 nt, 1000 nt, 1100 nt, 1200 nt, 1300 nt, 1400 nt, 1500 nt, 2000 nt, 2500 nt, 3000 nt, 3500 nt, 4000 nt, 4500 nt, 5000 nt, 6000 nt, 7000 nt, 8000 nt, 9000 nt, 10000 nt, 11000 nt, 12000 nt, 13000 nt, 14000 nt, or 15000 nt in length. In some embodiments, the polynucleotide is no more than 3000 nt, 3500 nt, 4000 nt, 4500 nt, 5000 nt, 6000 nt, 7000 nt, 8000 nt, 9000 nt, 10000 nt, 11000 nt, 12000 nt, 13000 nt, 14000 nt, 15000 nt, or 16000 nt in length. In some embodiments, the length of a DNA, linear RNA, and/or circular RNA polynucleotide provided herein is about 300 nt, 400 nt, 500 nt, 600 nt, 700 nt, 800 nt, 900 nt, 1000 nt, 1100 nt, 1200 nt, 1300 nt, 1400 nt, 1500 nt, 2000 nt, 2500 nt, 3000 nt, 3500 nt, 4000 nt, 4500 nt, 5000 nt, 6000 nt, 7000 nt, 8000 nt, 9000 nt, 10000 nt, 11000 nt, 12000 nt, 13000 nt, 14000 nt, or 15000 nt.

[0288] In some embodiments, provided herein is a vector. In certain embodiments, the vector comprises, in the following order, a) a 5' duplex forming region, b) a 3' group I intron fragment, c) optionally, a first spacer sequence, d) an IRES, e) a first expression sequence, f) a polynucleotide sequence encoding a cleavage site, g) a second expression sequence, h) optionally, a second spacer sequence, i) a 5' group I intron fragment, and g) a 3' duplex forming region. In some embodiments, the vector comprises a transcriptional promoter upstream of the 5' duplex forming region.

[0289] In some embodiments, provided herein is a vector. In certain embodiments, the vector comprises, in the following order, a) a 5' duplex forming region, b) a 3' group I intron fragment, c) optionally, a first spacer sequence, d) a first IRES, e) a first expression sequence, f) a second IRES, g) a second expression sequence, h) optionally, a second spacer sequence, i) a 5' group I intron fragment, and g) a 3' duplex forming region. In some embodiments, the vector comprises a transcriptional promoter upstream of the 5' duplex forming region.

[0290] In some embodiments, provided herein is a precursor RNA. In certain embodiments, the precursor RNA is a linear RNA produced by in vitro transcription of a vector provided herein. In some embodiments, the precursor RNA comprises, in the following order, a) optionally, a 5' duplex forming region, b) a 3' group I intron fragment, c) optionally, a first spacer sequence, d) an IRES, e) a first expression sequence, f) a polynucleotide sequence encoding a cleavage site, g) a second expression sequence, h)

optionally, a second spacer sequence, i) a 5' group I intron fragment, and j) optionally, a 3' duplex forming region. In some embodiments, the precursor RNA comprises, in the following order, a) a 5' duplex forming region, b) a 3' group I intron fragment, c) optionally, a first spacer sequence, d) a first IRES, e) a first expression sequence, f) a second IRES, g) a second expression sequence, h) optionally, a second spacer sequence, i) a 5' group I intron fragment, and j) a 3' duplex forming region. The precursor RNA can be unmodified, partially modified or completely modified.

[0291] In certain embodiments, provided herein is a circular RNA. In certain embodiments, the circular RNA is a circular RNA produced by a vector provided herein. In some embodiments, the circular RNA is circular RNA produced by circularization of a precursor RNA provided herein. In certain embodiments, transcription of a vector provided herein results in the formation of a precursor linear RNA capable of circularizing. In some embodiments, this precursor linear RNA polynucleotide circularizes when incubated in the presence of guanosine nucleotide or nucleoside (*e.g.*, GTP) and divalent cation (*e.g.*, Mg^{2+}). In some embodiments, the circular RNA comprises, in the following sequence, a) a first spacer sequence, b) an IRES, c) a first expression sequence, d) a polynucleotide sequence encoding a cleavage site, e) a second expression sequence, and f) a second spacer sequence. In some embodiments, the circular RNA comprises, in the following sequence, a) a post splicing 3' group I intron fragment, b) a first spacer sequence, c) an IRES, d) a first expression sequence, e) a polynucleotide sequence encoding a cleavage site, f) a second expression sequence, and g) a second spacer sequence, h) a post splicing 5' group I intron fragment. In some embodiments, the circular RNA comprises, in the following sequence, a) a first spacer sequence, b) a first IRES, c) a first expression sequence, d) a second IRES, e) a second expression sequence, and f) a second spacer sequence. In some embodiments, the circular RNA further comprises the portion of the 3' group I intron fragment that is 3' of the 3' splice site. In some embodiments, the circular RNA further comprises the portion of the 5' group I intron fragment that is 5' of the 5' splice site. In some embodiments, the circular RNA is at least 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000, 14000, or 15000 nucleotides in size. The circular RNA can be unmodified, partially modified or completely modified.

[0292] In some embodiments, the vectors and precursor RNA polynucleotides provided herein comprise a first (5') duplex forming region and a second (3') duplex forming region. In certain embodiments, the first and second homology regions may form perfect or imperfect duplexes. Thus, in certain embodiments at least 75%, 80%, 85%, 90%, 91%, 92%, 93%,

94%, 95%, 96%, 97%, 98%, 99% or 100% of the first and second duplex forming regions may be base paired with one another. In some embodiments, the duplex forming regions are predicted to have less than 50% (*e.g.*, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%) base pairing with unintended sequences in the RNA (*e.g.*, non-duplex forming region sequences). In some embodiments, including such duplex forming regions on the ends of the precursor RNA strand, and adjacent or very close to the group I intron fragment, bring the group I intron fragments in close proximity to each other, increasing splicing efficiency. In some embodiments, the duplex forming regions are 3 to 100 nucleotides in length (*e.g.*, 3-75 nucleotides in length, 3-50 nucleotides in length, 20-50 nucleotides in length, 35-50 nucleotides in length, 5-25 nucleotides in length, 9-19 nucleotides in length). In some embodiments, the duplex forming regions are about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 nucleotides in length. In some embodiments, the duplex forming regions have a length of about 9 to about 50 nucleotides. In one embodiment, the duplex forming regions have a length of about 9 to about 19 nucleotides. In some embodiments, the duplex forming regions have a length of about 20 to about 40 nucleotides. In certain embodiments, the duplex forming regions have a length of about 30 nucleotides.

[0293] In some embodiments, the circular RNA provided herein has higher functional stability than mRNA comprising the same expression sequence. In some embodiments, the circular RNA provided herein has higher functional stability than mRNA comprising the same expression sequence, 5mU modifications, an optimized UTR, a cap, and/or a polyA tail.

[0294] In some embodiments, the circular RNA polynucleotide provided herein has a functional half-life of at least 5 hours, 10 hours, 15 hours, 20 hours, 30 hours, 40 hours, 50 hours, 60 hours, 70 hours or 80 hours. In some embodiments, the circular RNA polynucleotide provided herein has a functional half-life of 5-80, 10-70, 15-60, and/or 20-50 hours. In some embodiments, the circular RNA polynucleotide provided herein has a functional half-life greater than (*e.g.*, at least 1.5-fold greater than, at least 2-fold greater than) that of an equivalent linear RNA polynucleotide encoding the same protein. In some embodiments, functional half-life can be assessed through the detection of functional protein synthesis.

[0295] In certain embodiments, the vectors, precursor RNA and circular RNA provided herein comprise a first (5') and/or a second (3') spacer. In some embodiments, including a spacer between the 3' group I intron fragment and the IRES may conserve secondary structures in those regions by preventing them from interacting, thus increasing splicing

efficiency. In some embodiments, the first (between 3' group I intron fragment and IRES) and second (between the two expression sequences and 5' group I intron fragment) spacers comprise additional base pairing regions that are predicted to base pair with each other and not to the first and second duplex forming regions. In other embodiments, the first (between 3' group I intron fragment and IRES) and second (between the one of the expression sequences and 5' group I intron fragment) spacers comprise additional base pairing regions that are predicted to base pair with each other and not to the first and second duplex forming regions. In some embodiments, such spacer base pairing brings the group I intron fragments in close proximity to each other, further increasing splicing efficiency. Additionally, in some embodiments, the combination of base pairing between the first and second duplex forming regions, and separately, base pairing between the first and second spacers, promotes the formation of a splicing bubble containing the group I intron fragments flanked by adjacent regions of base pairing. Typical spacers are contiguous sequences with one or more of the following qualities: 1) predicted to avoid interfering with proximal structures, for example, the IRES, expression sequence, or intron; 2) is at least 7 nt long and no longer than 100 nt; 3) is located after and adjacent to the 3' intron fragment and/or before and adjacent to the 5' intron fragment; and 4) contains one or more of the following: a) an unstructured region at least 5 nt long, b) a region of base pairing at least 5 nt long to a distal sequence, including another spacer, and c) a structured region at least 7 nt long limited in scope to the sequence of the spacer. Spacers may have several regions, including an unstructured region, a base pairing region, a hairpin/structured region, and combinations thereof. In an embodiment, the spacer has a structured region with high GC content. In an embodiment, a region within a spacer base pairs with another region within the same spacer. In an embodiment, a region within a spacer base pairs with a region within another spacer. In an embodiment, a spacer comprises one or more hairpin structures. In an embodiment, a spacer comprises one or more hairpin structures with a stem of 4 to 12 nucleotides and a loop of 2 to 10 nucleotides. In an embodiment, there is an additional spacer between the 3' group I intron fragment and the IRES. In an embodiment, this additional spacer prevents the structured regions of the IRES from interfering with the folding of the 3' group I intron fragment or reduces the extent to which this occurs. In some embodiments, the 5' spacer sequence is at least 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25 or 30 nucleotides in length. In some embodiments, the 5' spacer sequence is no more than 100, 90, 80, 70, 60, 50, 45, 40, 35 or 30 nucleotides in length. In some embodiments the 5' spacer sequence is between 5 and 50, 10 and 50, 20 and 50, 20 and 40, and/or 25 and 35 nucleotides in length. In certain embodiments, the 5' spacer

sequence is 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 nucleotides in length. In one embodiment, the 5' spacer sequence is a polyA sequence. In another embodiment, the 5' spacer sequence is a polyAC sequence. In one embodiment, a spacer comprises about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% polyAC content. In one embodiment, a spacer comprises about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% polypyrimidine (C/T or C/U) content.

[0296] In certain embodiments, a 3' group I intron fragment is a contiguous sequence at least 75% identical (*e.g.*, at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical) to a 3' proximal fragment of a natural group I intron including the 3' splice site dinucleotide and optionally the adjacent exon sequence at least 1 nt in length (*e.g.*, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 or 30 nt in length) and at most the length of the exon. Typically, a 5' group I intron fragment is a contiguous sequence at least 75% identical (*e.g.*, at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical) to a 5' proximal fragment of a natural group I intron including the 5' splice site dinucleotide and optionally the adjacent exon sequence at least 1 nt in length (*e.g.*, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 or 30 nt in length) and at most the length of the exon. As described by Umekage *et al.* (2012), external portions of the 3' group I intron fragment and 5' group I intron fragment are removed in circularization, causing the circular RNA provided herein to comprise only the portion of the 3' group I intron fragment formed by the optional exon sequence of at least 1 nt in length and 5' group I intron fragment formed by the optional exon sequence of at least 1 nt in length, if such sequences were present on the non-circularized precursor RNA. The part of the 3' group I intron fragment that is retained by a circular RNA is referred to herein as the post splicing 3' group I intron fragment. The part of the 5' group I intron fragment that is retained by a circular RNA is referred to herein as the post splicing 5' group I intron fragment.

[0297] In some embodiments, the circular RNA provided herein may have a higher magnitude of expression than equivalent linear mRNA, *e.g.*, a higher magnitude of expression 24 hours after administration of RNA to cells. In some embodiments, the circular RNA provided herein has a higher magnitude of expression than mRNA comprising the same expression sequence, 5moU modifications, an optimized UTR, a cap, and/or a polyA tail.

[0298] In some embodiments, the circular RNA polynucleotide provided herein has a functional half-life of at least 5 hours, 10 hours, 15 hours, 20 hours, 30 hours, 40 hours, 50 hours, 60 hours, 70 hours or 80 hours. In some embodiments, the circular RNA

polynucleotide provided herein has a functional half-life of 5-80, 10-70, 15-60, and/or 20-50 hours. In some embodiments, the circular RNA polynucleotide provided herein has a functional half-life greater than (*e.g.*, at least 1.5-fold greater than, at least 2-fold greater than) that of an equivalent linear RNA polynucleotide encoding the same protein. In some embodiments, functional half-life can be assessed through the detection of functional protein synthesis.

[0299] In some embodiments, the circular RNA polynucleotide provided herein has a half-life of at least 5 hours, 10 hours, 15 hours, 20 hours, 30 hours, 40 hours, 50 hours, 60 hours, 70 hours or 80 hours. In some embodiments, the circular RNA polynucleotide provided herein has a half-life of 5-80, 10-70, 15-60, and/or 20-50 hours. In some embodiments, the circular RNA polynucleotide provided herein has a half-life greater than (*e.g.*, at least 1.5-fold greater than, at least 2-fold greater than) that of an equivalent linear RNA polynucleotide encoding the same protein. In some embodiments, the circular RNA polynucleotide, or pharmaceutical composition thereof, has a functional half-life in a human cell greater than or equal to that of a pre-determined threshold value. In some embodiments the functional half-life is determined by a functional protein assay. For example in some embodiments, the functional half-life is determined by an *in vitro* luciferase assay, wherein the activity of Gaussia luciferase (GLuc) is measured in the media of human cells (*e.g.* HepG2) expressing the circular RNA polynucleotide every 1, 2, 6, 12, or 24 hours over 1, 2, 3, 4, 5, 6, 7, or 14 days. In other embodiments, the functional half-life is determined by an *in vivo* assay, wherein levels of a protein encoded by the expression sequence of the circular RNA polynucleotide are measured in patient serum or tissue samples every 1, 2, 6, 12, or 24 hours over 1, 2, 3, 4, 5, 6, 7, or 14 days. In some embodiments, the pre-determined threshold value is the functional half-life of a reference linear RNA polynucleotide comprising the same expression sequence as the circular RNA polynucleotide.

[0300] In some embodiments, the circular RNA provided herein may have a higher magnitude of expression than equivalent linear mRNA, *e.g.*, a higher magnitude of expression 24 hours after administration of RNA to cells. In some embodiments, the circular RNA provided herein has a higher magnitude of expression than mRNA comprising the same expression sequence, 5moU modifications, an optimized UTR, a cap, and/or a polyA tail.

[0301]

[0302] In some embodiments, the circular RNA provided herein may be less immunogenic than an equivalent mRNA when exposed to an immune system of an organism or a certain type of immune cell. In some embodiments, the circular RNA provided herein is associated with

modulated production of cytokines when exposed to an immune system of an organism or a certain type of immune cell. For example, in some embodiments, the circular RNA provided herein is associated with reduced production of TNF α , RIG-I, IL-2, IL-6, IFN γ , and/or a type I interferon, *e.g.*, IFN- β 1, when exposed to an immune system of an organism or a certain type of immune cell as compared to mRNA comprising the same expression sequence. In some embodiments, the circular RNA provided herein is associated with less TNF α , RIG-I, IL-2, IL-6, IFN γ , and/or type I interferon, *e.g.*, IFN- β 1, transcript induction when exposed to an immune system of an organism or a certain type of immune cell as compared to mRNA comprising the same expression sequence. In some embodiments, the circular RNA provided herein is less immunogenic than mRNA comprising the same expression sequences. In some embodiments, the circular RNA provided herein is less immunogenic than mRNA comprising the same expression sequences, 5moU modifications, an optimized UTR, a cap, and/or a polyA tail.

[0303] In some embodiments, the compositions and methods described herein provide RNA (*e.g.*, circRNA) with higher stability or functional stability than an equivalent linear RNA without the need for nucleoside modifications. In some embodiments, methods for producing RNA lacking nucleoside modifications produce higher percentages of full length transcripts than methods for producing RNA containing nucleoside modifications due to reduced abortive transcription. In some embodiments, the compositions and methods described herein are capable of producing large (*e.g.*, 5kb, 6kb, 7kb, 8kb, 9kb, 10kb, 11kb, 12kb, 13kb, 14kb, or 15kb) RNA constructs without the added abortive transcription associated with RNA containing nucleoside modifications.

[0304] In certain embodiments, the circular RNA provided herein can be transfected into a cell as is, or can be transfected in DNA vector form and transcribed in the cell. Transcription of circular RNA from a transfected DNA vector can be via added polymerases or polymerases encoded by nucleic acids transfected into the cell, or preferably via endogenous polymerases.

[0305] In certain embodiments, a circular RNA polynucleotide provided herein comprises modified RNA nucleotides and/or modified nucleosides. In some embodiments, the modified nucleoside is m⁵C (5-methylcytidine). In another embodiment, the modified nucleoside is m⁵U (5-methyluridine). In another embodiment, the modified nucleoside is m⁶A (N⁶-methyladenosine). In another embodiment, the modified nucleoside is s²U (2-thiouridine). In another embodiment, the modified nucleoside is Ψ (pseudouridine). In another embodiment, the modified nucleoside is Um (2'-O-methyluridine). In other embodiments, the modified nucleoside is m¹A (1-methyladenosine); m²A (2-methyladenosine); Am (2'-O-

methyladenosine); ms²m⁶A (2-methylthio-N⁶-methyladenosine); i⁶A (N⁶-isopentenyladenosine); ms²i⁶A (2-methylthio-N⁶-isopentenyladenosine); io⁶A (N⁶-(cis-hydroxyisopentenyl)adenosine); ms²io⁶A (2-methylthio-N⁶-(cis-hydroxyisopentenyl)adenosine); g⁶A (N⁶-glycinylicarbamoyladenosine); t⁶A (N⁶-threonylcarbamoyladenosine); ms²t⁶A (2-methylthio-N⁶-threonyl carbamoyladenosine); m⁶t⁶A (N⁶-methyl-N⁶-threonylcarbamoyladenosine); hn⁶A (N⁶-hydroxynorvalylcarbamoyladenosine); ms²hn⁶A (2-methylthio-N⁶-hydroxynorvalyl carbamoyladenosine); Ar(p) (2'-O-ribosyladenosine (phosphate)); I (inosine); m¹I (1-methylinosine); m¹Im (1,2'-O-dimethylinosine); m³C (3-methylcytidine); Cm (2'-O-methylcytidine); s²C (2-thiocytidine); ac⁴C (N⁴-acetylcytidine); f⁵C (5-formylcytidine); m⁵Cm (5,2'-O-dimethylcytidine); ac⁴Cm (N⁴-acetyl-2'-O-methylcytidine); k²C (lysidine); m¹G (1-methylguanosine); m²G (N²-methylguanosine); m⁷G (7-methylguanosine); Gm (2'-O-methylguanosine); m²₂G (N²,N²-dimethylguanosine); m²Gm (N²,2'-O-dimethylguanosine); m²₂Gm (N²,N²,2'-O-trimethylguanosine); Gr(p) (2'-O-ribosylguanosine(phosphate)); yW (wybutosine); ozyW (peroxywybutosine); OHyW (hydroxywybutosine); OHyW* (undermodified hydroxywybutosine); imG (wyosine); mimG (methylwyosine); Q (queuosine); oQ (epoxyqueuosine); galQ (galactosyl-queuosine); manQ (mannosyl-queuosine); preQ₀ (7-cyano-7-deazaguanosine); preQ₁ (7-aminomethyl-7-deazaguanosine); G⁺ (archaeosine); D (dihydrouridine); m⁵Um (5,2'-O-dimethyluridine); s⁴U (4-thiouridine); m⁵s²U (5-methyl-2-thiouridine); s²Um (2-thio-2'-O-methyluridine); acp³U (3-(3-amino-3-carboxypropyl)uridine); ho⁵U (5-hydroxyuridine); mo⁵U (5-methoxyuridine); cmo⁵U (uridine 5-oxyacetic acid); mcmo⁵U (uridine 5-oxyacetic acid methyl ester); chm⁵U (5-(carboxyhydroxymethyl)uridine); mchm⁵U (5-(carboxyhydroxymethyl)uridine methyl ester); mcm⁵U (5-methoxycarbonylmethyluridine); mcm⁵Um (5-methoxycarbonylmethyl-2'-O-methyluridine); mcm⁵s²U (5-methoxycarbonylmethyl-2-thiouridine); nm⁵S²U (5-aminomethyl-2-thiouridine); mnm⁵U (5-methylaminomethyluridine); mnm⁵s²U (5-methylaminomethyl-2-thiouridine); mnm⁵se²U (5-methylaminomethyl-2-selenouridine); ncm⁵U (5-carbamoylmethyluridine); ncm⁵Um (5-carbamoylmethyl-2'-O-methyluridine); cmnm⁵U (5-carboxymethylaminomethyluridine); cmnm⁵Um (5-carboxymethylaminomethyl-2'-O-methyluridine); cmnm⁵s²U (5-carboxymethylaminomethyl-2-thiouridine); m⁶₂A (N⁶,N⁶-dimethyladenosine); Im (2'-O-methylinosine); m⁴C (N⁴-methylcytidine); m⁴Cm (N⁴,2'-O-dimethylcytidine); hm⁵C (5-hydroxymethylcytidine); m³U (3-methyluridine); cm⁵U (5-carboxymethyluridine); m⁶Am (N⁶,2'-O-dimethyladenosine); m⁶₂Am (N⁶,N⁶,O-2'-

trimethyladenosine); $m^{2,7}G$ ($N^2,7$ -dimethylguanosine); $m^{2,2,7}G$ ($N^2,N^2,7$ -trimethylguanosine); m^3Um ($3,2'$ -O-dimethyluridine); m^5D (5-methyldihydrouridine); f^5Cm (5-formyl-2'-O-methylcytidine); m^1Gm ($1,2'$ -O-dimethylguanosine); m^1Am ($1,2'$ -O-dimethyladenosine); τm^5U (5-taurinomethyluridine); τm^5s^2U (5-taurinomethyl-2-thiouridine)); imG-14 (4-demethylwyosine); imG2 (isowyosine); or ac^6A (N^6 -acetyladenosine).

[0306] In some embodiments, the modified nucleoside may include a compound selected from the group of: 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-glycylcarbamoyladenosine, 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. In another embodiment, the modifications are independently selected from the group consisting of 5-methylcytosine, pseudouridine and 1-methylpseudouridine.

[0307] In some embodiments, the modified ribonucleosides include 5-methylcytidine, 5-methoxyuridine, 1-methyl-pseudouridine, N6-methyladenosine, and/or pseudouridine. In some embodiments, such modified nucleosides provide additional stability and resistance to immune activation.

[0308] In particular embodiments, polynucleotides may be codon-optimized. A codon optimized sequence may be one in which codons in a polynucleotide encoding a polypeptide have been substituted in order to increase the expression, stability and/or activity of the polypeptide. Factors that influence codon optimization include, but are not limited to one or more of: (i) variation of codon biases between two or more organisms or genes or synthetically constructed bias tables, (ii) variation in the degree of codon bias within an organism, gene, or set of genes, (iii) systematic variation of codons including context, (iv) variation of codons according to their decoding tRNAs, (v) variation of codons according to GC %, either overall or in one position of the triplet, (vi) variation in degree of similarity to a reference sequence for example a naturally occurring sequence, (vii) variation in the codon frequency cutoff, (viii) structural properties of mRNAs transcribed from the DNA sequence, (ix) prior knowledge about the function of the DNA sequences upon which design of the codon substitution set is to be based, and/or (x) systematic variation of codon sets for each amino acid. In some embodiments, a codon optimized polynucleotide may minimize ribozyme collisions and/or limit structural interference between the expression sequence and the IRES.

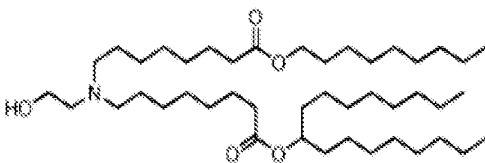
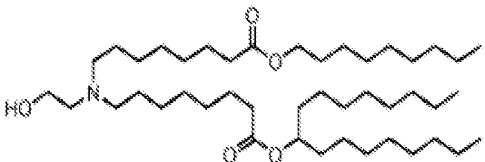
[0309] In certain embodiments circular RNA provided herein is produced inside a cell. In some embodiments, precursor RNA is transcribed using a DNA template (*e.g.*, in some embodiments, using a vector provided herein) in the cytoplasm by a bacteriophage RNA polymerase, or in the nucleus by host RNA polymerase II and then circularized.

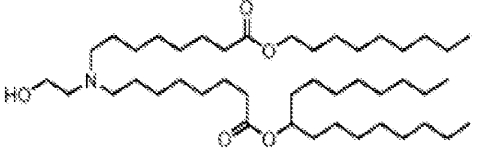
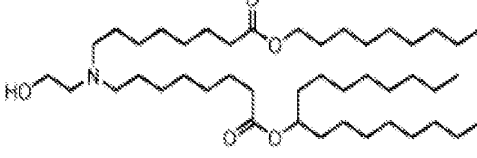
[0310] In certain embodiments, the circular RNA provided herein is injected into an animal (*e.g.*, a human), such that a polypeptide encoded by the circular RNA molecule is expressed inside the animal.

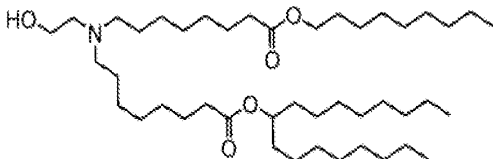
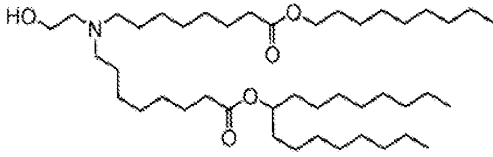
3. Payload

[0311] In some embodiments, the first expression sequence encodes a therapeutic protein. In some embodiments, the second expression sequence encodes a therapeutic protein. In some

embodiments, one or both of the therapeutic proteins are selected from the proteins listed in the following table.

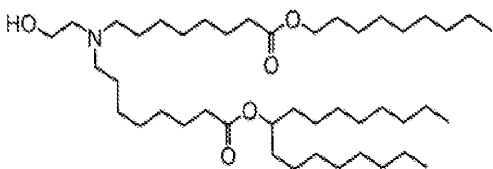
Payload	Sequence	Target cell / organ	Preferred delivery formulation
CD19 CAR	Any of sequences 309-314	T cells	 <p>(50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)</p>
BCMA CAR	MALPVTALLLPLALLL HAARPDIVLTQSPASLA VSLGERATINCRASESV SVIGAHLIHWYQQKPG QPPKLLIYLASNLETGV PARFSGSGSGTDFLTIS SLQAEDAAIYYCLQSRI FPRTFGQGTKLEIKGST SGSGKPGSGEGSTKGQ VQLVQSGSELKKPGAS VKVSCKASGYTFTDYSI NWVRQAPGQGLEWMG WINTETREPAYAYDFR GRFVFSLDTSVSTAYLQ ISSLKAEDTAVYYCAR DYSYAMDYWGQGTLV TVSSAAATTTTPAPRPPT PAPTIASQPLSLRPEACR PAAGGAVHTRGLDFAC DIYIWAPLAGTCGVLLL SLVITLYCKRGRKKLLY IFKQPFMRPVQTTQEED GCSCRFPEEEEGGCELR VKFSRSADAPAYQQGQ NQLYNELNLGRREEYD VLDKRRGRDPEMGGKP RRKNPQEGLYNELQKD KMAEAYSEIGMKGERR RGKGHDGLYQGLSTAT KDTYDALHMQALPPR	T cells	 <p>(50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)</p>

MAGE-A4 TCR	<p>TCR alpha chain: KNQVEQSPQSLILEGK NCTLQCNYTVSPFSNLR WYKQDTGRGPVSLTIM TFSENTKSNGRYTATLD ADTKQSSLHITASQLSD SASYICVVNHSGGSYIP TFGRGTS LIVHPYIQKP DPAVYQLRDSKSSDKS VCLFTDFDSQTNVSQSK DSDVYITDKTVLDMRS MDFKSNSAVAWSNKS DFACANAFNNSIIPEDT FFPSPESS</p> <p>TCR beta chain: DVKVTQSSRYLVKRTG EKVFECEVQDMDHEN MFWYRQDPGLGLRLIY FSYDVKMKEKGDIPG YSVSREKKERFSLILES ASTNQTSMYLCASSFL MTSGDPYEQYFGPGTR LTVTEDLKNVFPPEVA VFEPSEAEISHTQKATL VCLATGFYPDHVELSW WVNGKEVHSGVSTDPO PLKEQPALNDSRYCLSS RLRVSATFWQNP RNHF RCQVQFYGLSENDEWT QDRAKPVTQIVSAEAW GRAD</p>	T cells	 <p>(50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)</p>
NY-ESO TCR	<p>TCR alpha extracellular sequence MQEVTQIPAALSVPEGE NLVLNCSFTDSAIYNLQ WFRQDPGKGLTSLLLIQ SSQREQTSGRNLNASLKD SSGRSTLYIAASQPGDS ATYLCAVRPTSGGSYIP TFGRGTS LIVHPY</p> <p>TCR beta extracellular sequence MGVITQTPKFQVLKTGQ SMTLQCAQDMNHEYM SWYRQDPGMGLRLIHY SVGAGITDQGEVPNGY NVSRSTTEDFPLRLLSA</p>	T cells	 <p>(50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)</p>

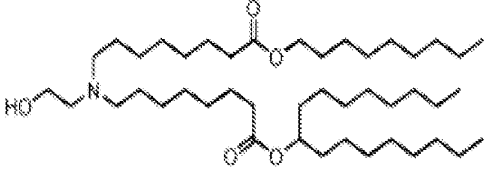
	APSQTSVYFCASSYVG NTGELFFGEGSRLTVL		
EPO	APPRLICDSRVLERYLL EAKEAENITTGCAEHCS LNENITVPDTKVNFYA WKRMEVGQQAVEVW QGLALLSEAVLRGQAL LVNSSQPWEPLQLHVD KAVSGLRSLTTLRLALG AQKEAISPPDAASAAPL RTITADTFRKLFRVYSN FLRGKLKLYTGEACRT GDR	Kidney or bone marrow	
PAH	MSTAVLENPGLGRKLS DFGQETSYIEDNCNQN GAISLIFSLKEEVGALA KVLRLFEENDVNLTHIE SRPSRLKKDEYEFFTHL DKRSLPALTNIIKILRHD IGATVHELSDRDKKDT VPWFPRTIQELDRFANQ ILSYGAELDADHPGFKD PVYRARRKQFADIAYN YRHGQPIPRVEYMEEE KKTWGTVFKTLKSLYK THACYEYNHIFLLEKY CGFHEDNIPQLEDVSQF LQTCTGFRLRPVAGLLS SRDFLGGLAFRVFHCT QYIRHGSKPMYTPEPDI CHELLGHVPLFSDRSFA QFSQEIGLASLGAPDEY IEKLATIWFTVEFGLC KQGDSIKAYGAGLLSSF GELQYCLSEKPKLLPLE LEKTAIQNYTVTEFOPL YYVAESFNDAKEKVRN FAATIPRPFSVRYDPYT QRIEVLDNTQQLKILAD SINSEIGILCSALQKIK	Hepatic cells	 <p>(50 mol %) DSPC (10 mol %) Cholesterol (38.5% mol %) PEG-DMG (1.5%)</p> <p>OR</p> <p>MC3 (50 mol %) DSPC (10 mol %) Cholesterol (38.5% mol %) PEG-DMG (1.5%)</p>
CPSI	LSVKAQTAHIVLEDGT KMKGYSFGHPSSVAGE VVFNTGLGGYPEAITDP AYKGQILTMANPIIGNG GAPDTTALDELGLSKY LESNGIKVSGLLVLDYS KDYNHWLATKSLGQW LQEEKVPAIYGVDTRM LTKIIRDKGTMLGKIEF EGQPVDFVDPNKQNLI	Hepatic cells	 <p>(50 mol %) DSPC (10 mol %) Cholesterol (38.5% mol %) PEG-DMG (1.5%)</p>

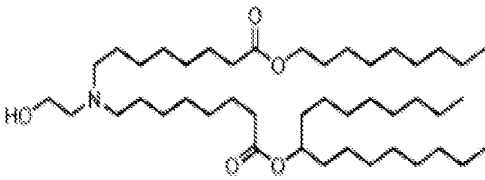
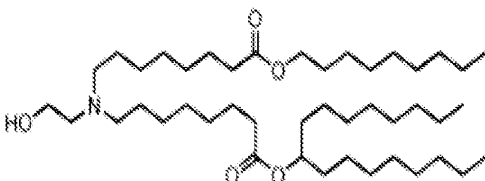
<p> AEVSTKDVKVYVGKGNP TKVVAVDCGIKNNVIR LLVKRGAEVHLPWN HDFTKMEYDGILIAGGP GNPALAEPLIQNVKIL ESDRKEPLFGISTGNLIT GLAAGAKTYKMSMAN RGQNQPVLNITNKQAFI TAQNHGYALDNTLPAG WKPLFVNVNDQTNELI MHESKPFFAVQFHPEV TPGPIDTEYLFDSFFSLI KKGKATTITSVLPKPAL VASRVEVSKVLILGSGG LSIGQAGEFDYSGSQAV KAMKEENVKTVLMNP NIASVQTNEVGLKQAD TVYFLPITPQFVTEVIKA EQPDGLILGMGGQTAL NCGVELFKRGVLKEYG VKVLGTSVESIMATED RQLFSDKLNEINEKIAPS FAVESIEDALKAADTIG YPVMIRSAYALGGLGS GICPNRETLMDLSTKAF AMTNQILVEKSVTGWK EIEYEVVRDADDNCVT VCNMENVNDAMGVHTG DSVVVAPAQTLSNAEF QMLRRTSINVVRHLGIV GECNIQFALHPTSMEYC IIEVNARLSRSSALASK ATGYPLAFIAAKIALGIP LPEIKNVVSGKTSACFE PSLDYMVTKIPRWDL RFHGTSSRIGSSMKSVG EVMAIGRTFEESFQKAL RMCHPSIEGFTPRLPMN KEWPSNLDLRKELSEPS STRIYAIAKAIDDNMSL DEIEKLTIDKWFLYK MRDILNMEKTLKGLNS ESMTEETLKRAKEIGFS DKQISKCLGLTEAQTRE LRLKKNIHPWVKQIDTL AAEPSTVNTYLYVTYN GQEHVDNFDDHGMMV LGCGPYHIGSSVEFDW CAVSSIRTLRQLGKKTV VVNCNPETVSTDFDEC </p>	<p>OR</p> <p> MC3 (50 mol %) DSPC (10 mol %) Cholesterol (38.5% mol %) PEG-DMG (1.5%) </p>
---	---

	<p>DKLYFEELSLERILDIYH QEACGGCIISVGGQIPN NLAVPLYKNGVKIMGT SPLQIDRAEDRSIFS AVL DELKVAQAPWKAVNT LNEALEFAKSVDYPCLL RPSYVLSGSAMNVVFS EDEMKKFLEEATRV SQ EHPVVLTKFVEGAREV EMDAVGKDGRVISHAI SEHVEDAGVHSGDATL MLPTQTISQGAIEKVKD ATR KIAKAF AISGPFNV QFLVKGNDVLVIECNL RASRSFPFVSKTLGVDF IDVATKVMIGENVDEK HLPTLDHPHIPADYVAIK APMFSWPRLRDADPILR CEMASTGEVACFGEI HTAFLKAMLSTGFKIPQ KGILIGIQQSFRPRFLGV AEQLHNEGFKLFATEA TSDWLNANNVPATPVA WPSQEGQNPSLSSIRKLI RDGSIDL VINLPNNNTK FVHDNYVIRRTAVDSGI PLL TNFQVTKLFAEAV QKSRKVDSKSLFH YRQ YSAGKAA</p>		
Cas9	<p>MKRNYILGLDIGITSVG YGIIDYETRDVIDAGVR LFKEANVENNEGRRSK RGARRLKRRRRHRIQR VKLLFDYNLLTDHSE LSGINPYEARVKGLSQK LSEEEFSAALLHLAKRR GVHNVNEVEEDTGNEL STKEQISRNSKALEEKY VAELQLERLKKDGEVR GSINRFKTS DYVKEAK QLLKVQKAYHQLDQSF IDTYIDLLETRRTYYEG PGE GSPFGWKDIKEWY EMLMGHCTYFPEELRS VKYAYNADLYNALND LNNLVITRDENEKLEYY EKFQI IENVFKQKKKPT LKQIAKEILVNEEDIKG YRVTSTGKPEFTNLKV YHDIKDITARKEIENAE</p>	Immun e cells	 (50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)

	LLDQIAKILTIYQSSEDI QEELTNLNSELQTQEEIE QISNLKGYTGTHNLSLK AINLILDELWHTNDNQI AIFNRLKLVPKKVDLSQ QKEIPTTLVDDFILSPVV KRSFIQSIKVINAIIKKY GLPNDIIHELAREKNSKD AQKMINEMQKRNRQT NERIEEIIIRTTGKENAKY LIEKIKLHDMQEGKCLY SLEAIPLEDLLNNPFNY EVDHIIIPRSVSFDNSFNN KVLVKQEENSKKGNRT PFQYLSSSDSKISYETFK KHILNLAAGKGRISKTK KEYLLEERDINRFSVQK DFINRNLDTRYATRG LMNLLRSYFRVNNLDV KVKSSINGGFTSFLRRK WKFKKERNKGYKHHA EDALIIANADFIFKEWK KLDKAKKVMENQMFE EKQAESMPEIETEQEYK EIFITPHQIKHIKDFKDY KYSHRVDKKPNRELIN DTLYSTRKDDKGNTLI VNNLNGLYDKDNDKL KKLINKSPEKLLMYHH DPQTYQKLKLIMEQYG DEKNPLYKYEEETGNY LTKYSKKDNGPVIKKIK YYGNKLNAHLDITDDY PNSRNKVVKLSLKPYR FDVYLDNGVYKFVTVK NLDVIKKENYYEVNSK CYEEAKKLKKISNQAEF IASFYNNDLIKINGELY RVIGVNNDLLNRIEVN MIDITYREYLENMNDK RPPRIIKTIASKTQSIKK YSTDILGNLYEVKSKK HPQIIKKG		
ADAM TS13	AAGGILHLELLVAVGP DVFQAHQEDTERYVLT NLNIGAELLRDPSLGAQ FRVHLVKMVILTEPEG APNITANLTSSLLSVCG WSQTINPEDDTPGHA DLVLYITRFDLELPDGN	Hepatic cells	 (50 mol %)

<p> RQVRGVTQLGGACSPT WCLITEDTGFDLGVTI AHEIGHSFGLHDGAPG SGCGPSGHVMASDGAA PRAGLAWSPCSRRLQLL SLLSAGRARCVDPPR PQPGSAGHPPDAQPGL YYSANEQCRVAFGPKA VACTFAREHLDMCQAL SCHTDPLDQSSCSRLLV PLLDGTECGVEKWCSK GRCRSLVELTPIAAVHG RWSSWGPRSPCSRSCG GGVVTRRRQCNNPRPA FGGRACVGADLQAEM CNTQACEKTQLEFMSQ QCARTDGQPLRSSPGG ASFYHWGAAPHSQG DALCRHMCRAIGESFIM KRGDSFLDGTRCMPSPG PREDGTLSLCVSGSRT FGCDGRMDSQQVWDR CQVCGGDNSTCSPRKG SFTAGRAREYVTFLVT PNLTSVYIANHRPLFTH LAVRIGGRYVVAGKMS ISPNTTYPSSLEDGRVE YRVALTEDRLPRLEEIRI WGPLQEDADIQVYRRY GEEYGNLTRPDITFTYF QPKPRQAWVWAAVRG PCSVSCGAGLRWVNYS CLDQARKELVETVQCQ GSQQPPAWPEACVLEP CPPYWAVGDFGPCSAS CGGGLRERPVRCEAQ GSLLKTLPPARCRAGA QQPAVALETCNPQPCP ARWEVSEPSSCTSAGG AGLALENETCVPGADG LEAPVTEGPGSVDEKLP APEPCVGMSCPPGWH LDATSAGEKAPSPWGS RTGAQAAHVWTPAAG SCSVSCGRGLMELRFLC MDSALRVPVQEELCGL ASKPGSRREVCQAVPC PARWQYKLAACSVSCG RGVVRRILYCARAHGE DDGEEILLDTQCQGLPR </p>	<p> DSPC (10 mol %) Cholesterol (38.5% mol %) PEG-DMG (1.5%) </p> <p>OR</p> <p> MC3 (50 mol %) DSPC (10 mol %) Cholesterol (38.5% mol %) PEG-DMG (1.5%) </p>
---	---

	PEPQEACSLGPCPPRWK VMSLGPCASCLGTA RRSVACVQLDQGDVE VDEAACAALVRPEASV PCLADCTYRWHVGTW MECSVSCGDGIQRRRD TCLGPQAQAPVPADFC QHLPKPVTVRGWAGP CVGQGTPSLVPHEEAA APGRTTATPAGASLEW SQARGLLFSPAPQPRRL LPGPQENSVQSSACGR QHLEPTGTIDMRGPGQ ADCAVAIGRPLGEVVT LRVLESSLNCASAGDML LLWGRLTWRKMCRKL LDMTFSSKTNTLVVRQ RCGRPGGGVLLRYGSQ LAPETFYRECDMQLFG PWGEIVSPSLSPATSNA GGCRLFINVAPHARIAI HALATNMGAGTEGAN ASYLIRDTHSLRTTAFH GQQVLYWESESSQAEM EFSEGFLKAQASLRGQ YWTLQSWVPQMOPQ SWKGKEGT		
FOXP3	MPNPRPGKPSAPSLALG PSPGASPSWRAAPKAS DLLGARGPGGTFQGRD LRGGAHASSSLNPMPP SQLQLPTLPLVMVAPSG ARLGPLPHLQALLQDR PHFMHQLSTVDAHART PVLQVHPLESPAMISLT PPTTATGVFSLKARPGL PPGINVASLEWVSREPA LLCTFPNPSAPRKDSTL SAVPQSSYPPLLANGVC KWPGCEKVFEEDFDL KHCQADHLLDEKGRA QCLLOREMVSLEQQQL VLEKEKLSAMQAHLA KMALTKASSVASSDKG SCCIVAAGSQGPVPA WSGPREAPDSLFAVRR HLWGSNGSTFPEFLH NMDYFKFHNMRPPFTY ATLIRWAILEAPEKQRT LNEIYHWFTRMFAFR	Immun e cells	 <p>(50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)</p>

	NHPATWKNAIRHNLSL HKCFVRVESEKGAVWT VDELEFRKKRSQRPSRC SNPTPGP		
IL-10	SPGQGTQSENSCTHFG NLPNMLRDLRDAFSRV KTFFQMKDQLDNLLK ESLLEDFKGYLGCAALS EMIQFYLEEVMQAEN QDPDIKAHVNSLGENL KTLRLRLRRCHRFLPCE NKSKA VEQVKNAFNKL QEKGIYKAMSEFDIFIN YIEAYMTMKIRN	Immun e cells	 <p>(50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)</p>
IL-2	APTSSSTKKTQLQLEHL LLDLQMILNGINNYKNP KLTRMLTFKFYMPKKA TELKHLQCLEEELKPLE EVLNLAQSKNFHLRPR DLISNINVIVLELKGSET TFMCEYADETATIVEFL NRWITFCQSIISTLT	Immun e cells	 <p>(50 mol %) DSPC (10 mol %) Beta-sitosterol (28.5% mol %) Cholesterol (10 mol %) PEG DMG (1.5 mol %)</p>

[0312] In some embodiments, the at least one of the expression sequences encodes a therapeutic protein. In some embodiments, the first or second expression sequence encodes a cytokine, *e.g.*, IL-12p70, IL-15, IL-2, IL-18, IL-21, IFN- α , IFN- β , IL-10, TGF- β , IL-4, or IL-35, or a functional fragment thereof. In some embodiments, the first or second expression sequence encodes an immune checkpoint inhibitor. In some embodiments, the first or second expression sequence encodes an agonist (*e.g.*, a TNFR family member such as CD137L, OX40L, ICOSL, LIGHT, or CD70). In some embodiments, the first or second expression sequence encodes a chimeric antigen receptor. In some embodiments, the first or second expression sequence encodes an inhibitory receptor agonist (*e.g.*, PDL1, PDL2, Galectin-9, VISTA, B7H4, or MHCII) or inhibitory receptor (*e.g.*, PD1, CTLA4, TIGIT, LAG3, or TIM3). In some embodiments, the first or second expression sequence encodes an inhibitory receptor antagonist. In some embodiments, the first or second expression sequence encodes one or more TCR chains (alpha and beta chains or gamma and delta chains). In some embodiments, the first or second expression sequence encodes a secreted T cell or immune

cell engager (*e.g.*, a bispecific antibody such as BiTE, targeting, *e.g.*, CD3, CD137, or CD28 and a tumor-expressed protein *e.g.*, CD19, CD20, or BCMA etc.). In some embodiments, the first or second expression sequence encodes a transcription factor (*e.g.*, *FOXP3*, *HELIOS*, *TOX1*, or *TOX2*). In some embodiments, the first or second expression sequence encodes an immunosuppressive enzyme (*e.g.*, IDO or CD39/CD73). In some embodiments, the first or second expression sequence encodes a GvHD (*e.g.*, anti-HLA-A2 CAR-Tregs).

[0313] In some embodiments, the first and second expression sequences encode the alpha and beta chains of a T cell receptor (TCR). In some embodiments, the first and second expression sequences encode the gamma and delta chains of a TCR. The invention includes methods of treating a subject suffering from cancer comprising administering a therapeutically effective amount of a composition comprising a circular RNA polynucleotide encoding a TCR alpha chain and a TCR beta chain or a TCR gamma chain and a TCR delta chain.

[0314] In some embodiments, the first and second expression sequences encode a chimeric antigen receptor (CAR) and an antagonist of PD1 or PDL1. In some embodiments, the first and second expression sequences encode a chimeric antigen receptor (CAR) and a cytokine. In some embodiments, the cytokine is IL-12p70, IL-15, IL-2, IL-18, IL-21, IFN- α , IFN- β , IL-10, TGF-beta, IL-4, or IL-35, or a functional fragment thereof. The invention includes methods of treating a subject suffering from cancer comprising administering a therapeutically effective amount of a composition comprising a circular RNA polynucleotide encoding a CAR and an antagonist of PD1 or PDL1. The invention includes methods of treating a subject suffering from cancer comprising administering a therapeutically effective amount of a composition comprising a circular RNA polynucleotide encoding a CAR and a cytokine.

[0315] In some embodiments, the first and second expression sequences encode a transcription factor and a cytokine. In some embodiments, the transcription factor is FOXP3, STAT5B, or HELIOS and the cytokine is IL10, IL12, or TGF beta. The invention includes methods of treating a subject suffering from an autoimmune disorder comprising administering a therapeutically effective amount of a composition comprising a circular RNA polynucleotide encoding a transcription factor, *e.g.*, FOXP3, and a cytokine.

[0316] In some embodiments, the first and second expression sequences encode a transcription factor and a CAR. In some embodiments, the transcription factor is FOXP3, STAT5B, or HELIOS. The invention includes methods of treating a subject suffering from an autoimmune disorder comprising administering a therapeutically effective amount of a

composition comprising a circular RNA polynucleotide encoding a transcription factor, *e.g.*, FOXP3, and a CAR.

[0317] In some embodiments, the first and second expression sequences encode a cytokine and an antigen. In some embodiments, the cytokine is IFN γ . In some embodiments, the antigen is a neoantigen. The invention includes methods of treating a subject suffering from cancer comprising administering a therapeutically effective amount of a composition comprising a circular RNA polynucleotide encoding a cytokine, *e.g.*, IFN γ , and a tumor antigen or fragment thereof.

[0318] In some embodiments, the first expression sequence encodes a first chimeric antigen receptor (CAR) and the second expression sequence encodes a second CAR. In some embodiments, the first CAR is specific for a first antigen and contains a costimulatory domain and an intracellular signaling domain, and the second CAR is specific for a second antigen and contains a costimulatory domain and an intracellular signaling domain. In some embodiments, expressing CARs targeting multiple tumor antigens provides a more effective therapy against a tumor with heterogeneous antigen expression. The invention includes methods of treating a subject suffering from cancer comprising administering a therapeutically effective amount of a composition comprising a circular RNA polynucleotide encoding a first CAR and a second CAR.

[0319] In some embodiments, the first expression sequence encodes a first cytokine and the second expression sequence encodes a second cytokine. In some embodiments, the first and second cytokines are in the group IL-10, TGF β , and IL-35. In some embodiments, the first and second cytokines are in the group IFN γ , IL-2, IL-7, IL-15, and IL-18.

[0320] In some embodiments, a polynucleotide encodes a protein that is made up of subunits that are encoded by more than one gene. For example, the protein may be a heterodimer, wherein each chain or subunit of the protein is encoded by a separate gene. It is possible that more than one circRNA molecule is delivered in the transfer vehicle and each circRNA encodes a separate subunit of the protein. Alternatively, a single circRNA may be engineered to encode more than one subunit. In certain embodiments, separate circRNA molecules encoding the individual subunits may be administered in separate transfer vehicles.

3.1 Cytokines

[0321] Descriptions and/or amino acid sequences of IL-2, IL-7, IL-10, IL-12, IL-15, IL-18, IL-27 β , IFN γ , and/or TGF β 1 are provided herein and at the www.uniprot.org database at accession numbers: P60568 (IL-2), P29459 (IL-12A), P29460 (IL-12B), P13232 (IL-7),

P22301 (IL-10), P40933 (IL-15), Q14116 (IL-18), Q14213 (IL-27 β), P01579 (IFN γ), and/or P01137 (TGF β 1).

3.2 PD-1 and PD-L1 antagonists

[0322] In some embodiments, a PD-1 inhibitor is pembrolizumab, pidilizumab, or nivolumab. In some embodiments, Nivolumab is described in International Patent Publication No. WO2006/121168. In some embodiments, Pembrolizumab is described in WO2009/114335. In some embodiments, Pidilizumab is described in International Patent Publication No. WO2009/101611. Additional anti-PD1 antibodies are described in US Patent No. 8,609,089, U.S. Patent Publication Nos. US 2010028330 and US 20120114649, and International Patent Publication Nos. WO2010/027827 and WO2011/066342.

[0323] In some embodiments, a PD-L1 inhibitor is atezolizumab, avelumab, durvalumab, BMS-936559, or CK-301.

[0324] Descriptions and/or amino acid sequences of heavy and light chains of PD-1, and/or PD-L1 antibodies are provided herein and at the www.drugbank.ca database at accession numbers: DB09037 (Pembrolizumab), DB09035 (Nivolumab), DB15383 (Pidilizumab), DB11595 (Atezolizumab), DB11945 (Avelumab), and DB11714 (Durvalumab).

3.3 Chimeric antigen receptors

[0325] Chimeric antigen receptors (CARs or CAR-Ts) are genetically-engineered receptors. These engineered receptors may be inserted into and expressed by immune cells, including T cells via circular RNA as described herein. With a CAR, a single receptor may be programmed to both recognize a specific antigen and, when bound to that antigen, activate the immune cell to attack and destroy the cell bearing that antigen. When these antigens exist on tumor cells, an immune cell that expresses the CAR may target and kill the tumor cell. In some embodiments, the CAR encoded by the polynucleotide comprises (i) an antigen-binding molecule that specifically binds to a target antigen, (ii) a hinge domain, a transmembrane domain, and an intracellular domain, and (iii) an activating domain.

[0326] In some embodiments, an orientation of the CARs in accordance with the disclosure comprises an antigen binding domain (such as scFv) in tandem with a costimulatory domain and an activating domain. The costimulatory domain may comprise one or more of an extracellular portion, a transmembrane portion, and an intracellular portion. In other embodiments, multiple costimulatory domains may be utilized in tandem.

Antigen binding domain

[0327] CARs may be engineered to bind to an antigen (such as a cell-surface antigen) by incorporating an antigen binding molecule that interacts with that targeted antigen. In some embodiments, the antigen binding molecule is an antibody fragment thereof, *e.g.*, one or more single chain antibody fragment (scFv). An scFv is a single chain antibody fragment having the variable regions of the heavy and light chains of an antibody linked together. See, for example, U.S. Patent Nos. 7,741,465, and 6,319,494 as well as Eshhar *et al.*, Cancer Immunol Immunotherapy (1997) 45: 131-136. An scFv retains the parent antibody's ability to specifically interact with target antigen. scFvs are useful in chimeric antigen receptors because they may be engineered to be expressed as part of a single chain along with the other CAR components. *Id.* See also Krause *et al.*, J. Exp. Med., Volume 188, No. 4, 1998 (619-626); Finney *et al.*, Journal of Immunology, 1998, 161 : 2791-2797. It will be appreciated that the antigen binding molecule is typically contained within the extracellular portion of the CAR such that it is capable of recognizing and binding to the antigen of interest. Bispecific and multispecific CARs are contemplated within the scope of the invention, with specificity to more than one target of interest.

[0328] In some embodiments, the antigen binding molecule comprises a single chain, wherein the heavy chain variable region and the light chain variable region are connected by a linker. In some embodiments, the VH is located at the N terminus of the linker and the VL is located at the C terminus of the linker. In other embodiments, the VL is located at the N terminus of the linker and the VH is located at the C terminus of the linker. In some embodiments, the linker comprises at least about 5, at least about 8, at least about 10, at least about 13, at least about 15, at least about 18, at least about 20, 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 60, at least about 70, at least about 80, at least about 90, or at least about 100 amino acids.

[0329] In some embodiments, the antigen binding molecule comprises a nanobody. In some embodiments, the antigen binding molecule comprises a DARPIn. In some embodiments, the antigen binding molecule comprises an anticalin or other synthetic protein capable of specific binding to target protein.

[0330] In some embodiments, the CAR comprises an antigen binding domain specific for an antigen selected from the group CD19, CD123, CD22, CD30, CD171, CS-1, C-type lectin-like molecule-1, CD33, epidermal growth factor receptor variant III (EGFRvIII), ganglioside G2 (GD2), ganglioside GD3, TNF receptor family member B cell maturation (BCMA), Tn antigen ((Tn Ag) or (GaINAc-Ser/Thr)), prostate-specific membrane antigen

(PSMA), Receptor tyrosine kinase-like orphan receptor 1 (ROR1), Fms-Like Tyrosine Kinase 3 (FLT3), Tumor-associated glycoprotein 72 (TAG72), CD38, CD44v6, Carcinoembryonic antigen (CEA), Epithelial cell adhesion molecule (EPCAM), B7H3 (CD276), KIT (CD117), Interleukin-13 receptor subunit alpha-2, mesothelin, Interleukin 11 receptor alpha (IL-11Ra), prostate stem cell antigen (PSCA), Protease Serine 21, vascular endothelial growth factor receptor 2 (VEGFR2), Lewis(Y) antigen, CD24, Platelet-derived growth factor receptor beta (PDGFR-beta), Stage-specific embryonic antigen-4 (SSEA-4), CD20, Folate receptor alpha, HER2, HER3, Mucin 1, cell surface associated (MUC1), epidermal growth factor receptor (EGFR), neural cell adhesion molecule (NCAM), Prostase, prostatic acid phosphatase (PAP), elongation factor 2 mutated (ELF2M), Ephrin B2, fibroblast activation protein alpha (FAP), insulin-like growth factor 1 receptor (IGF-I receptor), carbonic anhydrase IX (CAIX), Proteasome (Prosome, Macropain) Subunit, Beta Type, 9 (LMP2), glycoprotein 100 (gp100), oncogene fusion protein consisting of breakpoint cluster region (BCR) and Abelson murine leukemia viral oncogene homolog 1 (Abl) (bcr-abl), tyrosinase, ephrin type-A receptor 2 (EphA2), Fucosyl GM1, sialyl Lewis adhesion molecule (sLe), ganglioside GM3, transglutaminase 5 (TGS5), high molecular weight-melanoma-associated antigen (HMWMAA), o-acetyl-GD2 ganglioside (OAcGD2), Folate receptor beta, tumor endothelial marker 1 (TEM1/CD248), tumor endothelial marker 7-related (TEM7R), claudin 6 (CLDN6), thyroid stimulating hormone receptor (TSHR), G protein-coupled receptor class C group 5, member D (GPRC5D), chromosome X open reading frame 61 (CXORF61), CD97, CD179a, anaplastic lymphoma kinase (ALK), Polysialic acid, placenta-specific 1 (PLAC1), hexasaccharide portion of globoH glycosphingolipid (GloboH), mammary gland differentiation antigen (NY-BR-1), uroplakin 2 (UPK2), Hepatitis A virus cellular receptor 1 (HAVCR1), adrenoceptor beta 3 (ADRB3), pannexin 3 (PANX3), G protein-coupled receptor 20 (GPR20), lymphocyte antigen 6 complex, locus K 9 (LY6K), Olfactory receptor 51E2 (OR51E2), TCR Gamma Alternate Reading Frame Protein (TARP), Wilms tumor protein (WT1), Cancer/testis antigen 1 (NY-ESO-1), Cancer/testis antigen 2 (LAGE-1a), MAGE family members (including MAGE-A1, MAGE-A3 and MAGE-A4), ETS translocation-variant gene 6, located on chromosome 12p (ETV6-AML), sperm protein 17 (SPA17), X Antigen Family, Member 1A (XAGE1), angiopoietin-binding cell surface receptor 2 (Tie 2), melanoma cancer testis antigen-1 (MAD-CT-1), melanoma cancer testis antigen-2 (MAD-CT-2), Fos-related antigen 1, tumor protein p53 (p53), p53 mutant, prostatein, surviving, telomerase, prostate carcinoma tumor antigen-1, melanoma antigen recognized by T cells 1, Rat sarcoma (Ras) mutant, human Telomerase reverse transcriptase (hTERT), sarcoma

translocation breakpoints, melanoma inhibitor of apoptosis (ML-IAP), ERG (transmembrane protease, serine 2 (TMPRSS2) ETS fusion gene), N-Acetyl glucosaminyl-transferase V (NA17), paired box protein Pax-3 (PAX3), Androgen receptor, Cyclin B1, v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog (MYCN), Ras Homolog Family Member C (RhoC), Tyrosinase-related protein 2 (TRP-2), Cytochrome P450 1B1 (CYP1B1), CCCTC-Binding Factor (Zinc Finger Protein)-Like, Squamous Cell Carcinoma Antigen Recognized By T Cells 3 (SART3), Paired box protein Pax-5 (PAX5), proacrosin binding protein sp32 (OY-TES1), lymphocyte-specific protein tyrosine kinase (LCK), A kinase anchor protein 4 (AKAP-4), synovial sarcoma, X breakpoint 2 (SSX2), Receptor for Advanced Glycation Endproducts (RAGE-1), renal ubiquitous 1 (RU1), renal ubiquitous 2 (RU2), legumain, human papilloma virus E6 (HPV E6), human papilloma virus E7 (HPV E7), intestinal carboxyl esterase, heat shock protein 70-2 mutated (mut hsp70-2), CD79a, CD79b, CD72, Leukocyte-associated immunoglobulin-like receptor 1 (LAIR1), Fc fragment of IgA receptor (FCAR or CD89), Leukocyte immunoglobulin-like receptor subfamily A member 2 (LILRA2), CD300 molecule-like family member f (CD300LF), C-type lectin domain family 12 member A (CLEC12A), bone marrow stromal cell antigen 2 (BST2), EGF-like module-containing mucin-like hormone receptor-like 2 (EMR2), lymphocyte antigen 75 (LY75), Glypican-3 (GPC3), Fc receptor-like 5 (FCRL5), MUC16, 5T4, 8H9, $\alpha\beta\theta$ integrin, $\alpha\beta6$ integrin, alphafetoprotein (AFP), B7-H6, ca-125, CA9, CD44, CD44v7/8, CD52, E-cadherin, EMA (epithelial membrane antigen), epithelial glycoprotein-2 (EGP-2), epithelial glycoprotein-40 (EGP-40), ErbB4, epithelial tumor antigen (ETA), folate binding protein (FBP), kinase insert domain receptor (KDR), k-light chain, L1 cell adhesion molecule, MUC18, NKG2D, oncofetal antigen (h5T4), tumor/testis-antigen 1B, GAGE, GAGE-1, BAGE, SCP-1, CTZ9, SAGE, CAGE, CT10, MART-1, immunoglobulin lambda-like polypeptide 1 (IGLL1), Hepatitis B Surface Antigen Binding Protein (HBsAg), viral capsid antigen (VCA), early antigen (EA), EBV nuclear antigen (EBNA), HHV-6 p41 early antigen, HHV-6B U94 latent antigen, HHV-6B p98 late antigen, cytomegalovirus (CMV) antigen, large T antigen, small T antigen, adenovirus antigen, respiratory syncytial virus (RSV) antigen, haemagglutinin (HA), neuraminidase (NA), parainfluenza type 1 antigen, parainfluenza type 2 antigen, parainfluenza type 3 antigen, parainfluenza type 4 antigen, Human Metapneumovirus (HMPV) antigen, hepatitis C virus (HCV) core antigen, HIV p24 antigen, human T-cell lymphotropic virus (HTLV-1) antigen, Merkel cell polyoma virus small T antigen, Merkel cell polyoma virus large T antigen, Kaposi sarcoma-associated herpesvirus (KSHV) lytic nuclear antigen and KSHV latent nuclear antigen. In some

embodiments, an antigen binding domain comprises SEQ ID NO: 321 and/or 322.

Hinge / spacer domain

[0331] In some embodiments, a CAR of the instant disclosure comprises a hinge or spacer domain. In some embodiments, the hinge/spacer domain may comprise a truncated hinge/spacer domain (THD) the THD domain is a truncated version of a complete hinge/spacer domain ("CHD"). In some embodiments, an extracellular domain is from or derived from (*e.g.*, comprises all or a fragment of) ErbB2, glycophorin A (GpA), CD2, CD3 delta, CD3 epsilon, CD3 gamma, CD4, CD7, CD8a, CD8[T CD11a (ITGAL), CD11b (ITGAM), CD11c (ITGAX), CD11d (ITGAD), CD18 (ITGB2), CD19 (B4), CD27 (TNFRSF7), CD28, CD28T, CD29 (ITGB1), CD30 (TNFRSF8), CD40 (TNFRSF5), CD48 (SLAMF2), CD49a (ITGA1), CD49d (ITGA4), CD49f (ITGA6), CD66a (CEACAM1), CD66b (CEACAM8), CD66c (CEACAM6), CD66d (CEACAM3), CD66e (CEACAM5), CD69 (CLEC2), CD79A (B-cell antigen receptor complex-associated alpha chain), CD79B (B-cell antigen receptor complex-associated beta chain), CD84 (SLAMF5), CD96 (Tactile), CD100 (SEMA4D), CD103 (ITGAE), CD134 (OX40), CD137 (4-1BB), CD150 (SLAMF1), CD158A (KIR2DL1), CD158B1 (KIR2DL2), CD158B2 (KIR2DL3), CD158C (KIR3DP1), CD158D (KIRDL4), CD158F1 (KIR2DL5A), CD158F2 (KIR2DL5B), CD158K (KIR3DL2), CD160 (BY55), CD162 (SELPLG), CD226 (DNAM1), CD229 (SLAMF3), CD244 (SLAMF4), CD247 (CD3-zeta), CD258 (LIGHT), CD268 (BAFFR), CD270 (TNFSF14), CD272 (BTLA), CD276 (B7-H3), CD279 (PD-1), CD314 (NKG2D), CD319 (SLAMF7), CD335 (NK-p46), CD336 (NK-p44), CD337 (NK-p30), CD352 (SLAMF6), CD353 (SLAMF8), CD355 (CRT AM), CD357 (TNFRSF18), inducible T cell co-stimulator (ICOS), LFA-1 (CD11a/CD18), NKG2C, DAP-10, ICAM-1, NKp80 (KLRP1), IL-2R beta, IL-2R gamma, IL-7R alpha, LFA-1, SLAMF9, LAT, GADS (GrpL), SLP-76 (LCP2), PAG1/CBP, a CD83 ligand, Fc gamma receptor, MHC class I molecule, MHC class II molecule, a TNF receptor protein, an immunoglobulin protein, a cytokine receptor, an integrin, activating NK cell receptors, a Toll ligand receptor, and fragments or combinations thereof. A hinge or spacer domain may be derived either from a natural or from a synthetic source.

[0332] In some embodiments, a hinge or spacer domain is positioned between an antigen binding molecule (*e.g.*, an scFv) and a transmembrane domain. In this orientation, the hinge/spacer domain provides distance between the antigen binding molecule and the surface of a cell membrane on which the CAR is expressed. In some embodiments, a hinge or spacer domain is from or derived from an immunoglobulin. In some embodiments, a hinge or spacer

domain is selected from the hinge/spacer regions of IgG1, IgG2, IgG3, IgG4, IgA, IgD, IgE, and IgM, or a fragment thereof. In some embodiments, a hinge or spacer domain comprises, is from, or is derived from the hinge/spacer region of CD8 alpha. In some embodiments, a hinge or spacer domain comprises, is from, or is derived from the hinge/spacer region of CD28. In some embodiments, a hinge or spacer domain comprises a fragment of the hinge/spacer region of CD8 alpha or a fragment of the hinge/spacer region of CD28, wherein the fragment is anything less than the whole hinge/spacer region. In some embodiments, the fragment of the CD8 alpha hinge/spacer region or the fragment of the CD28 hinge/spacer region comprises an amino acid sequence that excludes at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, or at least 20 amino acids at the N-terminus or C-Terminus, or both, of the CD8 alpha hinge/spacer region, or of the CD28 hinge/spacer region.

Transmembrane domain

[0333] The CAR of the present disclosure may further comprise a transmembrane domain and/or an intracellular signaling domain. The transmembrane domain may be designed to be fused to the extracellular domain of the CAR. It may similarly be fused to the intracellular domain of the CAR. In some embodiments, the transmembrane domain that naturally is associated with one of the domains in a CAR is used. In some instances, the transmembrane domain may be selected or modified (*e.g.*, by an amino acid substitution) to avoid binding of such domains to the transmembrane domains of the same or different surface membrane proteins to minimize interactions with other members of the receptor complex. The transmembrane domain may be derived either from a natural or from a synthetic source. Where the source is natural, the domain may be derived from any membrane-bound or transmembrane protein.

[0334] Transmembrane regions may be derived from (*i.e.* comprise) a receptor tyrosine kinase (*e.g.*, ErbB2), glycoporphin A (GpA), 4-1BB/CD137, activating NK cell receptors, an Immunoglobulin protein, B7-H3, BAFFR, BFAME (SEAMF8), BTEA, CD100 (SEMA4D), CD103, CD160 (BY55), CD 18, CD 19, CD 19a, CD2, CD247, CD27, CD276 (B7-H3), CD28, CD29, CD3 delta, CD3 epsilon, CD3 gamma, CD30, CD4, CD40, CD49a, CD49D, CD49f, CD69, CD7, CD84, CD8alpha, CD8beta, CD96 (Tactile), CD1 la, CD1 lb, CD1 lc, CD1 ld, CDS, CEACAM1, CRT AM, cytokine receptor, DAP-10, DNAM1 (CD226), Fc gamma receptor, GADS, GITR, HVEM (EIGHTR), IA4, ICAM-1, ICAM-1, Ig alpha (CD79a), IE-2R beta, IE-2R gamma, IE-7R alpha, inducible T cell costimulator (ICOS),

integrins, ITGA4, ITGA4, ITGA6, ITGAD, ITGAE, ITGAE, ITGAM, ITGAX, ITGB2, ITGB7, ITGB1, KIRDS2, EAT, LFA-1, LFA-1, a ligand that specifically binds with CD83, LIGHT, LIGHT, LTBR, Ly9 (CD229), lymphocyte function-associated antigen-1 (LFA-1; CD11a/CD18), MHC class 1 molecule, NKG2C, NKG2D, NKp30, NKp44, NKp46, NKp80 (KLRF1), OX-40, PAG/Cbp, programmed death-1 (PD-1), PSGL1, SELPLG (CD162), Signaling Lymphocytic Activation Molecules (SLAM proteins), SLAM (SLAMF1; CD150; IPO-3), SLAMF4 (CD244; 2B4), SLAMF6 (NTB-A; Lyl08), SLAMF7, SLP-76, TNF receptor proteins, TNFR2, TNFSF14, a Toll ligand receptor, TRANCE/RANKL, VLA1, or VLA-6, or a fragment, truncation, or a combination thereof.

[0335] In some embodiments, suitable intracellular signaling domain include, but are not limited to, activating Macrophage/Myeloid cell receptors CSFR1, MYD88, CD14, TIE2, TLR4, CR3, CD64, TREM2, DAP10, DAP12, CD169, DECTIN1, CD206, CD47, CD163, CD36, MARCO, TIM4, MERTK, F4/80, CD91, C1QR, LOX-1, CD68, SRA, BAI-1, ABCA7, CD36, CD31, Lactoferrin, or a fragment, truncation, or combination thereof.

[0336] In some embodiments, a receptor tyrosine kinase may be derived from (*e.g.*, comprise) Insulin receptor (InsR), Insulin-like growth factor I receptor (IGF1R), Insulin receptor-related receptor (IRR), platelet derived growth factor receptor alpha (PDGFR α), platelet derived growth factor receptor beta (PDGFR β), KIT proto-oncogene receptor tyrosine kinase (Kit), colony stimulating factor 1 receptor (CSFR), fms related tyrosine kinase 3 (FLT3), fms related tyrosine kinase 1 (VEGFR-1), kinase insert domain receptor (VEGFR-2), fms related tyrosine kinase 4 (VEGFR-3), fibroblast growth factor receptor 1 (FGFR1), fibroblast growth factor receptor 2 (FGFR2), fibroblast growth factor receptor 3 (FGFR3), fibroblast growth factor receptor 4 (FGFR4), protein tyrosine kinase 7 (CCK4), neurotrophic receptor tyrosine kinase 1 (trkA), neurotrophic receptor tyrosine kinase 2 (trkB), neurotrophic receptor tyrosine kinase 3 (trkC), receptor tyrosine kinase like orphan receptor 1 (ROR1), receptor tyrosine kinase like orphan receptor 2 (ROR2), muscle associated receptor tyrosine kinase (MuSK), MET proto-oncogene, receptor tyrosine kinase (MET), macrophage stimulating 1 receptor (Ron), AXL receptor tyrosine kinase (Axl), TYRO3 protein tyrosine kinase (Tyro3), MER proto-oncogene, tyrosine kinase (Mer), tyrosine kinase with immunoglobulin like and EGF like domains 1 (TIE1), TEK receptor tyrosine kinase (TIE2), EPH receptor A1 (EphA1), EPH receptor A2 (EphA2), (EPH receptor A3) EphA3, EPH receptor A4 (EphA4), EPH receptor A5 (EphA5), EPH receptor A6 (EphA6), EPH receptor A7 (EphA7), EPH receptor A8 (EphA8), EPH receptor A10 (EphA10), EPH receptor B1 (EphB1), EPH receptor B2 (EphB2), EPH receptor B3 (EphB3), EPH receptor B4 (EphB4),

EPH receptor B6 (EphB6), ret proto oncogene (Ret), receptor-like tyrosine kinase (RYK), discoidin domain receptor tyrosine kinase 1 (DDR1), discoidin domain receptor tyrosine kinase 2 (DDR2), c-ros oncogene 1, receptor tyrosine kinase (ROS), apoptosis associated tyrosine kinase (Lmr1), lemur tyrosine kinase 2 (Lmr2), lemur tyrosine kinase 3 (Lmr3), leukocyte receptor tyrosine kinase (LTK), ALK receptor tyrosine kinase (ALK), or serine/threonine/tyrosine kinase 1 (STYK1).

Costimulatory domain

[0337] In certain embodiments, the CAR comprises a costimulatory domain. In some embodiments, the costimulatory domain comprises 4-1BB (CD137), CD28, or both, and/or an intracellular T cell signaling domain. In a preferred embodiment, the costimulatory domain is human CD28, human 4-1BB, or both, and the intracellular T cell signaling domain is human CD3 zeta (ζ). The 4-1BB, CD28, CD3 zeta, or any of these may comprise less than the whole 4-1BB, CD28 or CD3 zeta, respectively. Chimeric antigen receptors may incorporate costimulatory (signaling) domains to increase their potency. See U.S. Patent Nos. 7,741,465, and 6,319,494, as well as Krause *et al.* and Finney *et al.* (*supra*), Song *et al.*, Blood 119:696-706 (2012); Kalos *et al.*, Sci Transl. Med. 3:95 (2011); Porter *et al.*, N. Engl. J. Med. 365:725-33 (2011), and Gross *et al.*, Amur. Rev. Pharmacol. Toxicol. 56:59-83 (2016).

[0338] In some embodiments, a costimulatory domain comprises the amino acid sequence of SEQ ID NO: 318 or 320.

Intracellular signaling domain

[0339] The intracellular (signaling) domain of the engineered T cells disclosed herein may provide signaling to an activating domain, which then activates at least one of the normal effector functions of the immune cell. Effector function of a T cell, for example, may be cytolytic activity or helper activity including the secretion of cytokines.

[0340] In some embodiments, suitable intracellular signaling domain include (*e.g.*, comprise), but are not limited to 4-1BB/CD137, activating NK cell receptors, an immunoglobulin protein, B7-H3, BAFFR, BLAME (SLAMF8), BTLA, CD100 (SEMA4D), CD103, CD160 (BY55), CD18, CD19, CD 19a, CD2, CD247, CD27, CD276 (B7-H3), CD28, CD29, CD3 delta, CD3 epsilon, CD3 gamma, CD30, CD4, CD40, CD49a, CD49D, CD49f, CD69, CD7, CD84, CD8alpha, CD8beta, CD96 (Tactile), CD1 la, CD1 lb, CD1 lc, CD1 ld, CDS, CEACAM1, CRT AM, cytokine receptor, DAP-10, DNAM1 (CD226), Fc gamma receptor, GADS, GITR, HVEM (LIGHTR), IA4, ICAM-1, ICAM-1, Ig alpha (CD79a), IL-2R beta, IL-2R gamma, IL-7R alpha, inducible T cell costimulator (ICOS),

integrins, ITGA4, ITGA4, ITGA6, ITGAD, ITGAE, ITGAL, ITGAM, ITGAX, ITGB2, ITGB7, ITGB1, KIRDS2, LAT, LFA-1, LFA-1, ligand that specifically binds with CD83, LIGHT, LIGHT, LTBR, Ly9 (CD229), Lyl08), lymphocyte function-associated antigen- 1 (LFA-1; CD11a/CD18), MHC class 1 molecule, NKG2C, NKG2D, NKp30, NKp44, NKp46, NKp80 (KLRF1), OX-40, PAG/Cbp, programmed death-1 (PD-1), PSGL1, SELPLG (CD162), Signaling Lymphocytic Activation Molecules (SLAM proteins), SLAM (SLAMF1; CD150, IPO-3), SLAMF4 (CD244; 2B4), SLAMF6 (NTB-A, SLAMF7, SLP-76, TNF receptor proteins, TNFR2, TNFSF14, a Toll ligand receptor, TRANCE/RANKL, VLA1, or VLA-6, or a fragment, truncation, or a combination thereof.

[0341] CD3 is an element of the T cell receptor on native T cells, and has been shown to be an important intracellular activating element in CARs. In some embodiments, the CD3 is CD3 zeta. In some embodiments, the activating domain comprises an amino acid sequence at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identical to the polypeptide sequence of SEQ ID NO: 319.

3.4 T cell receptors

[0342] TCRs are described using the International Immunogenetics (IMGT) TCR nomenclature, and links to the IMGT public database of TCR sequences. Native alpha-beta heterodimeric TCRs have an alpha chain and a beta chain. Broadly, each chain may comprise variable, joining and constant regions, and the beta chain also usually contains a short diversity region between the variable and joining regions, but this diversity region is often considered as part of the joining region. Each variable region may comprise three CDRs (Complementarity Determining Regions) embedded in a framework sequence, one being the hypervariable region named CDR3. There are several types of alpha chain variable ($V\alpha$) regions and several types of beta chain variable ($V\beta$) regions distinguished by their framework, CDR1 and CDR2 sequences, and by a partly defined CDR3 sequence. The $V\alpha$ types are referred to in IMGT nomenclature by a unique TRAV number. Thus “TRAV21” defines a TCR $V\alpha$ region having unique framework and CDR1 and CDR2 sequences, and a CDR3 sequence which is partly defined by an amino acid sequence which is preserved from TCR to TCR but which also includes an amino acid sequence which varies from TCR to TCR. In the same way, “TRBV5-1” defines a TCR $V\beta$ region having unique framework and CDR1 and CDR2 sequences, but with only a partly defined CDR3 sequence.

[0343] The joining regions of the TCR are similarly defined by the unique IMGT TRAJ and TRBJ nomenclature, and the constant regions by the IMGT TRAC and TRBC nomenclature.

[0344] The beta chain diversity region is referred to in IMGT nomenclature by the abbreviation TRBD, and, as mentioned, the concatenated TRBD/TRBJ regions are often considered together as the joining region.

[0345] The unique sequences defined by the IMGT nomenclature are widely known and accessible to those working in the TCR field. For example, they can be found in the IMGT public database. The “T cell Receptor Factsbook”, (2001) LeFranc and LeFranc, Academic Press, ISBN 0-12-441352-8 also discloses sequences defined by the IMGT nomenclature, but because of its publication date and consequent time-lag, the information therein sometimes needs to be confirmed by reference to the IMGT database.

[0346] Native TCRs exist in heterodimeric $\alpha\beta$ or $\gamma\delta$ forms. However, recombinant TCRs consisting of $\alpha\alpha$ or $\beta\beta$ homodimers have previously been shown to bind to peptide MHC molecules. Therefore, the TCR of the invention may be a heterodimeric $\alpha\beta$ TCR or may be an $\alpha\alpha$ or $\beta\beta$ homodimeric TCR.

[0347] For use in adoptive therapy, an $\alpha\beta$ heterodimeric TCR may, for example, be transfected as full length chains having both cytoplasmic and transmembrane domains. In certain embodiments TCRs of the invention may have an introduced disulfide bond between residues of the respective constant domains, as described, for example, in WO 2006/000830.

[0348] TCRs of the invention, particularly alpha-beta heterodimeric TCRs, may comprise an alpha chain TRAC constant domain sequence and/or a beta chain TRBC1 or TRBC2 constant domain sequence. The alpha and beta chain constant domain sequences may be modified by truncation or substitution to delete the native disulfide bond between Cys4 of exon 2 of TRAC and Cys2 of exon 2 of TRBC1 or TRBC2. The alpha and/or beta chain constant domain sequence(s) may also be modified by substitution of cysteine residues for Thr 48 of TRAC and Ser 57 of TRBC1 or TRBC2, the said cysteines forming a disulfide bond between the alpha and beta constant domains of the TCR.

[0349] Binding affinity (inversely proportional to the equilibrium constant K_D) and binding half-life (expressed as $T_{1/2}$) can be determined by any appropriate method. It will be appreciated that doubling the affinity of a TCR results in halving the K_D . $T_{1/2}$ is calculated as $\ln 2$ divided by the off-rate (k_{off}). So doubling of $T_{1/2}$ results in a halving in k_{off} . K_D and k_{off} values for TCRs are usually measured for soluble forms of the TCR, *i.e.* those forms which are truncated to remove cytoplasmic and transmembrane domain residues. Therefore it is to

be understood that a given TCR has an improved binding affinity for, and/or a binding half-life for the parental TCR if a soluble form of that TCR has the said characteristics. Preferably the binding affinity or binding half-life of a given TCR is measured several times, for example 3 or more times, using the same assay protocol, and an average of the results is taken.

[0350] Since the TCRs of the invention have utility in adoptive therapy, the invention includes a non-naturally occurring and/or purified and/or or engineered cell, especially a T-cell, presenting a TCR of the invention. There are a number of methods suitable for the transfection of T cells with nucleic acid (such as DNA, cDNA or RNA) encoding the TCRs of the invention (see for example Robbins *et al.*, (2008) J Immunol. 180: 6116-6131). T cells expressing the TCRs of the invention will be suitable for use in adoptive therapy-based treatment of cancers such as those of the pancreas and liver. As will be known to those skilled in the art, there are a number of suitable methods by which adoptive therapy can be carried out (see for example Rosenberg *et al.*, (2008) Nat Rev Cancer 8(4): 299-308).

[0351] As is well-known in the art TCRs of the invention may be subject to post-translational modifications when expressed by transfected cells. Glycosylation is one such modification, which may comprise the covalent attachment of oligosaccharide moieties to defined amino acids in the TCR chain. For example, asparagine residues, or serine/threonine residues are well-known locations for oligosaccharide attachment. The glycosylation status of a particular protein depends on a number of factors, including protein sequence, protein conformation and the availability of certain enzymes. Furthermore, glycosylation status (i.e. oligosaccharide type, covalent linkage and total number of attachments) can influence protein function. Therefore, when producing recombinant proteins, controlling glycosylation is often desirable. Glycosylation of transfected TCRs may be controlled by mutations of the transfected gene (Kuball J *et al.* (2009), J Exp Med 206(2):463-475). Such mutations are also encompassed in this invention.

[0352] A TCR may be specific for an antigen in the group MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A10, MAGE-A11, MAGE-A12, MAGE-A13, GAGE-1, GAGE-2, GAGE-3, GAGE-4, GAGE-5, GAGE-6, GAGE-7, GAGE-8, BAGE-1, RAGE-1, LB33/MUM-1, PRAME, NAG, MAGE-Xp2 (MAGE-B2), MAGE-Xp3 (MAGE-B3), MAGE-Xp4 (AGE-B4), tyrosinase, brain glycogen phosphorylase, Melan-A, MAGE-C1, MAGE-C2, NY-ESO-1, LAGE-1, SSX-1, SSX-2(HOM-MEL-40), SSX-1, SSX-4, SSX-5, SCP-1, CT-7, alpha-actinin-4, Bcr-Abl fusion protein, Casp-8, beta-catenin, cdc27, cdk4, cdkn2a, coa-1, dek-can

fusion protein, EF2, ETV6-AML1 fusion protein, LDLR-fucosyltransferaseAS fusion protein, HLA-A2, HLA-A11, hsp70-2, KIAA0205, Mart2, Mum-2, and 3, neo-PAP, myosin class I, OS-9, pml-RARa fusion protein, PTPRK, K-ras, N-ras, Triosephosphate isomeras, GnTV, Herv-K-mel, Lage-1, Mage-C2, NA-88, Lage-2, SP17, and TRP2-Int2, (MART-I), gp100 (Pmel 17), TRP-1, TRP-2, MAGE-1, MAGE-3, p15(58), CEA, NY-ESO (LAGE), SCP-1, Hom/Mel-40, p53, H-Ras, HER-2/neu, BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR, Epstein Barr virus antigens, EBNA, human papillomavirus (HPV) antigens E6 and E7, TSP-180, MAGE-4, MAGE-5, MAGE-6, p185erbB2, p180erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CA 19-9, CA 72-4, CAM 17.1, NuMa, K-ras, beta-catenin, CDK4, Mum-1, p16, TAGE, PSMA, PSCA, CT7, telomerase, 43-9F, 5T4, 791Tgp72, α -fetoprotein, 13HCG, BCA225, BTAA, CA 125, CA 15-3 (CA 27.29\BCAA), CA 195, CA 242, CA-50, CAM43, CD68\KP1, CO-029, FGF-5, G250, Ga733 (EpCAM), HTgp-175, M344, MA-50, MG7-Ag, MOV18, NB\170K, NY-CO-1, RCAS1, SDCCAG16, TA-90 (Mac-2 binding protein/cyclophilin C-associated protein), TAAL6, TAG72, TLP, and TPS.

3.5 Transcription factors

[0353] Regulatory T cells (Treg) are important in maintaining homeostasis, controlling the magnitude and duration of the inflammatory response, and in preventing autoimmune and allergic responses.

[0354] In general, Tregs are thought to be mainly involved in suppressing immune responses, functioning in part as a “self-check” for the immune system to prevent excessive reactions. In particular, Tregs are involved in maintaining tolerance to self-antigens, harmless agents such as pollen or food, and abrogating autoimmune disease.

[0355] Tregs are found throughout the body including, without limitation, the gut, skin, lung, and liver. Additionally, Treg cells may also be found in certain compartments of the body that are not directly exposed to the external environment such as the spleen, lymph nodes, and even adipose tissue. Each of these Treg cell populations is known or suspected to have one or more unique features and additional information may be found in Lehtimäki and Lahesmaa, Regulatory T cells control immune responses through their non-redundant tissue specific features, 2013, FRONTIERS IN IMMUNOL., 4(294): 1-10, the disclosure of which is hereby incorporated in its entirety.

[0356] Typically, Tregs are known to require TGF- β and IL-2 for proper activation and development. Tregs, expressing abundant amounts of the IL-2 receptor (IL-2R), are reliant on IL-2 produced by activated T cells. Tregs are known to produce both IL-10 and TGF- β , both

potent immune suppressive cytokines. Additionally, Tregs are known to inhibit the ability of antigen presenting cells (APCs) to stimulate T cells. One proposed mechanism for APC inhibition is via CTLA-4, which is expressed by Foxp3+ Tregs. It is thought that CTLA-4 may bind to B7 molecules on APCs and either block these molecules or remove them by causing internalization resulting in reduced availability of B7 and an inability to provide adequate co-stimulation for immune responses. Additional discussion regarding the origin, differentiation and function of Tregs may be found in Dhamne *et al.*, Peripheral and thymic Foxp3+ regulatory T cells in search of origin, distinction, and function, 2013, *Frontiers in Immunol.*, 4 (253): 1-11, the disclosure of which is hereby incorporated in its entirety.

[0357] Descriptions and/or amino acid sequences of FOXP3, STAT5B, and/or HELIOS are provided herein and at the www.uniprot.org database at accession numbers: Q9BZS1 (FOXP3), P51692 (STAT5b), and/or Q9UKS7 (HELIOS).

Foxp3

[0358] In some embodiments, a transcription factor is the Forkhead box P3 transcription factor (Foxp3). Foxp3 has been shown to be a key regulator in the differentiation and activity of Tregs. In fact, loss-of-function mutations in the Foxp3 gene have been shown to lead to the lethal IPEX syndrome (immune dysregulation, polyendocrinopathy, enteropathy, X-linked). Patients with IPEX suffer from severe autoimmune responses, persistent eczema, and colitis. Treg cells expressing Foxp3 play a key role in limiting inflammatory responses in the intestine (Josefowicz, S. Z. *et al.* *Nature*, 2012, 482, 395-U1510).

STAT

[0359] Members of the signal transducer and activator of transcription (STAT) protein family are intracellular transcription factors that mediate many aspects of cellular immunity, proliferation, apoptosis and differentiation. They are primarily activated by membrane receptor-associated Janus kinases (JAK). Dysregulation of this pathway is frequently observed in primary tumors and leads to increased angiogenesis, enhanced survival of tumors and immunosuppression. Gene knockout studies have provided evidence that STAT proteins are involved in the development and function of the immune system and play a role in maintaining immune tolerance and tumor surveillance.

[0360] There are seven mammalian STAT family members that have been identified: STAT1, STAT2, STAT3, STAT4, STAT5 (including STAT5A and STAT5B), and STAT6.

[0361] Extracellular binding of cytokines or growth factors induce activation of receptor-associated Janus kinases, which phosphorylate a specific tyrosine residue within the STAT protein promoting dimerization via their SH2 domains. The phosphorylated dimer is then

actively transported to the nucleus via an importin α/β ternary complex. Originally, STAT proteins were described as latent cytoplasmic transcription factors as phosphorylation was thought to be required for nuclear retention. However, unphosphorylated STAT proteins also shuttle between the cytosol and nucleus, and play a role in gene expression. Once STAT reaches the nucleus, it binds to a consensus DNA-recognition motif called gamma-activated sites (GAS) in the promoter region of cytokine-inducible genes and activates transcription. The STAT protein can be dephosphorylated by nuclear phosphatases, which leads to inactivation of STAT and subsequent transport out of the nucleus by an exportin-RanGTP complex.

[0362] In some embodiments, a STAT protein of the present disclosure may be a STAT protein that comprises a modification that modulates its expression level or activity. In some embodiments such modifications include, among other things, mutations that effect STAT dimerization, STAT protein binding to signaling partners, STAT protein localization or STAT protein degradation. In some embodiments, a STAT protein of the present disclosure is constitutively active. In some embodiments, a STAT protein of the present disclosure is constitutively active due to constitutive dimerization. In some embodiments, a STAT protein of the present disclosure is constitutively active due to constitutive phosphorylation as described in Onishi, M. *et al.*, Mol. Cell. Biol. July 1998 vol. 18 no. 7 3871-3879 the entirety of which is herein incorporated by reference.

3.6 Vaccines

[0363] In an embodiment, one or more expression sequences encodes an antigen, *e.g.*, a tumor antigen, or a fragment thereof. In some embodiments, expression of such a sequence produces an immunogenic composition, *e.g.*, a vaccine composition capable of raising a specific T-cell response. In some embodiments, an antigen is a neoantigen.

4. Cleavage site

[0364] In some embodiments, two or more expression sequences in a polynucleotide construct may be separated by one or more cleavage site sequences.

[0365] A cleavage site may be any sequence which enables the two or more polypeptides to become separated. A cleavage site may be self-cleaving, such that when the polypeptide is produced, it is immediately cleaved into individual polypeptides without the need for any external cleavage activity.

[0366] A cleavage site may be a furin cleavage site.

[0367] Furin is an enzyme which belongs to the subtilisin-like proprotein convertase family. The members of this family are proprotein convertases that process latent precursor proteins into their biologically active products. Furin is a calcium-dependent serine endoprotease that can efficiently cleave precursor proteins at their paired basic amino acid processing sites. Examples of furin substrates include parathyroid hormone, transforming growth factor beta 1 precursor, proalbumin, pro-beta-secretase, membrane type-1 matrix metalloproteinase, beta subunit of pro-nerve growth factor and von Willebrand factor. Furin cleaves proteins just downstream of a basic amino acid target sequence (canonically, Arg-X-(Arg/Lys)-Arg) and is enriched in the Golgi apparatus.

[0368] A cleavage site may encode a self-cleaving peptide.

[0369] A cleavage site may operate by ribosome skipping such as the skipping of a glycyl-propyl bond at the C-terminus of a 2A self-cleaving peptide. In some embodiments, steric hinderance causes ribosome skipping. In some embodiments, a 2A self-cleaving peptide contains the sequence GDVEXNPGP (SEQ ID NO: 324), wherein X is E or S. In some embodiments, the protein encoded upstream of the 2A self-cleaving peptide is attached to the 2A self-cleaving peptide except the C-terminal proline post translation. In some embodiments, the protein encoded downstream of the 2A self-cleaving peptide is attached to a proline at its N-terminus post translation.

[0370] A self-cleaving peptide may be a 2A self-cleaving peptide from an aphtho- or a cardiovirus. The primary 2A/2B cleavage of the aphtho- and cardioviruses is mediated by 2A cleaving at its own C-terminus. In aphthoviruses, such as foot-and-mouth disease viruses (FMDV) and equine rhinitis A virus, the 2A region is a short section of about 18 amino acids, which, together with the N-terminal residue of protein 2B (a conserved proline residue) represents an autonomous element capable of mediating cleavage at its own C-terminus (Donnelly *et al.* (2001)).

[0371] 2A-like sequences have been found in picornaviruses other than aphtho- or cardioviruses, 'picornavirus-like' insect viruses, type C rotaviruses and repeated sequences within *Trypanosoma* spp and a bacterial sequence (Donnelly *et al.* (2001)). The cleavage site may comprise one of these 2A-like sequences, such as those listed in Table 8.

[0372] In some embodiments, a self-cleaving peptide is F2A. In some embodiments, a self-cleaving peptide is derived from foot-and-mouth disease virus. In some embodiments, a self-cleaving peptide is E2A. In some embodiments, a self-cleaving peptide is derived from equine rhinitis A virus. In some embodiments, a self-cleaving peptide is P2A. In some embodiments, a self-cleaving peptide is derived from porcine teschovirus-1. In some

embodiments, a self-cleaving peptide is T2A. In some embodiments, a self-cleaving peptide is derived from *thossea asigna* virus. In some embodiments, a self-cleaving peptide has a sequence listed in Table 8.

[0373] In an embodiment, expression sequences encoding therapeutic proteins separated by a cleavage site have the same level of protein expression.

[0374] In some embodiments, a self-cleaving peptide is described in Liu, Z., Chen, O., Wall, J.B.J. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* 7, 2193 (2017).

5. Polynucleotides containing a second IRES

[0375] In some embodiments, the ratios of expression of the therapeutic proteins encoded by the first and second expression sequences can be controlled or influenced by the IRES used in the circRNA and whether a cleavage site or a second IRES separates the first and second expression sequences. When equal expression of the proteins encoded by the first and second expression sequences is desired, the circRNA may encode a cleavage site, *e.g.*, a 2A self-cleaving peptide, between the first expression sequence and the second expression sequence. When greater expression of the protein encoded by the first expression sequence is desired, the circRNA may encode a first IRES and a second IRES, wherein the first IRES is associated with greater expression than the second IRES, or wherein the second IRES is an intergenic region (IGR) IRES. When greater expression of the protein encoded by the second expression sequence is desired, the circRNA may encode a first IRES and a second IRES, wherein the second IRES is associated with greater expression than the first IRES.

[0376] In some embodiments, an RNA polynucleotide contains a first IRES and a second IRES as described herein. In some embodiments, a DNA vector encodes a first IRES and a second IRES as described herein.

[0377] In an embodiment, the first IRES and the second IRES have the same sequence. In an embodiment, the first IRES and the second IRES have different sequences. In an embodiment, the first IRES is an IRES having a sequence as listed in Table 1 (SEQ ID NO: 1-72). In some embodiments, the first IRES is a Salivirus IRES. In some embodiments, the first IRES is a Salivirus SZ1 IRES. In an embodiment, the second IRES is an IRES having a sequence as listed in Table 1 (SEQ ID NO: 1-72). In some embodiments, the second IRES is a Salivirus IRES. In some embodiments, the first IRES is a Salivirus SZ1 IRES.

[0378] In some embodiments, the first IRES is associated with greater expression than the second IRES (*e.g.*, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% greater

expression when compared using constructs containing a single IRES). In some embodiments, the second IRES is associated with greater expression than the first IRES (*e.g.*, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% greater expression when compared using constructs containing a single IRES). In some embodiments, the second IRES is an intergenic region (IGR) IRES.

[0379] Expressing 2 proteins with a single circRNA polynucleotide has advantages over expression using multiple polynucleotides. In some embodiments, expression of 2 proteins from an inventive circRNA polynucleotide leads to more consistent ratios of expression than expression from multiple polynucleotides. In some embodiments, expression of 2 proteins from an inventive circRNA polynucleotide leads to transient expression, which may be desirable over the lasting expression of DNA.

6. Production of polynucleotides

[0380] The vectors provided herein can be made using standard techniques of molecular biology. For example, the various elements of the vectors provided herein can be obtained using recombinant methods, such as by screening cDNA and genomic libraries from cells, or by deriving the polynucleotides from a vector known to include the same.

[0381] The various elements of the vectors provided herein can also be produced synthetically, rather than cloned, based on the known sequences. The complete sequence can be assembled from overlapping oligonucleotides prepared by standard methods and assembled into the complete sequence. See, *e.g.*, Edge, *Nature* (1981) 292:756; Nambair *et al.*, *Science* (1984) 223 : 1299; and Jay *et al.*, *J. Biol. Chem.* (1984) 259:631 1.

[0382] Thus, particular nucleotide sequences can be obtained from vectors harboring the desired sequences or synthesized completely, or in part, using various oligonucleotide synthesis techniques known in the art, such as site-directed mutagenesis and polymerase chain reaction (PCR) techniques where appropriate. One method of obtaining nucleotide sequences encoding the desired vector elements is by annealing complementary sets of overlapping synthetic oligonucleotides produced in a conventional, automated polynucleotide synthesizer, followed by ligation with an appropriate DNA ligase and amplification of the ligated nucleotide sequence via PCR. See, *e.g.*, Jayaraman *et al.*, *Proc. Natl. Acad. Sci. USA* (1991) 88:4084-4088. Additionally, oligonucleotide-directed synthesis (Jones *et al.*, *Nature* (1986) 54:75-82), oligonucleotide directed mutagenesis of preexisting nucleotide regions (Riechmann *et al.*, *Nature* (1988) 332:323-327 and Verhoeven *et al.*, *Science* (1988) 239:

1534-1536), and enzymatic filling-in of gapped oligonucleotides using T4 DNA polymerase (Queen *et al.*, Proc. Natl. Acad. Sci. USA (1989) 86: 10029-10033) can be used.

[0383] The precursor RNA provided herein can be generated by incubating a vector provided herein under conditions permissive of transcription of the precursor RNA encoded by the vector. For example, in some embodiments a precursor RNA is synthesized by incubating a vector provided herein that comprises an RNA polymerase promoter upstream of its 5' duplex forming region and/or expression sequences with a compatible RNA polymerase enzyme under conditions permissive of *in vitro* transcription. In some embodiments, the vector is incubated inside of a cell by a bacteriophage RNA polymerase or in the nucleus of a cell by host RNA polymerase II.

[0384] In certain embodiments, provided herein is a method of generating precursor RNA by performing *in vitro* transcription using a vector provided herein as a template (*e.g.*, a vector provided herein with a RNA polymerase promoter positioned upstream of the 5' duplex forming region).

[0385] In certain embodiments, the resulting precursor RNA can be used to generate circular RNA (*e.g.*, a circular RNA polynucleotide provided herein) by incubating it in the presence of magnesium ions and guanosine nucleotide or nucleoside at a temperature at which RNA circularization occurs (*e.g.*, between 20 °C and 60 °C).

[0386] Thus, in certain embodiments provided herein is a method of making circular RNA. In certain embodiments, the method comprises synthesizing precursor RNA by transcription (*e.g.*, run-off transcription) using a vector provided herein (*e.g.*, a post splicing 3' group I intron fragment, an Internal Ribosome Entry Site (IRES), an expression sequence, a polynucleotide sequence encoding a cleavage site, a second expression sequence, and a 5' group I intron fragment) as a template, and incubating the resulting precursor RNA in the presence of divalent cations (*e.g.*, magnesium ions) and GTP such that it circularizes to form circular RNA. In some embodiments, the precursor RNA disclosed herein is capable of circularizing in the absence of magnesium ions and GTP and/or without the step of incubation with magnesium ions and GTP. It has been discovered that circular RNA has reduced immunogenicity relative to a corresponding mRNA, at least partially because the mRNA contains an immunogenic 5' cap. When transcribing a DNA vector from certain promoters (*e.g.*, a T7 promoter) to produce a precursor RNA, it is understood that the 5' end of the precursor RNA is G. To reduce the immunogenicity of a circular RNA composition that contains a low level of contaminant linear mRNA, an excess of GMP relative to GTP can be provided during transcription such that most transcripts contain a 5' GMP, which cannot

be capped. Therefore, in some embodiments, transcription is carried out in the presence of an excess of GMP. In some embodiments, transcription is carried out where the ratio of GMP concentration to GTP concentration is within the range of about 3:1 to about 15:1, for example, about 3:1 to about 10:1, about 3:1 to about 5:1, about 3:1, about 4:1, or about 5:1.

[0387] In some embodiments, a composition comprising circular RNA has been purified. Circular RNA may be purified by any known method commonly used in the art, such as column chromatography, gel filtration chromatography, and size exclusion chromatography.

In some embodiments, purification comprises one or more of the following steps: phosphatase treatment, HPLC size exclusion purification, and RNase R digestion. In some embodiments, purification comprises the following steps in order: RNase R digestion, phosphatase treatment, and HPLC size exclusion purification. In some embodiments, purification comprises reverse phase HPLC. In some embodiments, a purified composition contains less double stranded RNA, DNA splints, triphosphorylated RNA, phosphatase proteins, protein ligases, capping enzymes and/or nicked RNA than unpurified RNA. In some embodiments, a purified composition is less immunogenic than an unpurified composition. In some embodiments, immune cells exposed to a purified composition produce less TNF α , RIG-I, IL-2, IL-6, IFN γ , and/or a type 1 interferon, *e.g.*, IFN- β 1, than immune cells exposed to an unpurified composition.

7. Nanoparticles

[0388] In certain aspects, provided herein are pharmaceutical compositions comprising the circular RNA provided herein. In certain embodiments, such pharmaceutical compositions are formulated with nanoparticles to facilitate delivery.

[0389] In certain embodiments, the circular RNA provided herein may be delivered and/or targeted to a cell in a transfer vehicle, *e.g.*, a nanoparticle, or a composition comprising a nanoparticle. In some embodiments, the circular RNA may also be delivered to a subject in a transfer vehicle or a composition comprising a transfer vehicle. In some embodiments, the transfer vehicle is a nanoparticle. In some embodiments, the nanoparticle is a lipid nanoparticle, a solid lipid nanoparticle, a polymeric core-shell nanoparticle, or a biodegradable nanoparticle. In some embodiments, the transfer vehicle comprises or is coated with one or more cationic lipids, non-cationic lipids, ionizable lipids, PEG-modified lipids, polyglutamic acid polymers, Hyaluronic acid polymers, poly β -amino esters, poly beta amino peptides, or positively charged peptides.

[0390] In one embodiment, the transfer vehicle may be selected and/or prepared to optimize delivery of the circRNA to a target cell. For example, if the target cell is a hepatocyte, the properties of the transfer vehicle (*e.g.*, size, charge and/or pH) may be optimized to effectively deliver such transfer vehicle to the target cell, reduce immune clearance and/or promote retention in that target cell.

[0391] The use of transfer vehicles to facilitate the delivery of nucleic acids to target cells is contemplated by the present invention. Liposomes (*e.g.*, liposomal lipid nanoparticles) are generally useful in a variety of applications in research, industry, and medicine, particularly for their use as transfer vehicles of diagnostic or therapeutic compounds *in vivo* (Lasic, Trends Biotechnol., 16: 307-321, 1998; Drummond *et al.*, Pharmacol. Rev., 51: 691-743, 1999) and are usually characterized as microscopic vesicles having an interior aqueous space sequestered from an outer medium by a membrane of one or more bilayers. Bilayer membranes of liposomes are typically formed by amphiphilic molecules, such as lipids of synthetic or natural origin that comprise spatially separated hydrophilic and hydrophobic domains (Lasic, Trends Biotechnol., 16: 307-321, 1998). Bilayer membranes of the liposomes can also be formed by amphiphilic polymers and surfactants (*e.g.*, polymerosomes, niosomes, etc.).

[0392] In the context of the present invention, a transfer vehicle typically serves to transport the circRNA to the target cell. For the purposes of the present invention, the transfer vehicles are prepared to contain or encapsulate the desired nucleic acids. The process of incorporation of a desired entity (*e.g.*, a nucleic acid) into a liposome is often referred to as loading (Lasic, *et al.*, FEBS Lett., 312: 255-258, 1992). The liposome-incorporated nucleic acids may be completely or partially located in the interior space of the liposome, within the bilayer membrane of the liposome, or associated with the exterior surface of the liposome membrane. The purpose of incorporating a circRNA into a transfer vehicle, such as a liposome, is often to protect the nucleic acid from an environment which may contain enzymes or chemicals that degrade nucleic acids and/or systems or receptors that cause the rapid excretion of the nucleic acids. Accordingly, in an embodiment of the present invention, the selected transfer vehicle is capable of enhancing the stability of the circRNA contained therein. The liposome can allow the encapsulated circRNA to reach the target cell, or alternatively limit the delivery of such circRNA to other sites or cells where the presence of the administered circRNA may be useless or undesirable. Furthermore, incorporating the circRNA into a transfer vehicle, such as, for example, a cationic liposome, also facilitates the delivery of such circRNA into a target cell. In some embodiments, a transfer vehicle

disclosed herein may serve to promote endosomal or lysosomal release of, for example, contents that are encapsulated in the transfer vehicle (*e.g.*, lipid nanoparticle).

[0393] Ideally, transfer vehicles are prepared to encapsulate one or more desired circRNA such that the compositions demonstrate a high transfection efficiency and enhanced stability. While liposomes can facilitate introduction of nucleic acids into target cells, the addition of polycations (*e.g.*, poly L-lysine and protamine), as a copolymer can in some instances markedly enhance the transfection efficiency of several types of cationic liposomes by 2-28 fold in a number of cell lines both *in vitro* and *in vivo*. (See N J. Caplen, *et al.*, *Gene Ther.* 1995; 2: 603; S. Li, *et al.*, *Gene Ther.* 1997; 4, 891.)

[0394] In certain embodiments disclosed herein are ionizable lipids that may be used as a component of a transfer vehicle to facilitate or enhance the delivery and release of circular RNA to one or more target cells (*e.g.*, by permeating or fusing with the lipid membranes of such target cells). In certain embodiments, an ionizable lipid comprises one or more cleavable functional groups (*e.g.*, a disulfide) that allow, for example, a hydrophilic functional head-group to dissociate from a lipophilic functional tail-group of the compound (*e.g.*, upon exposure to oxidative, reducing or acidic conditions), thereby facilitating a phase transition in the lipid bilayer of the one or more target cells.

[0395] In some embodiments, an ionizable lipid is a lipid as described in international patent application PCT/US2018/058555.

[0396] In some of embodiments, a cationic lipid has the following formula:



wherein:

R1 and R2 are either the same or different and independently optionally substituted C₁₀-C₂₄ alkyl, optionally substituted C₁₀-C₂₄ alkenyl, optionally substituted C₁₀-C₂₄ alkynyl, or optionally substituted C₁₀-C₂₄ acyl;

R3 and R4 are either the same or different and independently optionally substituted C₁-C₆ alkyl, optionally substituted C₂-C₆ alkenyl, or optionally substituted C₂-C₆ alkynyl or R3 and R4 may join to form an optionally substituted heterocyclic ring of 4 to 6 carbon atoms and 1 or 2 heteroatoms chosen from nitrogen and oxygen;

R5 is either absent or present and when present is hydrogen or C₁-C₆ alkyl; m, n, and p are either the same or different and independently either 0 or 1 with the proviso that m, n, and

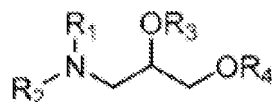
p are not simultaneously 0; q is 0, 1, 2, 3, or 4; and

Y and Z are either the same or different and independently O, S, or NH.

[0397] In one embodiment, R₁ and R₂ are each linoleyl, and the amino lipid is a dilinoleyl amino lipid.

[0398] In one embodiment, the amino lipid is a dilinoleyl amino lipid.

[0399] In various other embodiments, a cationic lipid has the following structure:



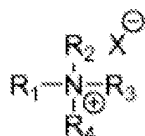
or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

R₁ and R₂ are each independently selected from the group consisting of H and C₁-C₃ alkyls; and

R₃ and R₄ are each independently an alkyl group having from about 10 to about 20 carbon atoms, wherein at least one of R₃ and R₄ comprises at least two sites of unsaturation.

[0400] In some embodiments, R₃ and R₄ are each independently selected from dodecadienyl, tetradecadienyl, hexadecadienyl, linoleyl, and icosadienyl. In an embodiment, R₃ and R₄ are both linoleyl. In some embodiments, R₃ and/or R₄ may comprise at least three sites of unsaturation (*e.g.*, R₃ and/or R₄ may be, for example, dodecatrienyl, tetradectrienyl, hexadecatrienyl, linolenyl, and icosatrienyl).

[0401] In some embodiments, a cationic lipid has the following structure:



or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

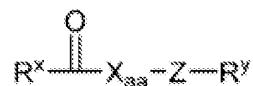
R₁ and R₂ are each independently selected from H and C₁-C₃ alkyls;

R₃ and R₄ are each independently an alkyl group having from about 10 to about 20 carbon atoms, wherein at least one of R₃ and R₄ comprises at least two sites of unsaturation.

[0402] In one embodiment, R₃ and R₄ are the same, for example, in some embodiments R₃ and R₄ are both linoleyl (C₁₈-alkyl). In another embodiment, R₃ and R₄ are different, for example, in some embodiments, R₃ is tetradectrienyl (C₁₄-alkyl) and R₄ is linoleyl (C₁₈-alkyl). In a preferred embodiment, the cationic lipid(s) of the present invention are symmetrical, *i.e.*, R₃ and R₄ are the same. In another preferred embodiment, both R₃ and R₄ comprise at least two sites of unsaturation. In some embodiments, R₃ and R₄ are each

independently selected from dodecadienyl, tetradecadienyl, hexadecadienyl, linoleyl, and icosadienyl. In an embodiment, R₃ and R₄ are both linoleyl. In some embodiments, R₃ and/or R₄ comprise at least three sites of unsaturation and are each independently selected from dodecatrienyl, tetradectrienyl, hexadecatrienyl, linolenyl, and icosatrienyl.

[0403] In various embodiments, a cationic lipid has the formula:



or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

X_{aa} is a D- or L-amino acid residue having the formula $-NR^N-CR^1R^2-C(=O)-$, or a peptide or a peptide of amino acid residues having the formula $- \{NR^N-CR^1R^2-C(=O)\}_n-$, wherein n is an integer from 2 to 20;

R¹ is independently, for each occurrence, a non-hydrogen or a substituted or unsubstituted side chain of an amino acid;

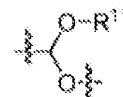
R² and R^N are independently, for each occurrence, hydrogen, an organic group consisting of carbon, oxygen, nitrogen, sulfur, and hydrogen atoms, or any combination of the foregoing, and having from 1 to 20 carbon atoms, C₍₁₋₅₎alkyl, cycloalkyl, cycloalkylalkyl, C₍₁₋₅₎alkenyl, C₍₁₋₅₎alkynyl, C₍₁₋₅₎alkanoyl, C₍₁₋₅₎alkanoyloxy, C₍₁₋₅₎alkoxy, C₍₁₋₅₎alkoxy- C₍₁₋₅₎alkyl, C₍₁₋₅₎alkoxy- C₍₁₋₅₎alkoxy, C₍₁₋₅₎alkyl-amino- C₍₁₋₅₎alkyl-, C₍₁₋₅₎dialkyl-amino- C₍₁₋₅₎alkyl-, nitro-C₍₁₋₅₎alkyl, cyano-C₍₁₋₅₎alkyl, aryl-C₍₁₋₅₎alkyl, 4-biphenyl-C₍₁₋₅₎alkyl, carboxyl, or hydroxyl;

Z is $-NH-$, $-O-$, $-S-$, $-CH_2S-$, $-CH_2S(O)-$, or an organic linker consisting of 1-40 atoms selected from hydrogen, carbon, oxygen, nitrogen, and sulfur atoms (preferably, Z is $-NH-$ or $-O-$);

R^x and R^y are, independently, (i) a lipophilic tail derived from a lipid (which can be naturally occurring or synthetic), *e.g.*, a phospholipid, a glycolipid, a triacylglycerol, a glycerophospholipid, a sphingolipid, a ceramide, a sphingomyelin, a cerebroside, or a ganglioside, wherein the tail optionally includes a steroid; (ii) an amino acid terminal group selected from hydrogen, hydroxyl, amino, and an organic protecting group; or (iii) a substituted or unsubstituted C₍₃₋₂₂₎alkyl, C₍₆₋₁₂₎cycloalkyl, C₍₆₋₁₂₎cycloalkyl- C₍₃₋₂₂₎alkyl, C₍₃₋₂₂₎alkenyl, C₍₃₋₂₂₎alkynyl, C₍₃₋₂₂₎alkoxy, or C₍₆₋₁₂₎-alkoxy C₍₃₋₂₂₎alkyl;

[0404] In some embodiments, one of R^x and R^y is a lipophilic tail as defined above and the other is an amino acid terminal group. In some embodiments, both R^x and R^y are lipophilic tails.

[0405] In some embodiments, at least one of R^x and R^y is interrupted by one or more biodegradable groups (e.g., $-\text{OC}(\text{O})-$, $-\text{C}(\text{O})\text{O}-$, $-\text{SC}(\text{O})-$, $-\text{C}(\text{O})\text{S}-$, $-\text{OC}(\text{S})-$, $-\text{C}(\text{S})\text{O}-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{O})(\text{NR}^5)-$, $-\text{N}(\text{R}^5)\text{C}(\text{O})-$, $-\text{C}(\text{S})(\text{NR}^5)-$, $-\text{N}(\text{R}^5)\text{C}(\text{O})-$, $-\text{N}(\text{R}^5)\text{C}(\text{O})\text{N}(\text{R}^5)-$, $-\text{OC}(\text{O})\text{O}-$, $-\text{OSi}(\text{R}^5)_2\text{O}-$, $-\text{C}(\text{O})(\text{CR}^3\text{R}^4)\text{C}(\text{O})\text{O}-$, $-\text{OC}(\text{O})(\text{CR}^3\text{R}^4)\text{C}(\text{O})-$, or



$\text{OC}(\text{O})\text{O}-$, $-\text{OSi}(\text{R}^5)_2\text{O}-$, $-\text{C}(\text{O})(\text{CR}^3\text{R}^4)\text{C}(\text{O})\text{O}-$, $-\text{OC}(\text{O})(\text{CR}^3\text{R}^4)\text{C}(\text{O})-$, or

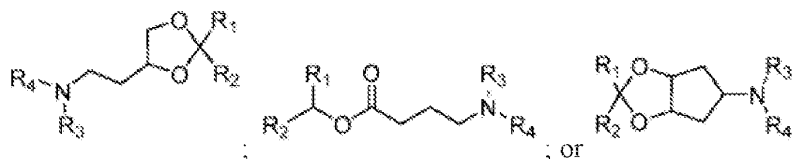
[0406] In some embodiments, R^{11} is a C_2 - C_8 alkyl or alkenyl.

[0407] In some embodiments, each occurrence of R^5 is, independently, H or alkyl.

[0408] In some embodiments, each occurrence of R^3 and R^4 are, independently H, halogen, OH, alkyl, alkoxy, $-\text{NH}_2$, alkylamino, or dialkylamino; or R_3 and R_4 , together with the carbon atom to which they are directly attached, form a cycloalkyl group. In some particular embodiments, each occurrence of R^3 and R^4 are, independently H or C_1 - C_4 alkyl.

[0409] In some embodiments, R^x and R^y each, independently, have one or more carbon-carbon double bonds.

[0410] In some embodiments, the cationic lipid is one of the following:

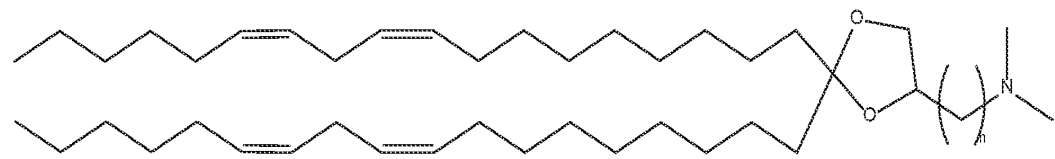


or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

R_1 and R_2 are each independently alkyl, alkenyl, or alkynyl, each of which can optionally substituted;

R_3 and R_4 are each independently a C_1 - C_6 alkyl, or R_3 and R_4 are taken together to form an optionally substituted heterocyclic ring.

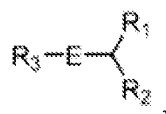
[0411] A representative useful dilinoleyl amino lipid has the formula:



wherein n is 0, 1, 2, 3, or 4.

[0412] In one embodiment, a cationic lipid is DLin-K-DMA. In one embodiment, a cationic lipid is DLin-KC2-DMA (DLin-K-DMA above, wherein n is 2).

[0413] In one embodiment, a cationic lipid has the following structure:



or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

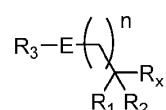
R₁ and R₂ are each independently for each occurrence optionally substituted C₁₀-C₃₀ alkyl, optionally substituted C₁₀-C₃₀ alkenyl, optionally substituted C₁₀-C₃₀ alkynyl or optionally substituted C₁₀-C₃₀ acyl;

R₃ is H, optionally substituted C₂-C₁₀ alkyl, optionally substituted C₂-C₁₀ alkenyl, optionally substituted C₂-C₁₀ alkyl, alkylheterocycle, alkylphosphate, alkylphosphorothioate, alkylphosphorodithioate, alkylphosphonate, alkylamine, hydroxyalkyl, ω-aminoalkyl, ω-(substituted)aminoalkyl, ω-phosphoalkyl, ω-thiophosphoalkyl, optionally substituted polyethylene glycol (PEG, mw 100-40K), optionally substituted mPEG (mw 120-40K), heteroaryl, or heterocycle, or a linker ligand, for example, in some embodiments, R₃ is (CH₃)₂N(CH₂)_n-, wherein n is 1, 2, 3 or 4;

E is O, S, N(Q), C(O), OC(O), C(O)O, N(Q)C(O), C(O)N(Q), (Q)N(CO)O, O(CO)N(Q), S(O), NS(O)₂N(Q), S(O)₂, N(Q)S(O)₂, SS, O=N, aryl, heteroaryl, cyclic or heterocycle, for example -C(O)O, wherein - is a point of connection to R₃; and

5 Q is H, alkyl, ω-aminoalkyl, ω-(substituted)aminoalkyl, ω-phosphoalkyl or ω-thiophosphoalkyl.

In one specific embodiment, the cationic lipid of Embodiments 1, 2, 3, 4 or 5 has the following structure:



10 or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

E is O, S, N(Q), C(O), N(Q)C(O), C(O)N(Q), (Q)N(CO)O, O(CO)N(Q), S(O), NS(O)₂N(Q), S(O)₂, N(Q)S(O)₂, SS, O=N, aryl, heteroaryl, cyclic or heterocycle;

Q is H, alkyl, ω-aminoalkyl, ω-(substituted)aminoalkyl, ω-

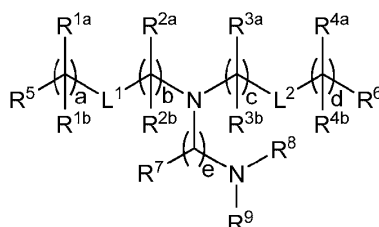
15 phosphoalkyl or ω-thiophosphoalkyl;

R₁ and R₂ and R_x are each independently for each occurrence H, optionally substituted C₁-C₁₀ alkyl, optionally substituted C₁₀-C₃₀ alkyl, optionally substituted C₁₀-C₃₀ alkenyl, optionally substituted C₁₀-C₃₀ alkynyl, optionally substituted C₁₀-C₃₀ acyl, or linker-ligand, provided that at least one of R₁, R₂ and R_x is
20 not H;

R₃ is H, optionally substituted C₁-C₁₀ alkyl, optionally substituted C₂-C₁₀ alkenyl, optionally substituted C₂-C₁₀ alkynyl, alkylheterocycle, alkylphosphate, alkylphosphorothioate, alkylphosphorodithioate, alkylphosphonate, alkylamine, hydroxyalkyl, ω-aminoalkyl, ω-(substituted)aminoalkyl, ω-phosphoalkyl, ω-
25 thiophosphoalkyl, optionally substituted polyethylene glycol (PEG, mw 100-40K), optionally substituted mPEG (mw 120-40K), heteroaryl, or heterocycle, or linker-ligand; and

n is 0, 1, 2, or 3.

In one embodiment, the cationic lipid of Embodiments 1, 2, 3, 4 or 5 has the structure of Formula I:



I

5 or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof,
wherein:

one of L^1 or L^2 is $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $SC(=O)-$, $-NR^aC(=O)-$, $-C(=O)NR^a$, $NR^aC(=O)NR^a$, $-OC(=O)NR^a$ or $-NR^aC(=O)O-$, and the other of L^1 or L^2 is $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $SC(=O)-$, $-NR^aC(=O)-$, $-C(=O)NR^a$, $NR^aC(=O)NR^a$, $-OC(=O)NR^a$ or $-NR^aC(=O)O-$ or a direct bond:

R^a is H or C₁-C₁₂ alkyl;

R^{1a} and R^{1b} are, at each occurrence, independently either (a) H or C₁-C₁₂ alkyl, or (b) R^{1a} is H or C₁-C₁₂ alkyl, and R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^{2a} and R^{2b} are, at each occurrence, independently either (a) H or C₁-C₁₂ alkyl, or (b) R^{2a} is H or C₁-C₁₂ alkyl, and R^{2b} together with the carbon atom to which it is bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^{3a} and R^{3b} are, at each occurrence, independently either (a) H or C₁-C₁₂ alkyl, or (b) R^{3a} is H or C₁-C₁₂ alkyl, and R^{3b} together with the carbon atom to which it is bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^{4a} and R^{4b} are, at each occurrence, independently either (a) H or C_1-C_{12} alkyl, or (b) R^{4a} is H or C_1-C_{12} alkyl, and R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

5 R^5 and R^6 are each independently methyl or cycloalkyl;
 R^7 is, at each occurrence, independently H or C_1-C_{12} alkyl;
 R^8 and R^9 are each independently unsubstituted C_1-C_{12} alkyl; or R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring comprising one nitrogen atom;

10 a and d are each independently an integer from 0 to 24;
b and c are each independently an integer from 1 to 24;
e is 1 or 2; and
x is 0, 1 or 2.

In some embodiments of Formula I, L^1 and L^2 are independently –
15 $O(C=O)-$ or $-(C=O)O-$.

In certain embodiments of Formula I, at least one of R^{1a} , R^{2a} , R^{3a} or R^{4a} is C_1-C_{12} alkyl, or at least one of L^1 or L^2 is $-O(C=O)-$ or $-(C=O)O-$. In other embodiments, R^{1a} and R^{1b} are not isopropyl when a is 6 or n-butyl when a is 8.

In still further embodiments of Formula I, at least one of R^{1a} , R^{2a} , R^{3a} or
20 R^{4a} is C_1-C_{12} alkyl, or at least one of L^1 or L^2 is $-O(C=O)-$ or $-(C=O)O-$; and
 R^{1a} and R^{1b} are not isopropyl when a is 6 or n-butyl when a is 8.

In other embodiments of Formula I, R^8 and R^9 are each independently unsubstituted C_1-C_{12} alkyl; or R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring comprising one nitrogen
25 atom;

In certain embodiments of Formula I, any one of L^1 or L^2 may be $-O(C=O)-$ or a carbon-carbon double bond. L^1 and L^2 may each be $-O(C=O)-$ or may each be a carbon-carbon double bond.

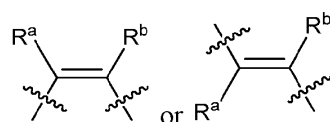
In some embodiments of Formula I, one of L^1 or L^2 is $-O(C=O)-$. In
30 other embodiments, both L^1 and L^2 are $-O(C=O)-$.

In some embodiments of Formula I, one of L^1 or L^2 is $-(C=O)O-$. In other embodiments, both L^1 and L^2 are $-(C=O)O-$.

In some other embodiments of Formula I, one of L^1 or L^2 is a carbon-carbon double bond. In other embodiments, both L^1 and L^2 are a carbon-carbon double bond.

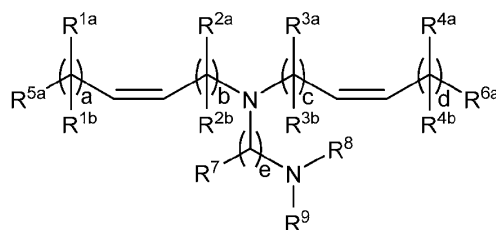
In still other embodiments of Formula I, one of L^1 or L^2 is $-O(C=O)-$ and the other of L^1 or L^2 is $-(C=O)O-$. In more embodiments, one of L^1 or L^2 is $-O(C=O)-$ and the other of L^1 or L^2 is a carbon-carbon double bond. In yet more embodiments, one of L^1 or L^2 is $-(C=O)O-$ and the other of L^1 or L^2 is a carbon-carbon double bond.

It is understood that "carbon-carbon" double bond, as used throughout the specification, refers to one of the following structures:



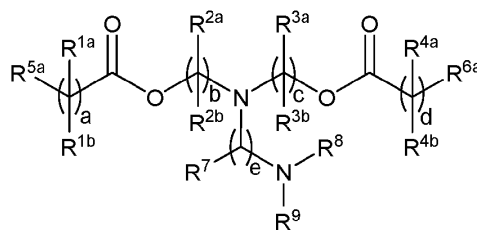
wherein R^a and R^b are, at each occurrence, independently H or a substituent. For example, in some embodiments R^a and R^b are, at each occurrence, independently H, C_1 - C_{12} alkyl or cycloalkyl, for example H or C_1 - C_{12} alkyl.

In other embodiments, the lipid compounds of Formula I have the following Formula (Ia):



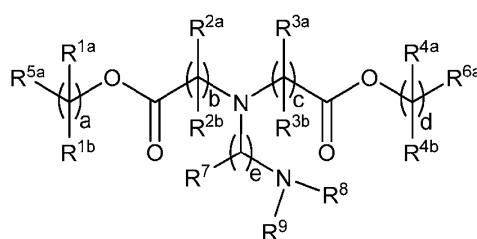
(Ia)

In other embodiments, the lipid compounds of Formula I have the following Formula (Ib):



(Ib)

In yet other embodiments, the lipid compounds of Formula I have the following Formula (Ic):



(Ic)

In certain embodiments of the lipid compound of Formula I, a, b, c and d are each independently an integer from 2 to 12 or an integer from 4 to 12. In other embodiments, a, b, c and d are each independently an integer from 8 to 12 or 5 to 9. In some certain embodiments, a is 0. In some embodiments, a is 1. In other embodiments, a is 2. In more embodiments, a is 3. In yet other embodiments, a is 4. In some embodiments, a is 5. In other embodiments, a is 6. In more embodiments, a is 7. In yet other embodiments, a is 8. In some embodiments, a is 9. In other embodiments, a is 10. In more embodiments, a is 11. In yet other embodiments, a is 12. In some embodiments, a is 13. In other embodiments, a is 14. In more embodiments, a is 15. In yet other embodiments, a is 16.

In some other embodiments of Formula I, b is 1. In other embodiments, b is 2. In more embodiments, b is 3. In yet other embodiments, b is 4. In some embodiments, b is 5. In other embodiments, b is 6. In more embodiments, b is 7. In yet other embodiments, b is 8. In some embodiments, b is 9. In other embodiments, b is 10. In more embodiments, b is 11. In yet other embodiments, b is 12. In some

embodiments, b is 13. In other embodiments, b is 14. In more embodiments, b is 15.
In yet other embodiments, b is 16.

In some more embodiments of Formula I, c is 1. In other embodiments, c is 2. In more embodiments, c is 3. In yet other embodiments, c is 4. In some
5 embodiments, c is 5. In other embodiments, c is 6. In more embodiments, c is 7. In yet other embodiments, c is 8. In some embodiments, c is 9. In other embodiments, c is 10. In more embodiments, c is 11. In yet other embodiments, c is 12. In some embodiments, c is 13. In other embodiments, c is 14. In more embodiments, c is 15. In yet other embodiments, c is 16.

10 In some certain other embodiments of Formula I, d is 0. In some embodiments, d is 1. In other embodiments, d is 2. In more embodiments, d is 3. In yet other embodiments, d is 4. In some embodiments, d is 5. In other embodiments, d is 6. In more embodiments, d is 7. In yet other embodiments, d is 8. In some embodiments, d is 9. In other embodiments, d is 10. In more embodiments, d is 11. In
15 yet other embodiments, d is 12. In some embodiments, d is 13. In other embodiments, d is 14. In more embodiments, d is 15. In yet other embodiments, d is 16.

In some other various embodiments of Formula I, a and d are the same. In some other embodiments, b and c are the same. In some other specific embodiments, a and d are the same and b and c are the same.

20 The sum of a and b and the sum of c and d in Formula I are factors which may be varied to obtain a lipid of formula I having the desired properties. In one embodiment, a and b are chosen such that their sum is an integer ranging from 14 to 24. In other embodiments, c and d are chosen such that their sum is an integer ranging from 14 to 24. In further embodiment, the sum of a and b and the sum of c and d are the
25 same. For example, in some embodiments the sum of a and b and the sum of c and d are both the same integer which may range from 14 to 24. In still more embodiments, a, b, c and d are selected such the sum of a and b and the sum of c and d is 12 or greater.

In some embodiments of Formula I, e is 1. In other embodiments, e is 2.

The substituents at R^{1a} , R^{2a} , R^{3a} and R^{4a} of Formula I are not particularly
30 limited. In certain embodiments R^{1a} , R^{2a} , R^{3a} and R^{4a} are H at each occurrence. In

certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_{12} alkyl. In certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_8 alkyl. In certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_6 alkyl. In some of the foregoing embodiments, the C_1 - C_8 alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In certain embodiments of Formula I, R^{1a} , R^{1b} , R^{4a} and R^{4b} are C_1 - C_{12} alkyl at each occurrence.

In further embodiments of Formula I, at least one of R^{1b} , R^{2b} , R^{3b} and R^{4b} is H or R^{1b} , R^{2b} , R^{3b} and R^{4b} are H at each occurrence.

In certain embodiments of Formula I, R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond. In other embodiments of the foregoing R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

The substituents at R^5 and R^6 of Formula I are not particularly limited in the foregoing embodiments. In certain embodiments one or both of R^5 or R^6 is methyl. In certain other embodiments one or both of R^5 or R^6 is cycloalkyl for example cyclohexyl. In these embodiments the cycloalkyl may be substituted or not substituted. In certain other embodiments the cycloalkyl is substituted with C_1 - C_{12} alkyl, for example tert-butyl.

The substituents at R^7 are not particularly limited in the foregoing embodiments of Formula I. In certain embodiments at least one R^7 is H. In some other embodiments, R^7 is H at each occurrence. In certain other embodiments R^7 is C_1 - C_{12} alkyl.

In certain other of the foregoing embodiments of Formula I, one of R^8 or R^9 is methyl. In other embodiments, both R^8 and R^9 are methyl.

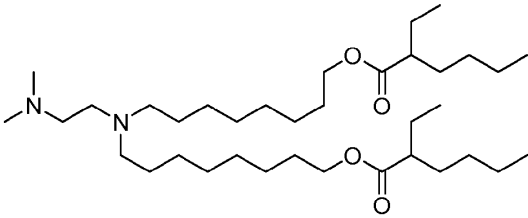
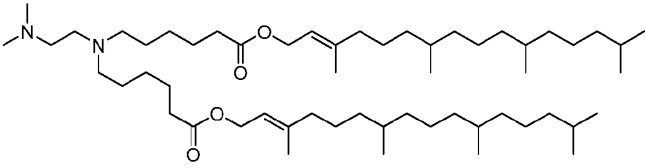
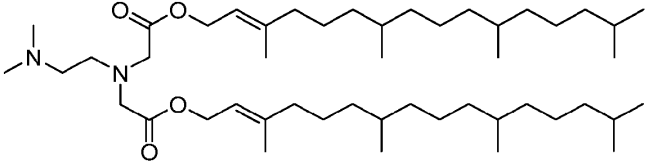
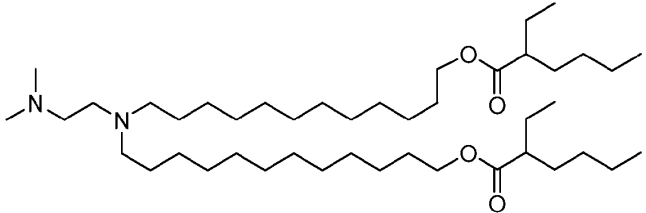
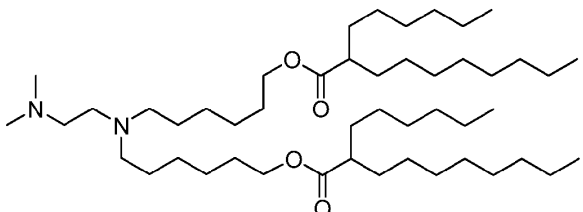
In some different embodiments of Formula I, R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring. In some embodiments of the foregoing, R^8 and R^9 , together with the nitrogen

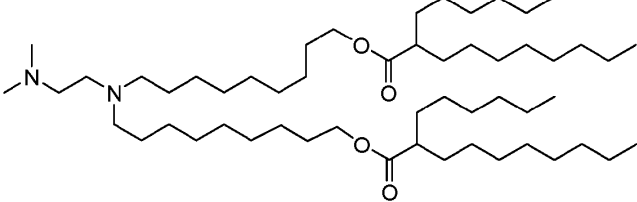
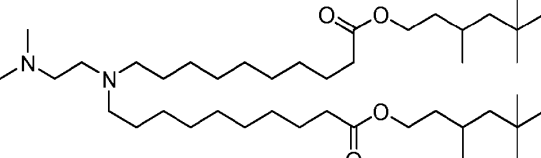
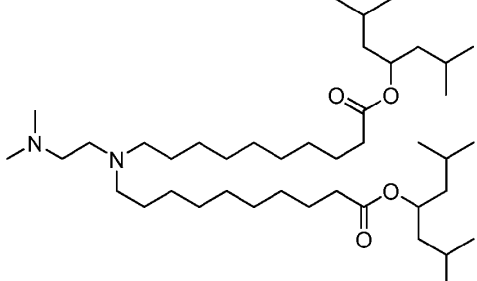
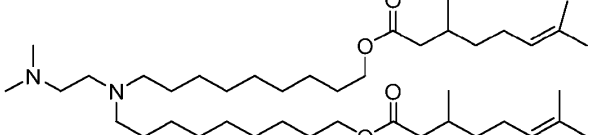
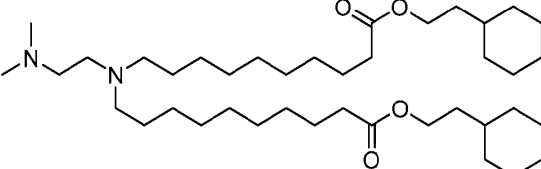
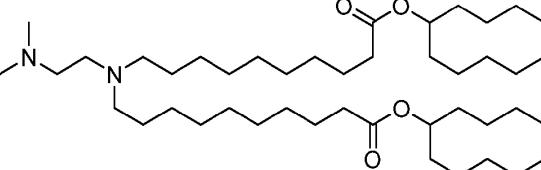
atom to which they are attached, form a 5-membered heterocyclic ring, for example a pyrrolidiny ring.

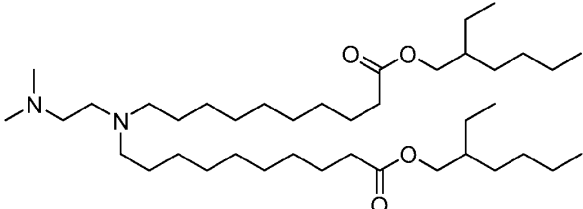
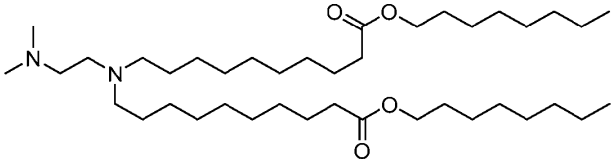
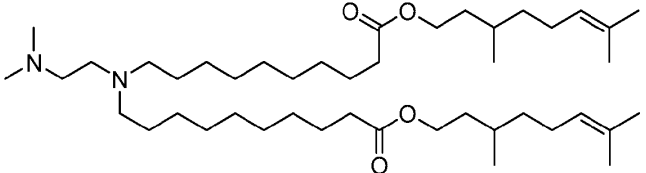
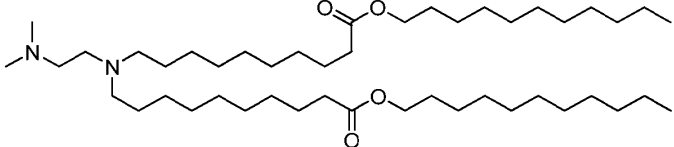
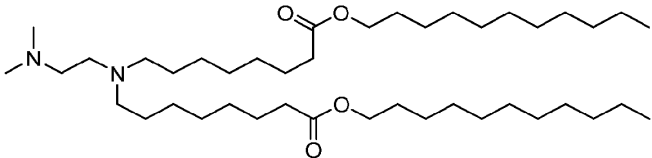
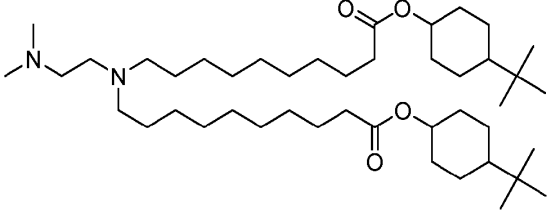
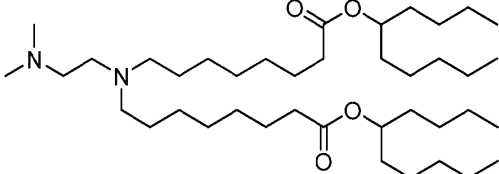
In some embodiments of Embodiment 3, the first and second cationic lipids are each, independently selected from a lipid of Formula I.

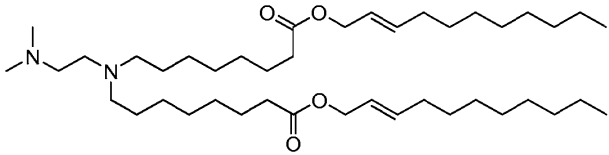
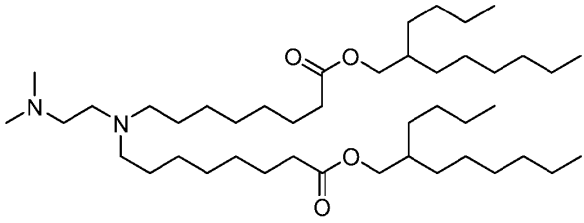
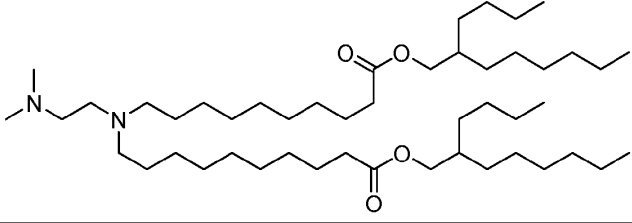
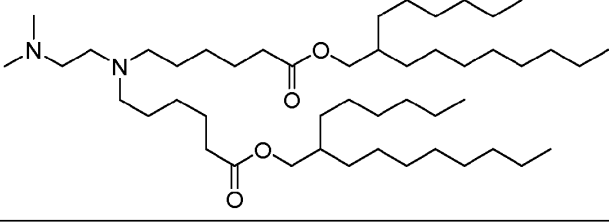
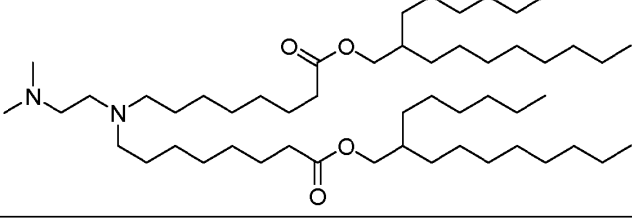
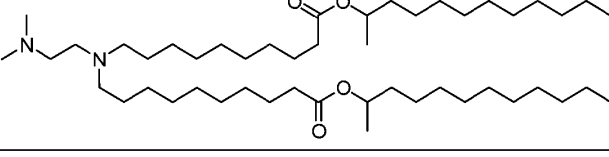
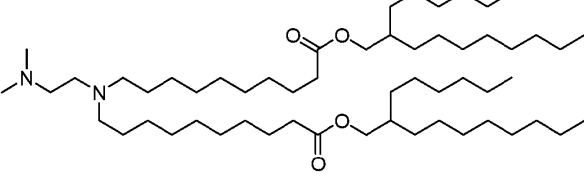
- 5 In various different embodiments, the lipid of Formula I has one of the structures set forth in Table 1 below.

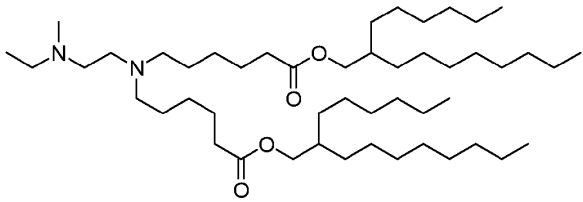
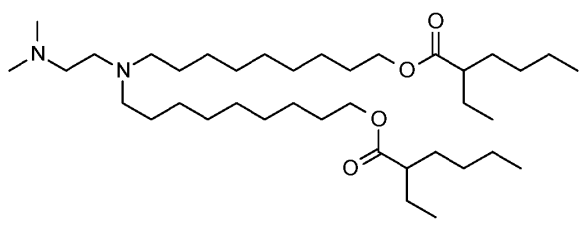
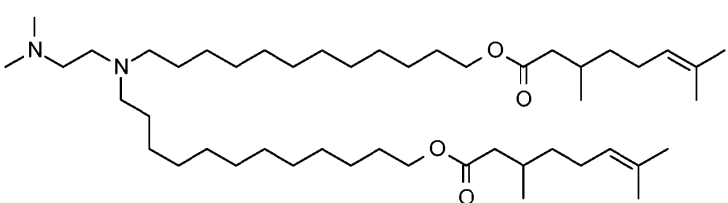
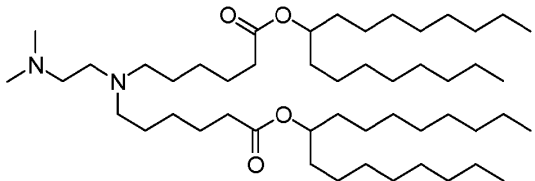
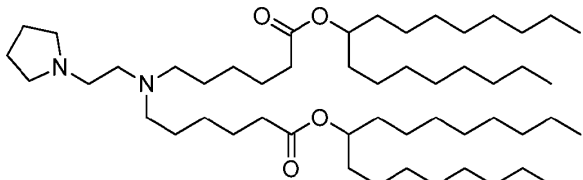
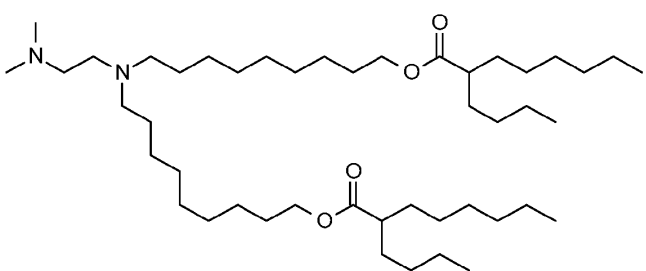
Table 1: Representative Lipids of Formula I

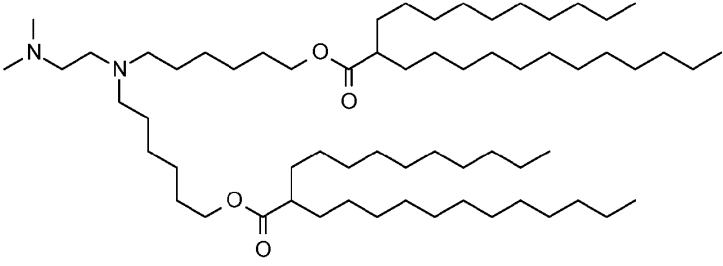
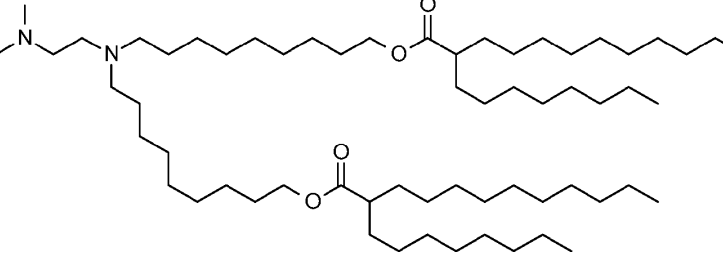
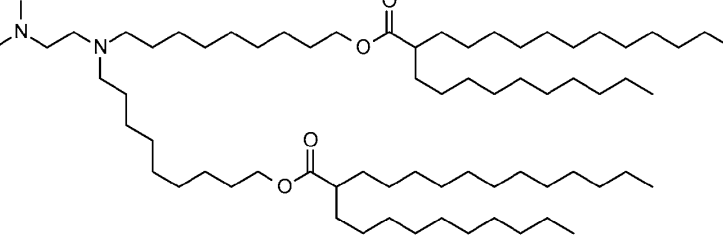
No.	Structure	pKa
I-1		-
I-2		5.64
I-3		7.15
I-4		6.43
I-5		6.28

No.	Structure	pKa
I-6		6.12
I-7		-
I-8		-
I-9		-
I-10		-
I-11		6.36

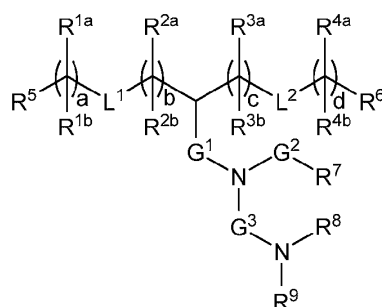
No.	Structure	pKa
I-12		-
I-13		6.51
I-14		-
I-15		6.30
I-16		6.63
I-17		-
I-18		-

No.	Structure	pKa
I-19		6.72
I-20		6.44
I-21		6.28
I-22		6.53
I-23		6.24
I-24		6.28
I-25		6.20

No.	Structure	pKa
I-33		6.27
I-34		-
I-35		6.21
I-36		-
I-37		-
I-38		6.24

No.	Structure	pKa
I-39		5.82
I-40		6.38
I-41		5.91

In some embodiments, the cationic lipid of Embodiments 1, 2, 3, 4 or 5 has a structure of Formula II:



II

or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof,
wherein:

- 5 one of L^1 or L^2 is $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $SC(=O)-$, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ or $-NR^aC(=O)O-$, and the other of L^1 or L^2 is $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $SC(=O)-$, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ or
- 10 $-NR^aC(=O)O-$ or a direct bond;
 G^1 is C_1 - C_2 alkylene, $-(C=O)-$, $-O(C=O)-$, $-SC(=O)-$, $-NR^aC(=O)-$ or a direct bond;
 G^2 is $-C(=O)-$, $-(C=O)O-$, $-C(=O)S-$, $-C(=O)NR^a-$ or a direct bond;
 G^3 is C_1 - C_6 alkylene;
- 15 R^a is H or C_1 - C_{12} alkyl;
 R^{1a} and R^{1b} are, at each occurrence, independently either: (a) H or C_1 - C_{12} alkyl; or (b) R^{1a} is H or C_1 - C_{12} alkyl, and R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond;
- 20 R^{2a} and R^{2b} are, at each occurrence, independently either: (a) H or C_1 - C_{12} alkyl; or (b) R^{2a} is H or C_1 - C_{12} alkyl, and R^{2b} together with the carbon atom to which it is bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound to form a carbon-carbon double bond;
- R^{3a} and R^{3b} are, at each occurrence, independently either (a): H or C_1 - C_{12} alkyl; or (b) R^{3a} is H or C_1 - C_{12} alkyl, and R^{3b} together with the carbon atom to which it
- 25

is bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^{4a} and R^{4b} are, at each occurrence, independently either: (a) H or C_1-C_{12} alkyl; or (b) R^{4a} is H or C_1-C_{12} alkyl, and R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^5 and R^6 are each independently H or methyl;

R^7 is C_4-C_{20} alkyl;

R^8 and R^9 are each independently C_1-C_{12} alkyl; or R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring;

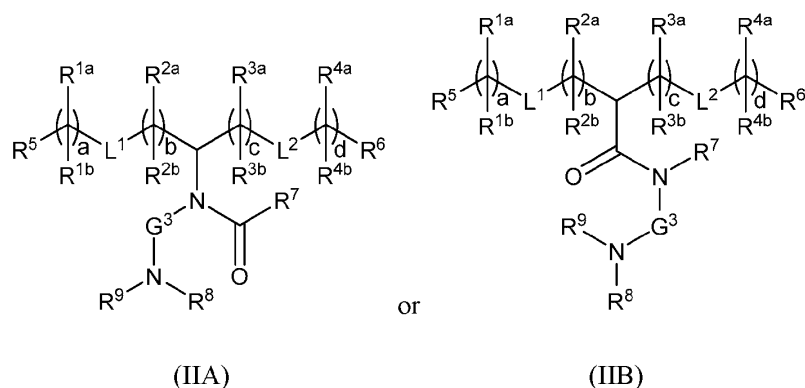
a, b, c and d are each independently an integer from 1 to 24; and

x is 0, 1 or 2.

In some embodiments of Formula (II), L^1 and L^2 are each independently $-O(C=O)-$, $-(C=O)O-$ or a direct bond. In other embodiments, G^1 and G^2 are each independently $-(C=O)-$ or a direct bond. In some different embodiments, L^1 and L^2 are each independently $-O(C=O)-$, $-(C=O)O-$ or a direct bond; and G^1 and G^2 are each independently $-(C=O)-$ or a direct bond.

In some different embodiments of Formula (II), L^1 and L^2 are each independently $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $-SC(=O)-$, $-NR^a-$, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $-NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$, $-NR^aC(=O)O-$, $-NR^aS(O)_xNR^a-$, $-NR^aS(O)_x-$ or $-S(O)_xNR^a-$.

In other of the foregoing embodiments of Formula (II), the lipid compound has one of the following Formulae (IIA) or (IIB):



In some embodiments of Formula (II), the lipid compound has Formula (IIA). In other embodiments, the lipid compound has Formula (IIB).

5 In any of the foregoing embodiments of Formula (II), one of L¹ or L² is -O(C=O)-. For example, in some embodiments each of L¹ and L² are -O(C=O)-.

In some different embodiments of Formula (II), one of L¹ or L² is -(C=O)O-. For example, in some embodiments each of L¹ and L² is -(C=O)O-.

In different embodiments of Formula (II), one of L^1 or L^2 is a direct
 10 bond. As used herein, a "direct bond" means the group (*e.g.*, L^1 or L^2) is absent. For
 example, in some embodiments each of L^1 and L^2 is a direct bond.

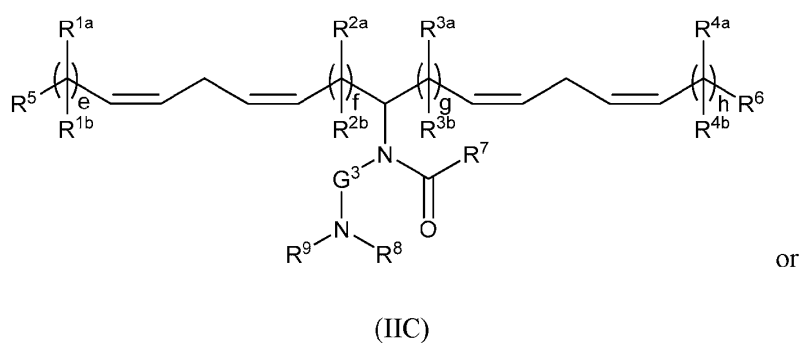
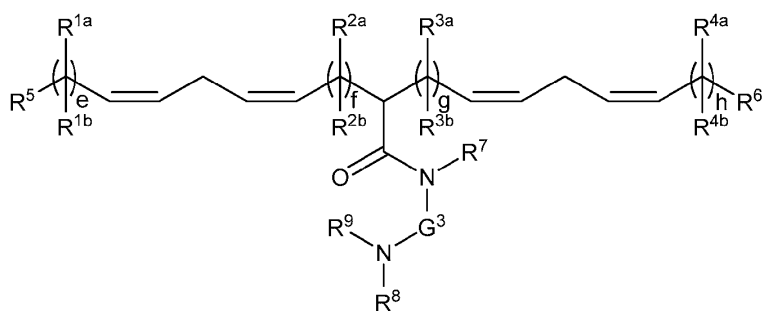
In other different embodiments of Formula (II), for at least one occurrence of R^{1a} and R^{1b} , R^{1a} is H or C_1 - C_{12} alkyl, and R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In still other different embodiments of Formula (II), for at least one occurrence of R^{4a} and R^{4b}, R^{4a} is H or C₁-C₁₂ alkyl, and R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

20 In more embodiments of Formula (II), for at least one occurrence of R^{2a} and R^{2b} , R^{2a} is H or C_1 - C_{12} alkyl, and R^{2b} together with the carbon atom to which it is bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In other different embodiments of Formula (II), for at least one occurrence of R^{3a} and R^{3b} , R^{3a} is H or C_1 - C_{12} alkyl, and R^{3b} together with the carbon atom to which it is bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

5 In various other embodiments of Formula (II), the lipid compound has one of the following Formulae (IIC) or (IID):



10

wherein e, f, g and h are each independently an integer from 1 to 12.

In some embodiments of Formula (II), the lipid compound has Formula (IIC). In other embodiments, the lipid compound has Formula (IID).

15 In various embodiments of Formulae (IIC) or (IID), e, f, g and h are each independently an integer from 4 to 10.

In certain embodiments of Formula (II), a, b, c and d are each independently an integer from 2 to 12 or an integer from 4 to 12. In other embodiments, a, b, c and d are each independently an integer from 8 to 12 or 5 to 9. In some certain embodiments, a is 0. In some embodiments, a is 1. In other embodiments,

a is 2. In more embodiments, a is 3. In yet other embodiments, a is 4. In some
embodiments, a is 5. In other embodiments, a is 6. In more embodiments, a is 7. In yet
other embodiments, a is 8. In some embodiments, a is 9. In other embodiments, a is
10. In more embodiments, a is 11. In yet other embodiments, a is 12. In some
5 embodiments, a is 13. In other embodiments, a is 14. In more embodiments, a is 15.
In yet other embodiments, a is 16.

In some embodiments of Formula (II), b is 1. In other embodiments, b is
2. In more embodiments, b is 3. In yet other embodiments, b is 4. In some
embodiments, b is 5. In other embodiments, b is 6. In more embodiments, b is 7. In
10 yet other embodiments, b is 8. In some embodiments, b is 9. In other embodiments, b
is 10. In more embodiments, b is 11. In yet other embodiments, b is 12. In some
embodiments, b is 13. In other embodiments, b is 14. In more embodiments, b is 15.
In yet other embodiments, b is 16.

In some embodiments of Formula (II), c is 1. In other embodiments, c is
15 2. In more embodiments, c is 3. In yet other embodiments, c is 4. In some
embodiments, c is 5. In other embodiments, c is 6. In more embodiments, c is 7. In yet
other embodiments, c is 8. In some embodiments, c is 9. In other embodiments, c is
10. In more embodiments, c is 11. In yet other embodiments, c is 12. In some
embodiments, c is 13. In other embodiments, c is 14. In more embodiments, c is 15.
20 In yet other embodiments, c is 16.

In some certain embodiments of Formula (II), d is 0. In some
embodiments, d is 1. In other embodiments, d is 2. In more embodiments, d is 3. In
yet other embodiments, d is 4. In some embodiments, d is 5. In other embodiments, d
is 6. In more embodiments, d is 7. In yet other embodiments, d is 8. In some
25 embodiments, d is 9. In other embodiments, d is 10. In more embodiments, d is 11. In
yet other embodiments, d is 12. In some embodiments, d is 13. In other embodiments,
d is 14. In more embodiments, d is 15. In yet other embodiments, d is 16.

In some embodiments of Formula (II), e is 1. In other embodiments, e is
2. In more embodiments, e is 3. In yet other embodiments, e is 4. In some
30 embodiments, e is 5. In other embodiments, e is 6. In more embodiments, e is 7. In yet

other embodiments, e is 8. In some embodiments, e is 9. In other embodiments, e is 10. In more embodiments, e is 11. In yet other embodiments, e is 12.

In some embodiments of Formula (II), f is 1. In other embodiments, f is 2. In more embodiments, f is 3. In yet other embodiments, f is 4. In some 5 embodiments, f is 5. In other embodiments, f is 6. In more embodiments, f is 7. In yet other embodiments, f is 8. In some embodiments, f is 9. In other embodiments, f is 10. In more embodiments, f is 11. In yet other embodiments, f is 12.

In some embodiments of Formula (II), g is 1. In other embodiments, g is 2. In more embodiments, g is 3. In yet other embodiments, g is 4. In some 10 embodiments, g is 5. In other embodiments, g is 6. In more embodiments, g is 7. In yet other embodiments, g is 8. In some embodiments, g is 9. In other embodiments, g is 10. In more embodiments, g is 11. In yet other embodiments, g is 12.

In some embodiments of Formula (II), h is 1. In other embodiments, e is 2. In more embodiments, h is 3. In yet other embodiments, h is 4. In some 15 embodiments, e is 5. In other embodiments, h is 6. In more embodiments, h is 7. In yet other embodiments, h is 8. In some embodiments, h is 9. In other embodiments, h is 10. In more embodiments, h is 11. In yet other embodiments, h is 12.

In some other various embodiments of Formula (II), a and d are the same. In some other embodiments, b and c are the same. In some other specific 20 embodiments and a and d are the same and b and c are the same.

The sum of a and b and the sum of c and d of Formula (II) are factors which may be varied to obtain a lipid having the desired properties. In one embodiment, a and b are chosen such that their sum is an integer ranging from 14 to 24. In other embodiments, c and d are chosen such that their sum is an integer ranging from 25 14 to 24. In further embodiment, the sum of a and b and the sum of c and d are the same. For example, in some embodiments the sum of a and b and the sum of c and d are both the same integer which may range from 14 to 24. In still more embodiments, a, b, c and d are selected such that the sum of a and b and the sum of c and d is 12 or greater.

The substituents at R^{1a} , R^{2a} , R^{3a} and R^{4a} of Formula (II) are not particularly limited. In some embodiments, at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is H. In certain embodiments R^{1a} , R^{2a} , R^{3a} and R^{4a} are H at each occurrence. In certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_{12} alkyl. In certain other
 5 embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_8 alkyl. In certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_6 alkyl. In some of the foregoing embodiments, the C_1 - C_8 alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In certain embodiments of Formula (II), R^{1a} , R^{1b} , R^{4a} and R^{4b} are C_1 - C_{12}
 10 alkyl at each occurrence.

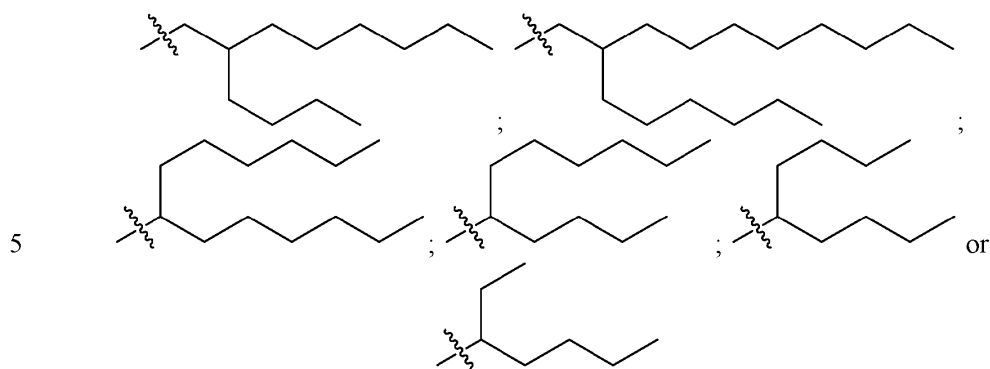
In further embodiments of Formula (II), at least one of R^{1b} , R^{2b} , R^{3b} and R^{4b} is H or R^{1b} , R^{2b} , R^{3b} and R^{4b} are H at each occurrence.

In certain embodiments of Formula (II), R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to
 15 which it is bound to form a carbon-carbon double bond. In other embodiments of the foregoing R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

The substituents at R^5 and R^6 of Formula (II) are not particularly limited
 20 in the foregoing embodiments. In certain embodiments one of R^5 or R^6 is methyl. In other embodiments each of R^5 or R^6 is methyl.

The substituents at R^7 of Formula (II) are not particularly limited in the foregoing embodiments. In certain embodiments R^7 is C_6 - C_{16} alkyl. In some other embodiments, R^7 is C_6 - C_9 alkyl. In some of these embodiments, R^7 is substituted
 25 with $-(C=O)OR^b$, $-O(C=O)R^b$, $-C(=O)R^b$, $-OR^b$, $-S(O)_xR^b$, $-S-SR^b$, $-C(=O)SR^b$, $-SC(=O)R^b$, $-NR^aR^b$, $-NR^aC(=O)R^b$, $-C(=O)NR^aR^b$, $-NR^aC(=O)NR^aR^b$, $-OC(=O)NR^aR^b$, $-NR^aC(=O)OR^b$, $-NR^aS(O)_xNR^aR^b$, $-NR^aS(O)_xR^b$ or $-S(O)_xNR^aR^b$, wherein: R^a is H or C_1 - C_{12} alkyl; R^b is C_1 - C_{15} alkyl; and x is 0, 1 or 2. For example, in some embodiments R^7 is substituted with $-(C=O)OR^b$ or $-O(C=O)R^b$.

In some of the foregoing embodiments of Formula (II), R^b is branched C_1 - C_{16} alkyl. For example, in some embodiments R^b has one of the following structures:



In certain other of the foregoing embodiments of Formula (II), one of R^8 or R^9 is methyl. In other embodiments, both R^8 and R^9 are methyl.

In some different embodiments of Formula (II), R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5, 6 or 7-membered heterocyclic ring. In some embodiments of the foregoing, R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 5-membered heterocyclic ring, for example a pyrrolidinyl ring. In some different embodiments of the foregoing, R^8 and R^9 , together with the nitrogen atom to which they are attached, form a 6-membered heterocyclic ring, for example a piperazinyl ring.

10

15

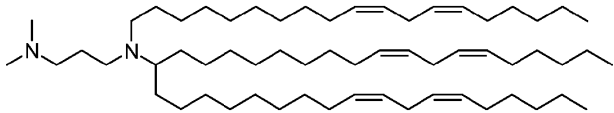
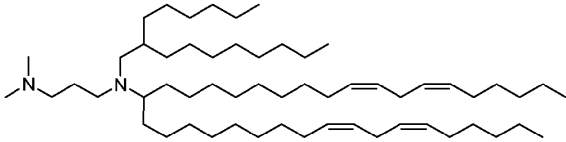
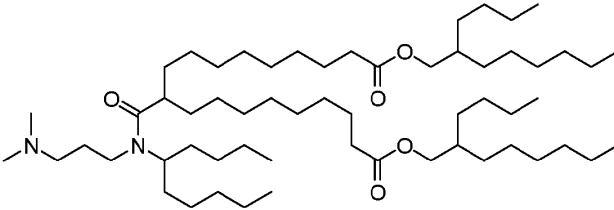
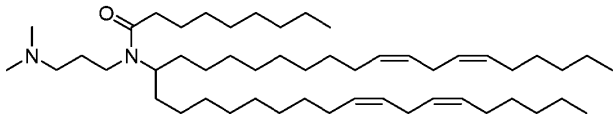
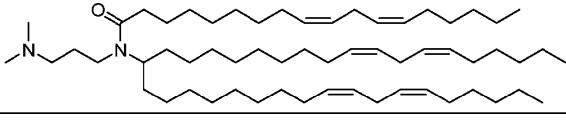
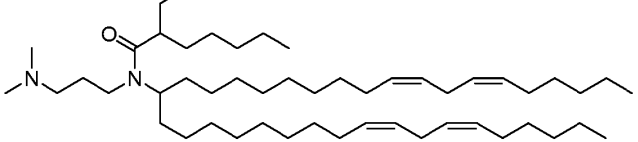
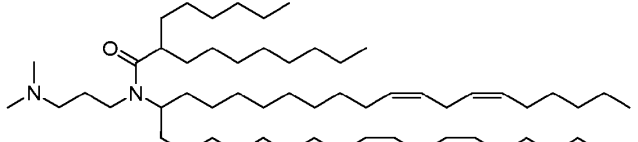
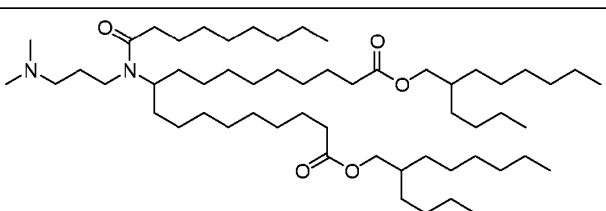
In certain embodiments of Embodiment 3, the first and second cationic lipids are each, independently selected from a lipid of Formula II.

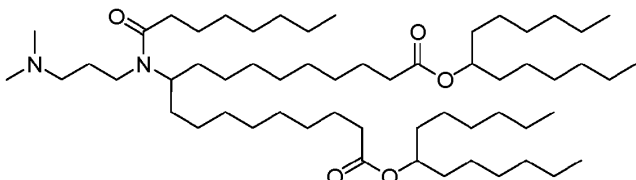
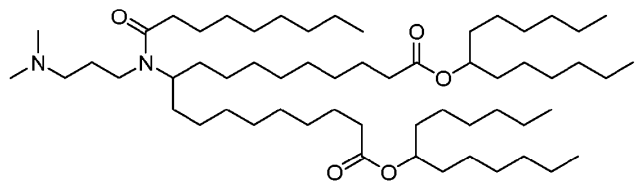
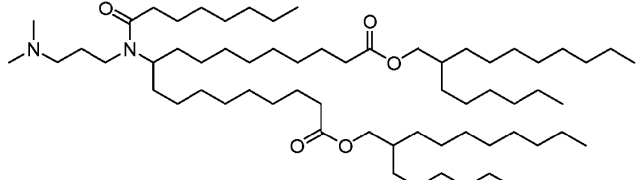
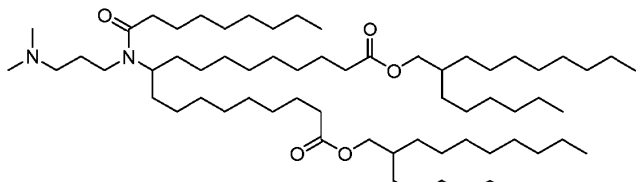
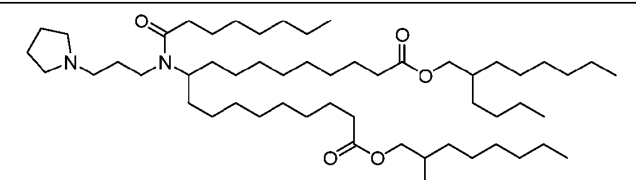
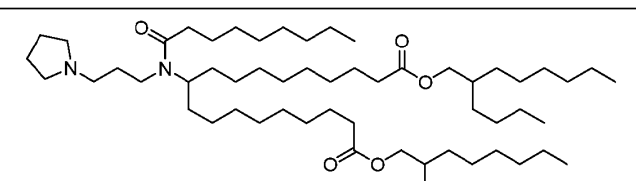
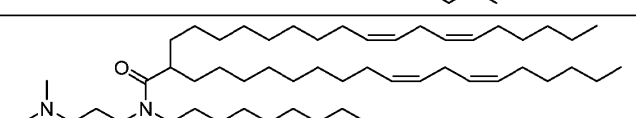
In still other embodiments of the foregoing lipids of Formula (II), G^3 is C_2 - C_4 alkylene, for example C_3 alkylene. In various different embodiments, the lipid compound has one of the structures set forth in Table 2 below

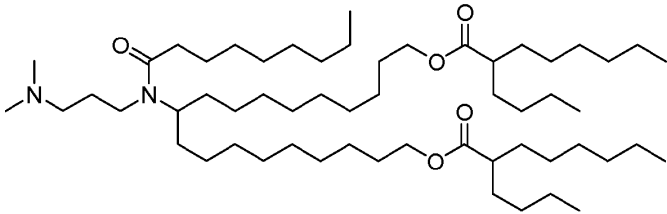
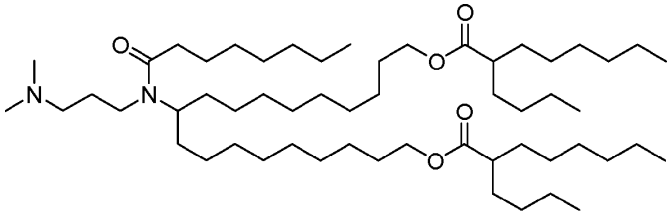
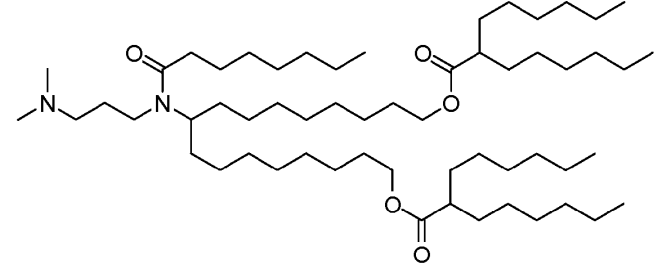
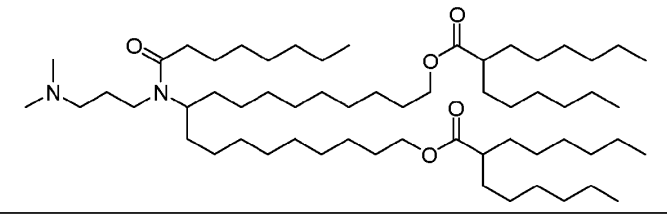
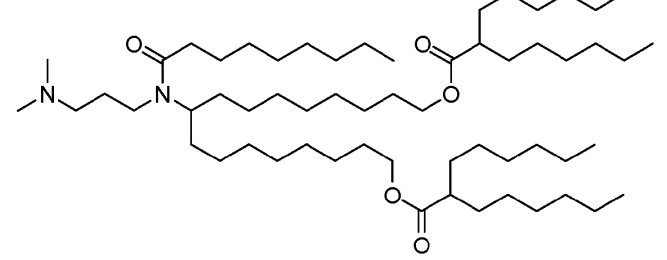
20

Table 2: Representative Lipids of Formula (II)

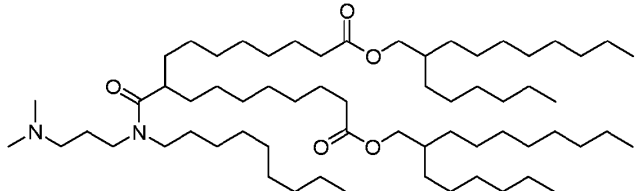
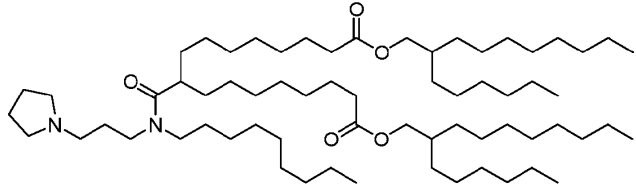
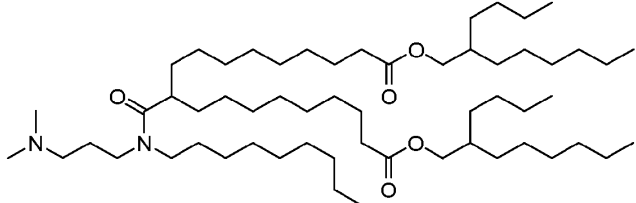
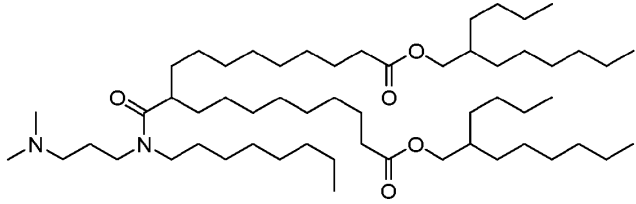
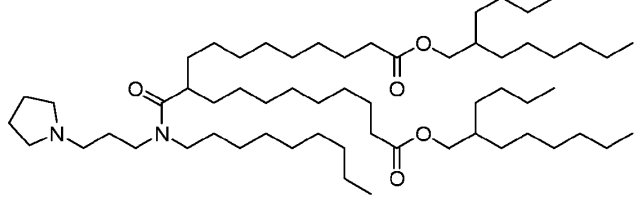
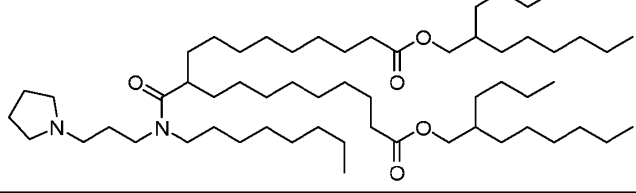
No.	Structure	pKa
II-1		5.64

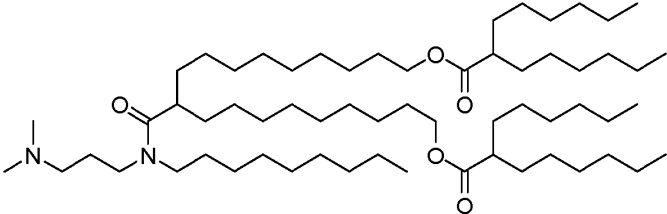
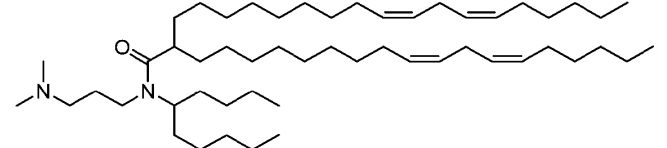
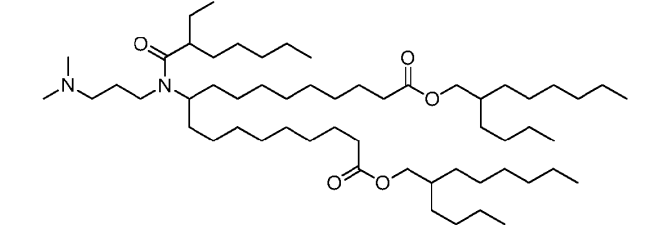
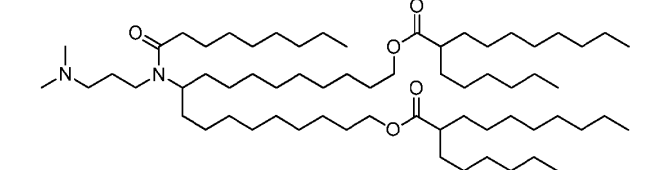
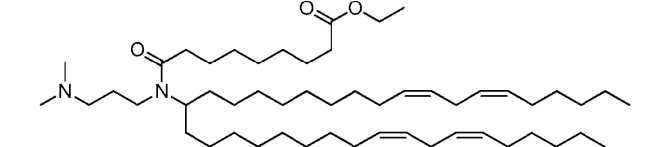
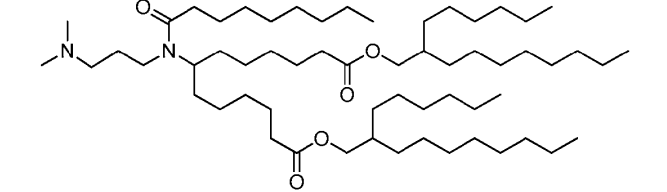
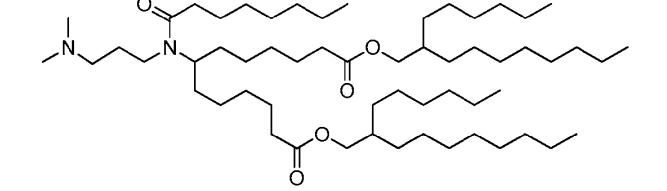
No.	Structure	pKa
II-2		-
II-3		-
II-4		-
II-5		6.27
II-6		6.14
II-7		5.93
II-8		5.35
II-9		6.27

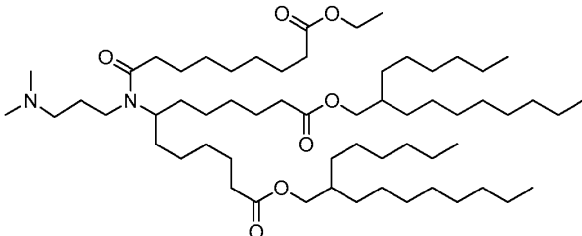
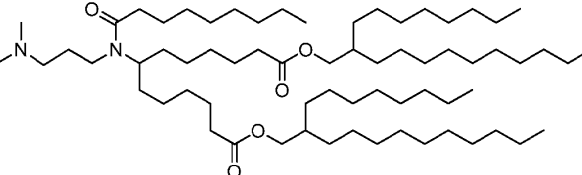
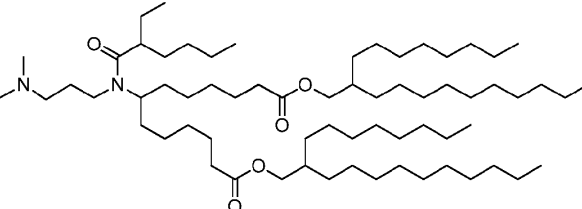
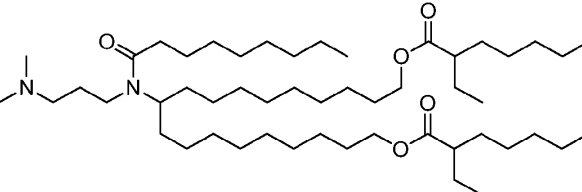
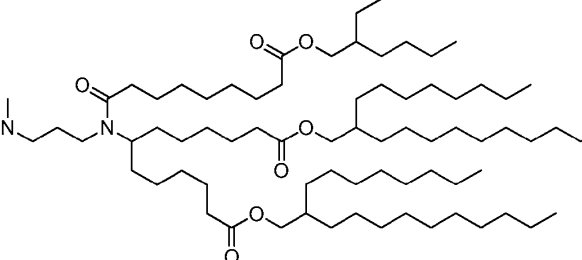
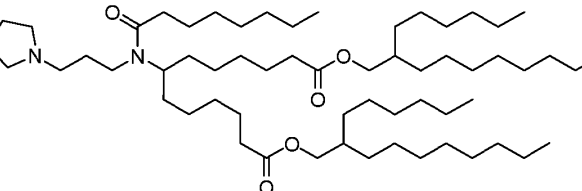
No.	Structure	pKa
II-10		6.16
II-11		6.13
II-12		6.21
II-13		6.22
II-14		6.33
II-15		6.32
II-16		6.37

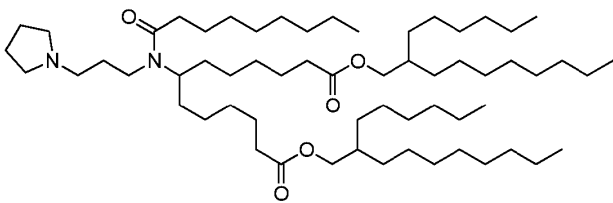
No.	Structure	pKa
II-17		6.27
II-18		-
II-19		-
II-20		-
II-21		-

No.	Structure	pKa
II-22		-
II-23		-
II-24		6.14
II-25		-
II-26		-

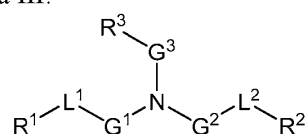
No.	Structure	pKa
II-27		-
II-28		-
II-29		-
II-30		-
II-31		-
II-32		-

No.	Structure	pKa
II-33		-
II-34		-
II-35		5.97
II-36		6.13
II-37		5.61
II-38		6.45
II-39		6.45

No.	Structure	pKa
II-40		6.57
II-41		-
II-42		-
II-43		-
II-44		-
II-45		-

No.	Structure	pKa
II-46		-

In some other embodiments, the cationic lipid of Embodiments 1, 2, 3, 4 or 5 has a structure of Formula III:



III

or a pharmaceutically acceptable salt, prodrug or stereoisomer thereof, wherein:

one of L^1 or L^2 is $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})-$, $-\text{O}-$, $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $\text{SC}(\text{=O})-$, $-\text{NR}^a\text{C}(\text{=O})-$, $-\text{C}(\text{=O})\text{NR}^a-$, $\text{NR}^a\text{C}(\text{=O})\text{NR}^a-$, $-\text{OC}(\text{=O})\text{NR}^a-$ or $-\text{NR}^a\text{C}(\text{=O})\text{O}-$, and the other of L^1 or L^2 is $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})-$, $-\text{O}-$, $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $\text{SC}(\text{=O})-$, $-\text{NR}^a\text{C}(\text{=O})-$, $-\text{C}(\text{=O})\text{NR}^a-$, $\text{NR}^a\text{C}(\text{=O})\text{NR}^a-$, $-\text{OC}(\text{=O})\text{NR}^a-$ or $-\text{NR}^a\text{C}(\text{=O})\text{O}-$ or a direct bond;

G^1 and G^2 are each independently unsubstituted C_1 - C_{12} alkylene or C_1 - C_{12} alkenylene;

G^3 is C_1 - C_{24} alkylene, C_1 - C_{24} alkenylene, C_3 - C_8 cycloalkylene, C_3 - C_8 cycloalkenylene;

R^a is H or C_1 - C_{12} alkyl;

R^1 and R^2 are each independently C_6 - C_{24} alkyl or C_6 - C_{24} alkenyl;

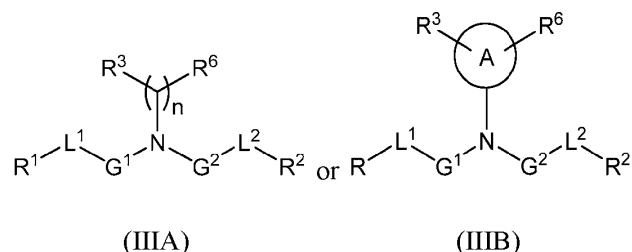
R^3 is H, OR^5 , CN, $-\text{C}(\text{=O})\text{OR}^4$, $-\text{OC}(\text{=O})\text{R}^4$ or $-\text{NR}^5\text{C}(\text{=O})\text{R}^4$;

R^4 is C_1 - C_{12} alkyl;

R^5 is H or C_1 - C_6 alkyl; and

x is 0, 1 or 2.

In some of the foregoing embodiments of Formula (III), the lipid has one of the following Formulae (IIIA) or (IIIB):



5 wherein:

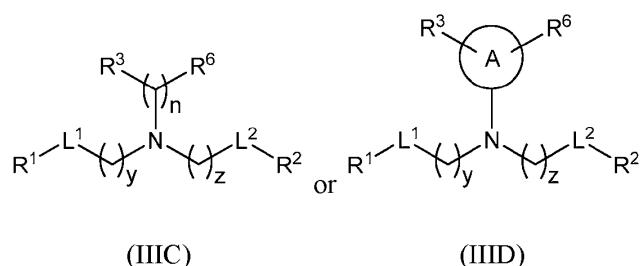
A is a 3 to 8-membered cycloalkyl or cycloalkylene ring;

R⁶ is, at each occurrence, independently H, OH or C₁-C₂₄ alkyl;

n is an integer ranging from 1 to 15.

In some of the foregoing embodiments of Formula (III), the lipid has
10 Formula (IIIA), and in other embodiments, the lipid has Formula (IIIB).

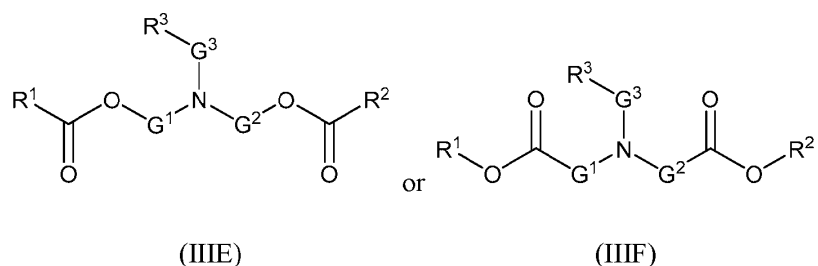
In other embodiments of Formula (III), the lipid has one of the following
Formulae (IIIC) or (IIID):



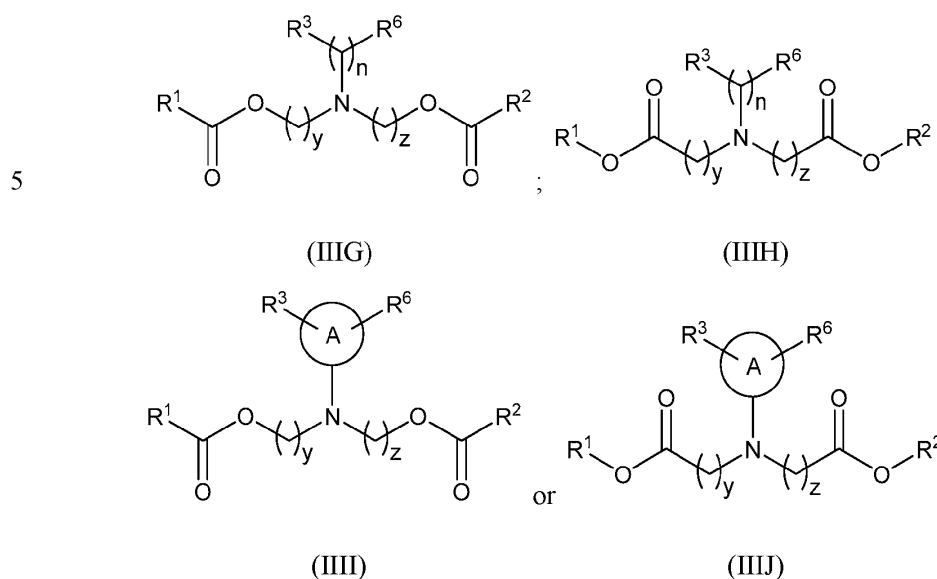
15 wherein y and z are each independently integers ranging from 1 to 12.

In any of the foregoing embodiments of Formula (III), one of L¹ or L²
is -O(C=O)-. For example, in some embodiments each of L¹ and L² are -O(C=O)-. In
some different embodiments of any of the foregoing, L¹ and L² are each
independently -(C=O)O- or -O(C=O)-. For example, in some embodiments each of L¹
20 and L² is -(C=O)O-.

In some different embodiments of Formula (III), the lipid has one of the
following Formulae (IIIE) or (IIIF):



In some of the foregoing embodiments of Formula (III), the lipid has one of the following Formulae (IIIG), (IIIH), (IIII), or (IIIJ):



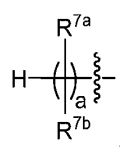
In some of the foregoing embodiments of Formula (III), n is an integer ranging from 2 to 12, for example from 2 to 8 or from 2 to 4. For example, in some embodiments, n is 3, 4, 5 or 6. In some embodiments, n is 3. In some embodiments, n is 4. In some embodiments, n is 5. In some embodiments, n is 6.

In some other of the foregoing embodiments of Formula (III), y and z are each independently an integer ranging from 2 to 10. For example, in some embodiments, y and z are each independently an integer ranging from 4 to 9 or from 4 to 6.

In some of the foregoing embodiments of Formula (III), R^6 is H. In other of the foregoing embodiments, R^6 is C_1 - C_{24} alkyl. In other embodiments, R^6 is OH.

In some embodiments of Formula (III), G^3 is unsubstituted. In other embodiments, G^3 is substituted. In various different embodiments, G^3 is linear C_1 - C_{24} alkylene or linear C_1 - C_{24} alkenylene.

In some other foregoing embodiments of Formula (III), R^1 or R^2 , or both, is C_6 - C_{24} alkenyl. For example, in some embodiments, R^1 and R^2 each, independently have the following structure:



wherein:

R^{7a} and R^{7b} are, at each occurrence, independently H or C_1 - C_{12} alkyl;

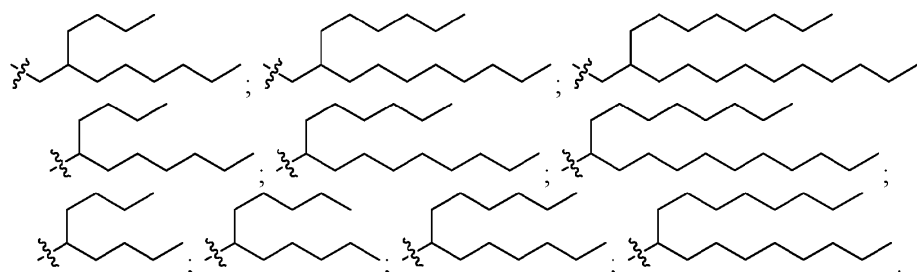
and

a is an integer from 2 to 12,

wherein R^{7a} , R^{7b} and a are each selected such that R^1 and R^2 each independently comprise from 6 to 20 carbon atoms. For example, in some embodiments a is an integer ranging from 5 to 9 or from 8 to 12.

In some of the foregoing embodiments of Formula (III), at least one occurrence of R^{7a} is H. For example, in some embodiments, R^{7a} is H at each occurrence. In other different embodiments of the foregoing, at least one occurrence of R^{7b} is C_1 - C_8 alkyl. For example, in some embodiments, C_1 - C_8 alkyl is methyl, ethyl, n-propyl, isopropyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In different embodiments of Formula (III), R^1 or R^2 , or both, has one of the following structures:

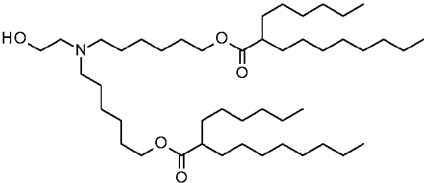
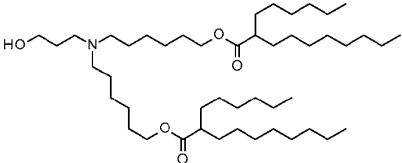
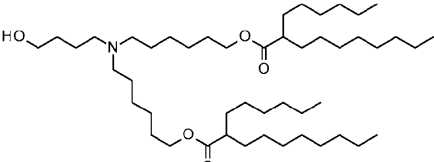
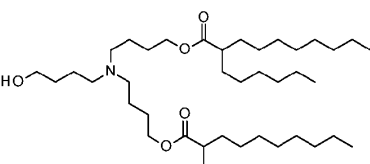


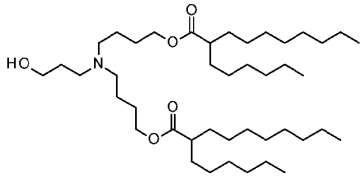
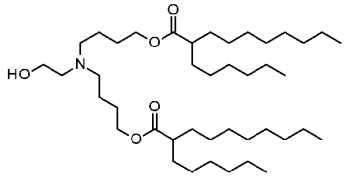
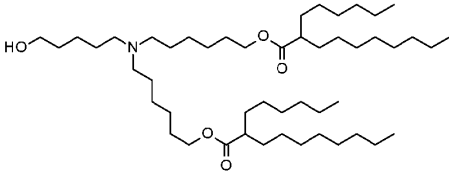
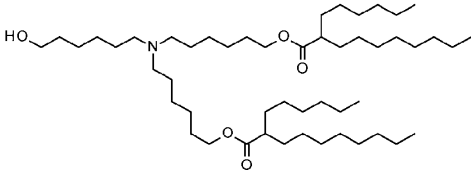
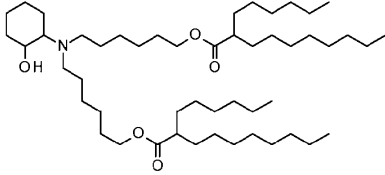
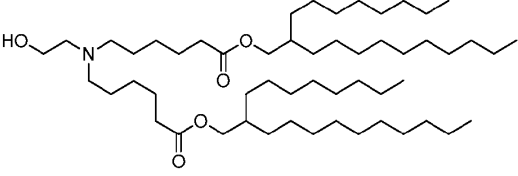
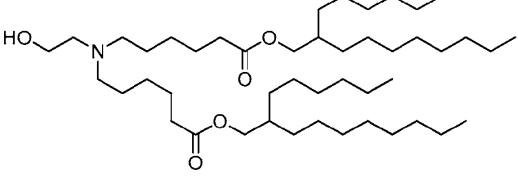
In some of the foregoing embodiments of Formula (III), R^3 is OH, CN, $-C(=O)OR^4$, $-OC(=O)R^4$ or $-NHC(=O)R^4$. In some embodiments, R^4 is methyl or ethyl.

In some specific embodiments of Embodiment 3, the first and second
5 cationic lipids are each, independently selected from a lipid of Formula III.

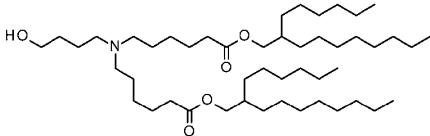
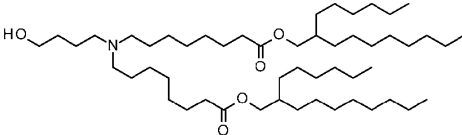
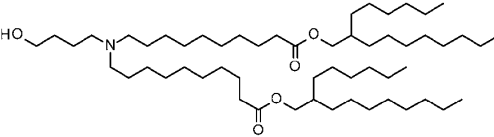
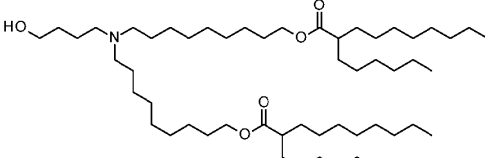
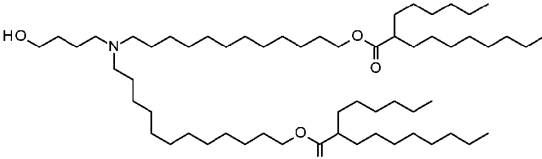
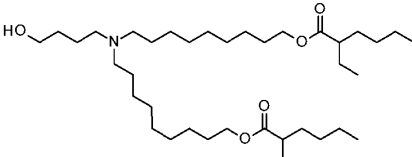
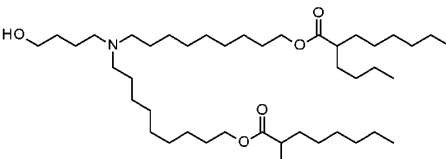
In various different embodiments, a cationic lipid of any one of the disclosed embodiments (*e.g.*, the cationic lipid, the first cationic lipid, the second cationic lipid) of Formula (III) has one of the structures set forth in Table 3 below.

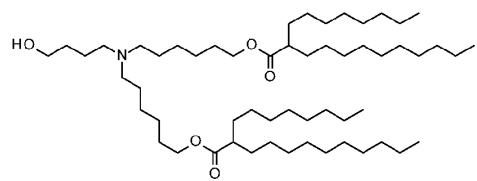
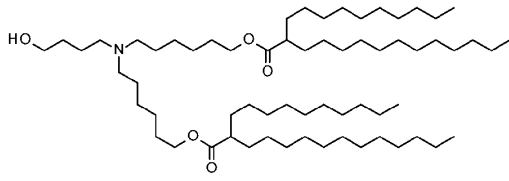
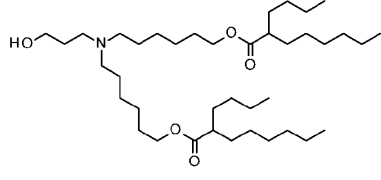
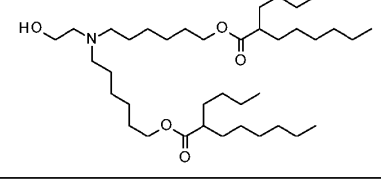
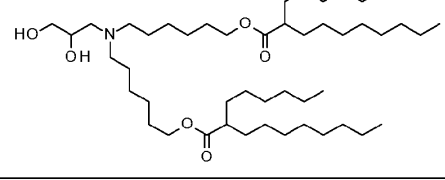
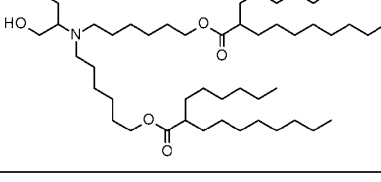
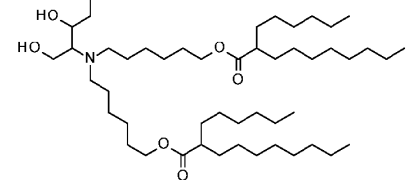
Table 3: Representative Compounds of Formula (III)

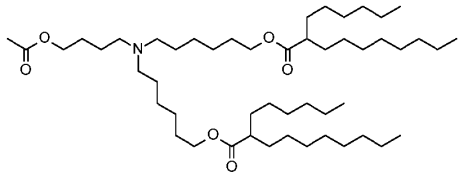
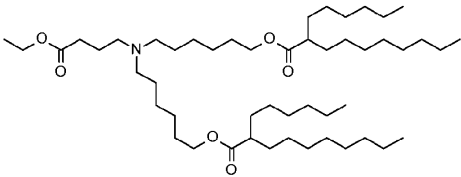
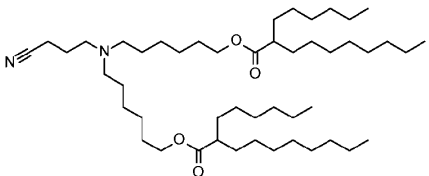
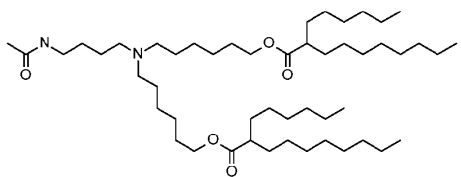
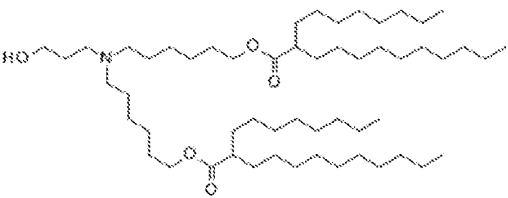
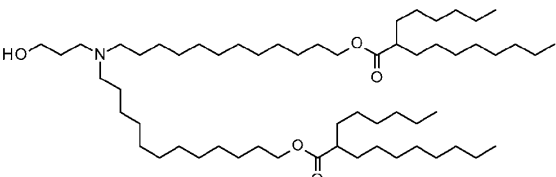
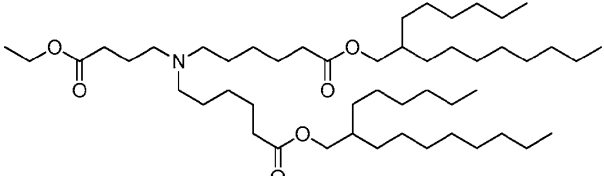
No.	Structure	pKa
III-1		5.89
III-2		6.05
III-3		6.09
III-4		5.60

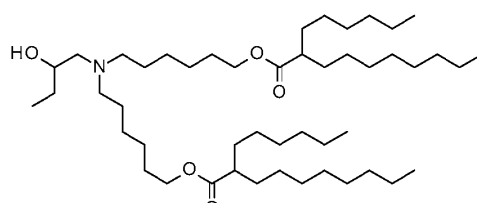
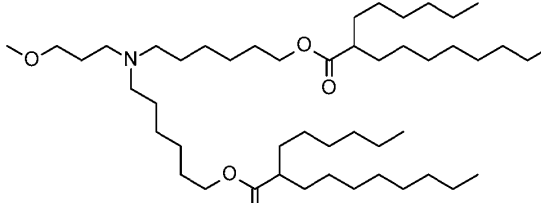
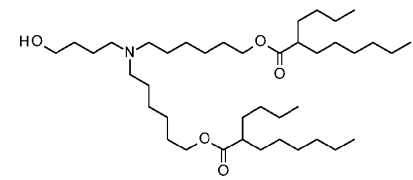
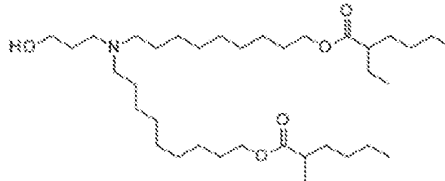
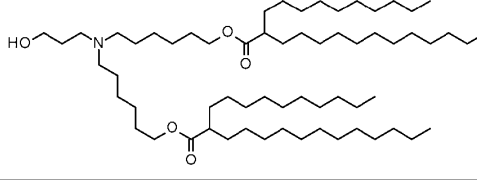
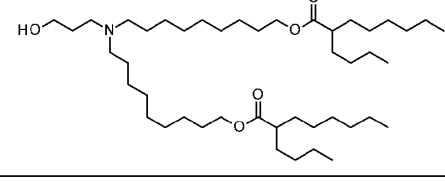
No.	Structure	pKa
III-5		5.59
III-6		5.42
III-7		6.11
III-8		5.84
III-9		-
III-10		-
III-11		-

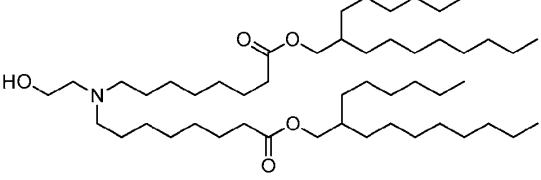
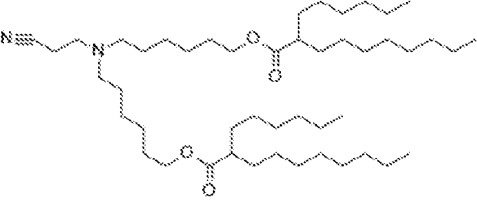
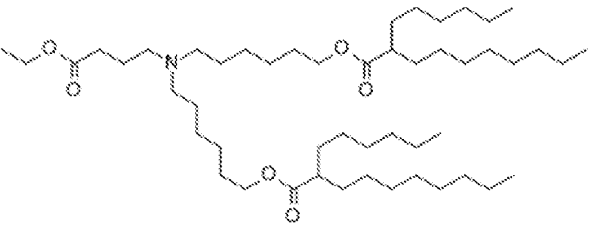
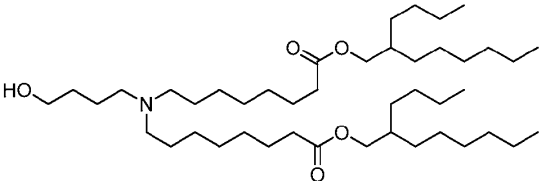
No.	Structure	pKa
III-12		-
III-13		-
III-14		-
III-15		6.14
III-16		6.31
III-17		6.28
III-18		-

No.	Structure	pKa
III-19		-
III-20		6.36
III-21		-
III-22		6.10
III-23		5.98
III-24		-
III-25		6.22

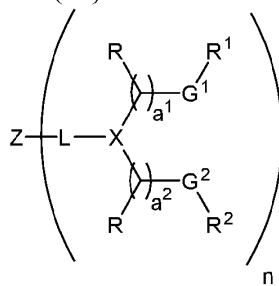
No.	Structure	pKa
III-26		5.84
III-27		5.77
III-28		-
III-29		-
III-30		6.09
III-31		-
III-32		-

No.	Structure	pKa
III-33		-
III-34		-
III-35		-
III-36		-
III-37		-
III-38		-
III-39		-

No.	Structure	pKa
III-40		-
III-41		-
III-42		-
III-43		-
III-44		-
III-45		-

No.	Structure	pKa
III-46		-
III-47		-
III-48		-
III-49		-

In one embodiment, the cationic lipid of any one of Embodiments 1, 2, 3, 4 or 5 has a structure of Formula (IV):



5

or a pharmaceutically acceptable salt, prodrug or stereoisomer thereof, wherein:

one of G^1 or G^2 is, at each occurrence, $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_y-$, $-S-S-$, $-C(=O)S-$, $SC(=O)-$, $-N(R^a)C(=O)-$, $-C(=O)N(R^a)-$, $-N(R^a)C(=O)N(R^a)-$, $-OC(=O)N(R^a)-$ or $-N(R^a)C(=O)O-$, and the other of G^1 or G^2 is, at each occurrence, $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_y-$, $-S-S-$, $-C(=O)S-$, $-SC(=O)-$, $-N(R^a)C(=O)-$, $-C(=O)N(R^a)-$, $-N(R^a)C(=O)N(R^a)-$, $-OC(=O)N(R^a)-$ or $-N(R^a)C(=O)O-$ or a direct bond;

L is, at each occurrence, $\sim O(C=O)-$, wherein $\sim$ represents a covalent bond to X;

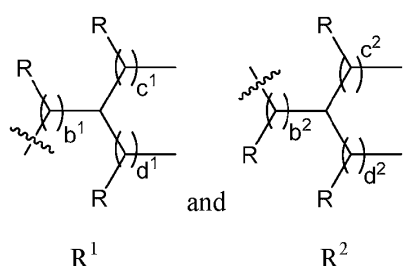
X is CR^a ;

Z is alkyl, cycloalkyl or a monovalent moiety comprising at least one polar functional group when n is 1; or Z is alkylene, cycloalkylene or a polyvalent moiety comprising at least one polar functional group when n is greater than 1;

R^a is, at each occurrence, independently H, C_1 - C_{12} alkyl, C_1 - C_{12} hydroxylalkyl, C_1 - C_{12} aminoalkyl, C_1 - C_{12} alkylaminylalkyl, C_1 - C_{12} alkoxyalkyl, C_1 - C_{12} alkoxycarbonyl, C_1 - C_{12} alkylcarbonyloxy, C_1 - C_{12} alkylcarbonyloxyalkyl or C_1 - C_{12} alkylcarbonyl;

R is, at each occurrence, independently either: (a) H or C_1 - C_{12} alkyl; or (b) R together with the carbon atom to which it is bound is taken together with an adjacent R and the carbon atom to which it is bound to form a carbon-carbon double bond;

R^1 and R^2 have, at each occurrence, the following structure, respectively:



a^1 and a^2 are, at each occurrence, independently an integer from 3 to 12;
 b^1 and b^2 are, at each occurrence, independently 0 or 1;
 c^1 and c^2 are, at each occurrence, independently an integer from 5 to 10;
 d^1 and d^2 are, at each occurrence, independently an integer from 5 to 10;

y is, at each occurrence, independently an integer from 0 to 2; and

n is an integer from 1 to 6,

wherein each alkyl, alkylene, hydroxylalkyl, aminoalkyl,

alkylaminylalkyl, alkoxyalkyl, alkoxycarbonyl, alkylcarbonyloxy,

5 alkylcarbonyloxyalkyl and alkylcarbonyl is optionally substituted with one or more substituent.

In some embodiments of Formula (IV), G^1 and G^2 are each independently

$-O(C=O)-$ or $-(C=O)O-$.

10 In other embodiments of Formula (IV), X is CH.

In different embodiments of Formula (IV), the sum of $a^1 + b^1 + c^1$ or the sum of $a^2 + b^2 + c^2$ is an integer from 12 to 26.

In still other embodiments of Formula (IV), a^1 and a^2 are independently an integer from 3 to 10. For example, in some embodiments a^1 and a^2 are

15 independently an integer from 4 to 9.

In various embodiments of Formula (IV), b^1 and b^2 are 0. In different embodiments, b^1 and b^2 are 1.

In more embodiments of Formula (IV), c^1 , c^2 , d^1 and d^2 are independently an integer from 6 to 8.

20 In other embodiments of Formula (IV), c^1 and c^2 are, at each occurrence, independently an integer from 6 to 10, and d^1 and d^2 are, at each occurrence, independently an integer from 6 to 10.

In other embodiments of Formula (IV), c^1 and c^2 are, at each occurrence, independently an integer from 5 to 9, and d^1 and d^2 are, at each occurrence,

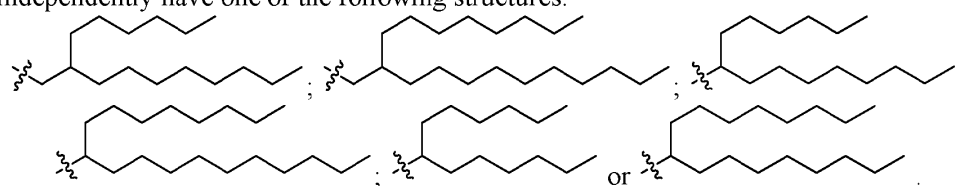
25 independently an integer from 5 to 9.

In more embodiments of Formula (IV), Z is alkyl, cycloalkyl or a monovalent moiety comprising at least one polar functional group when n is 1. In other embodiments, Z is alkyl.

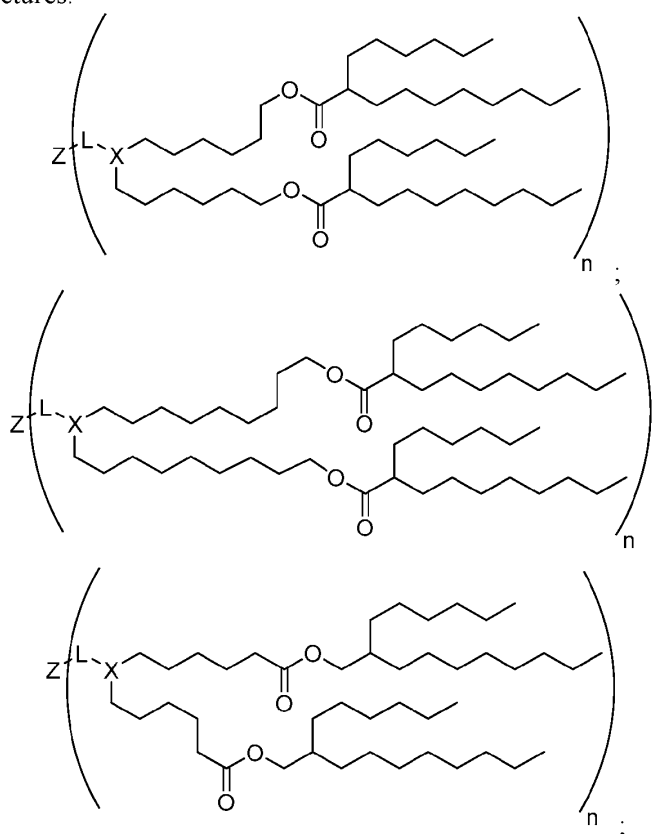
In various embodiments of the foregoing Formula (IV), R is, at each
30 occurrence, independently either: (a) H or methyl; or (b) R together with the carbon

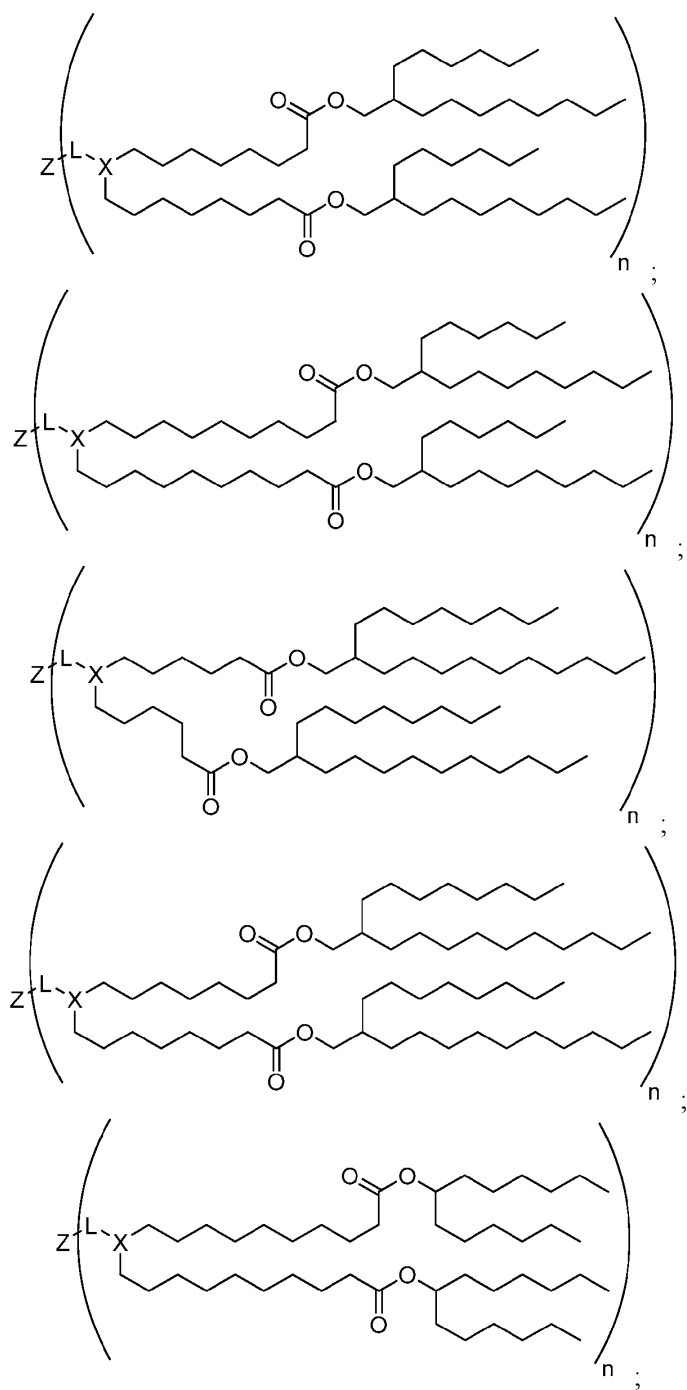
atom to which it is bound is taken together with an adjacent R and the carbon atom to which it is bound to form a carbon-carbon double bond. In certain embodiments, each R is H. In other embodiments at least one R together with the carbon atom to which it is bound is taken together with an adjacent R and the carbon atom to which it is bound to form a carbon-carbon double bond.

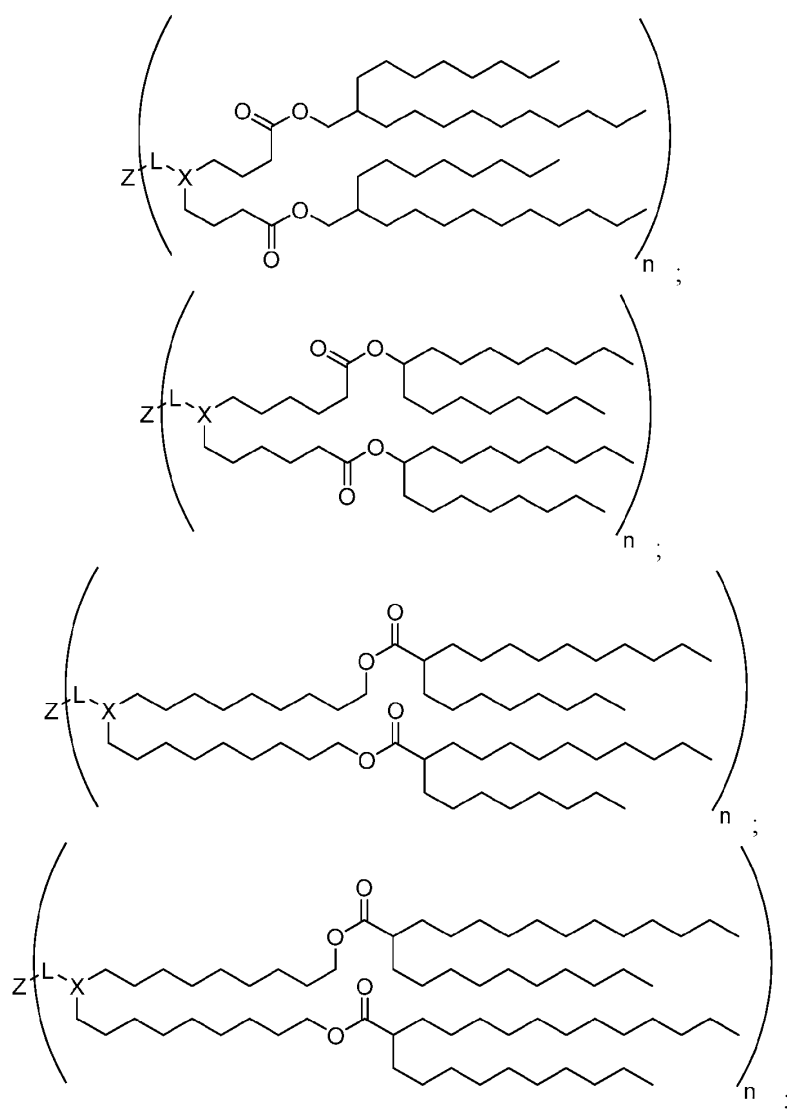
In other embodiments of the compound of Formula (IV), R^1 and R^2 independently have one of the following structures:

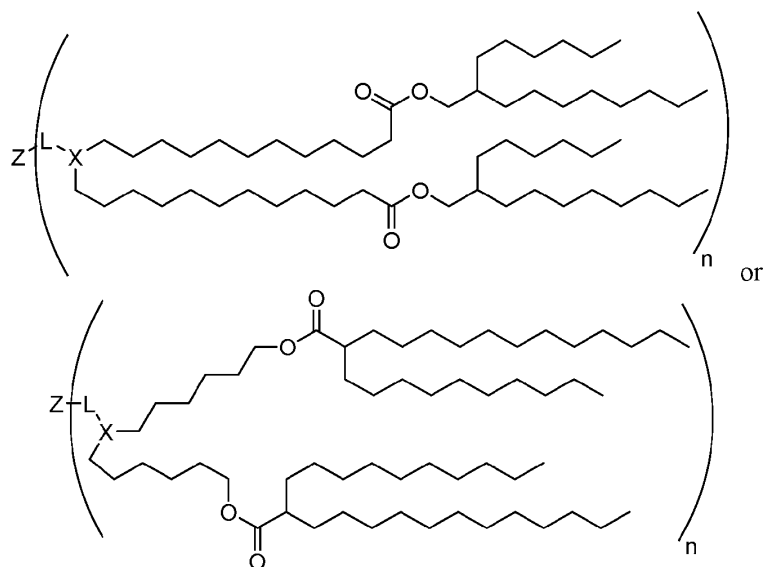


In certain embodiments of Formula (IV), the compound has one of the following structures:

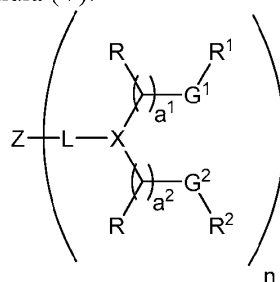








In still different embodiments the cationic lipid of Embodiments 1, 2, 3, 4 or 5 has the structure of Formula (V):



5

(V)

or a pharmaceutically acceptable salt, prodrug or stereoisomer thereof, wherein:

one of G^1 or G^2 is, at each occurrence, $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})-$, $-\text{O}-$, $-\text{S}(\text{O})_y-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $\text{SC}(\text{=O})-$, $-\text{N}(\text{R}^a)\text{C}(\text{=O})-$, $-\text{C}(\text{=O})\text{N}(\text{R}^a)-$, $-\text{N}(\text{R}^a)\text{C}(\text{=O})\text{N}(\text{R}^a)-$, $-\text{OC}(\text{=O})\text{N}(\text{R}^a)-$ or $-\text{N}(\text{R}^a)\text{C}(\text{=O})\text{O}-$, and the other of G^1 or G^2 is, at each occurrence, $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})-$, $-\text{O}-$, $-\text{S}(\text{O})_y-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $-\text{SC}(\text{=O})-$, $-\text{N}(\text{R}^a)\text{C}(\text{=O})-$, $-\text{C}(\text{=O})\text{N}(\text{R}^a)-$, $-\text{N}(\text{R}^a)\text{C}(\text{=O})\text{N}(\text{R}^a)-$, $-\text{OC}(\text{=O})\text{N}(\text{R}^a)-$ or $-\text{N}(\text{R}^a)\text{C}(\text{=O})\text{O}-$ or a direct bond;

L is, at each occurrence, $\sim\text{O}(\text{C}=\text{O})-$, wherein $\sim$ represents a covalent bond to X;

X is CR^a ;

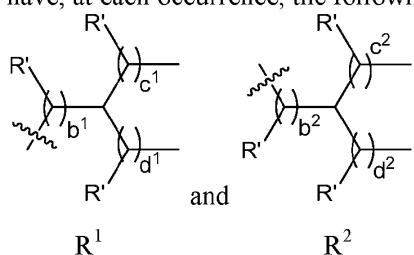
15

Z is alkyl, cycloalkyl or a monovalent moiety comprising at least one polar functional group when n is 1; or Z is alkylene, cycloalkylene or a polyvalent moiety comprising at least one polar functional group when n is greater than 1;

R^a is, at each occurrence, independently H, C₁-C₁₂ alkyl, C₁-C₁₂ hydroxylalkyl, C₁-C₁₂ aminoalkyl, C₁-C₁₂ alkylaminylalkyl, C₁-C₁₂ alkoxyalkyl, C₁-C₁₂ alkoxycarbonyl, C₁-C₁₂ alkylcarbonyloxy, C₁-C₁₂ alkylcarbonyloxyalkyl or C₁-C₁₂ alkylcarbonyl;

R is, at each occurrence, independently either: (a) H or C₁-C₁₂ alkyl; or (b) R together with the carbon atom to which it is bound is taken together with an adjacent R and the carbon atom to which it is bound to form a carbon-carbon double bond;

R¹ and R² have, at each occurrence, the following structure, respectively:



R' is, at each occurrence, independently H or C₁-C₁₂ alkyl;
 a¹ and a² are, at each occurrence, independently an integer from 3 to 12;
 b¹ and b² are, at each occurrence, independently 0 or 1;
 c¹ and c² are, at each occurrence, independently an integer from 2 to 12;
 d¹ and d² are, at each occurrence, independently an integer from 2 to 12;
 y is, at each occurrence, independently an integer from 0 to 2; and
 n is an integer from 1 to 6,
 wherein a¹, a², c¹, c², d¹ and d² are selected such that the sum of a¹+c¹+d¹ is an integer from 18 to 30, and the sum of a²+c²+d² is an integer from 18 to 30, and wherein each alkyl, alkylene, hydroxylalkyl, aminoalkyl, alkylaminylalkyl, alkoxyalkyl, alkoxycarbonyl, alkylcarbonyloxy, alkylcarbonyloxyalkyl and alkylcarbonyl is optionally substituted with one or more substituent.

In certain embodiments of Formula (V), G^1 and G^2 are each independently

$-O(C=O)-$ or $-(C=O)O-$.

In other embodiments of Formula (V), X is CH.

5 In some embodiments of Formula (V), the sum of $a^1+c^1+d^1$ is an integer from 20 to 30, and the sum of $a^2+c^2+d^2$ is an integer from 18 to 30. In other embodiments, the sum of $a^1+c^1+d^1$ is an integer from 20 to 30, and the sum of $a^2+c^2+d^2$ is an integer from 20 to 30. In more embodiments of Formula (V), the sum of $a^1+b^1+c^1$ or the sum of $a^2+b^2+c^2$ is an integer from 12 to 26. In other embodiments, a^1 , a^2 ,
10 c^1 , c^2 , d^1 and d^2 are selected such that the sum of $a^1+c^1+d^1$ is an integer from 18 to 28, and the sum of $a^2+c^2+d^2$ is an integer from 18 to 28,

In still other embodiments of Formula (V), a^1 and a^2 are independently an integer from 3 to 10, for example an integer from 4 to 9.

In yet other embodiments of Formula (V), b^1 and b^2 are 0. In different
15 embodiments b^1 and b^2 are 1.

In certain other embodiments of Formula (V), c^1 , c^2 , d^1 and d^2 are independently an integer from 6 to 8.

In different other embodiments of Formula (V), Z is alkyl or a monovalent moiety comprising at least one polar functional group when n is 1; or Z is
20 alkylene or a polyvalent moiety comprising at least one polar functional group when n is greater than 1.

In more embodiments of Formula (V), Z is alkyl, cycloalkyl or a monovalent moiety comprising at least one polar functional group when n is 1. In other
embodiments, Z is alkyl.

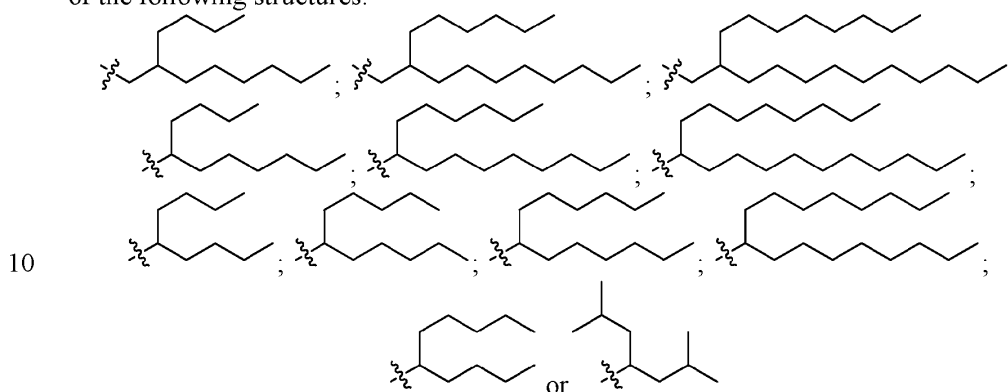
25 In other different embodiments of Formula (V), R is, at each occurrence, independently either: (a) H or methyl; or (b) R together with the carbon atom to which it is bound is taken together with an adjacent R and the carbon atom to which it is bound to form a carbon-carbon double bond. For example in some embodiments each R is H. In other embodiments at least one R together with the carbon atom to which it

is bound is taken together with an adjacent R and the carbon atom to which it is bound to form a carbon-carbon double bond.

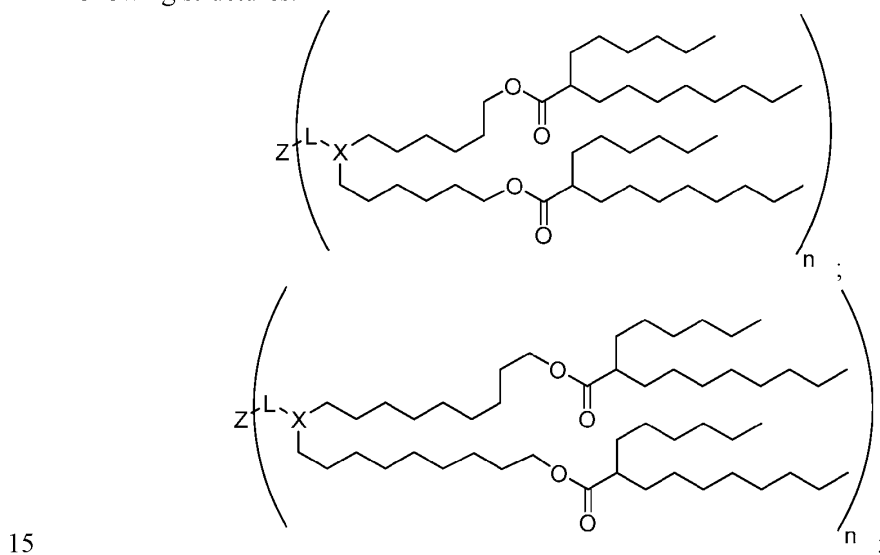
In more embodiments, each R' is H.

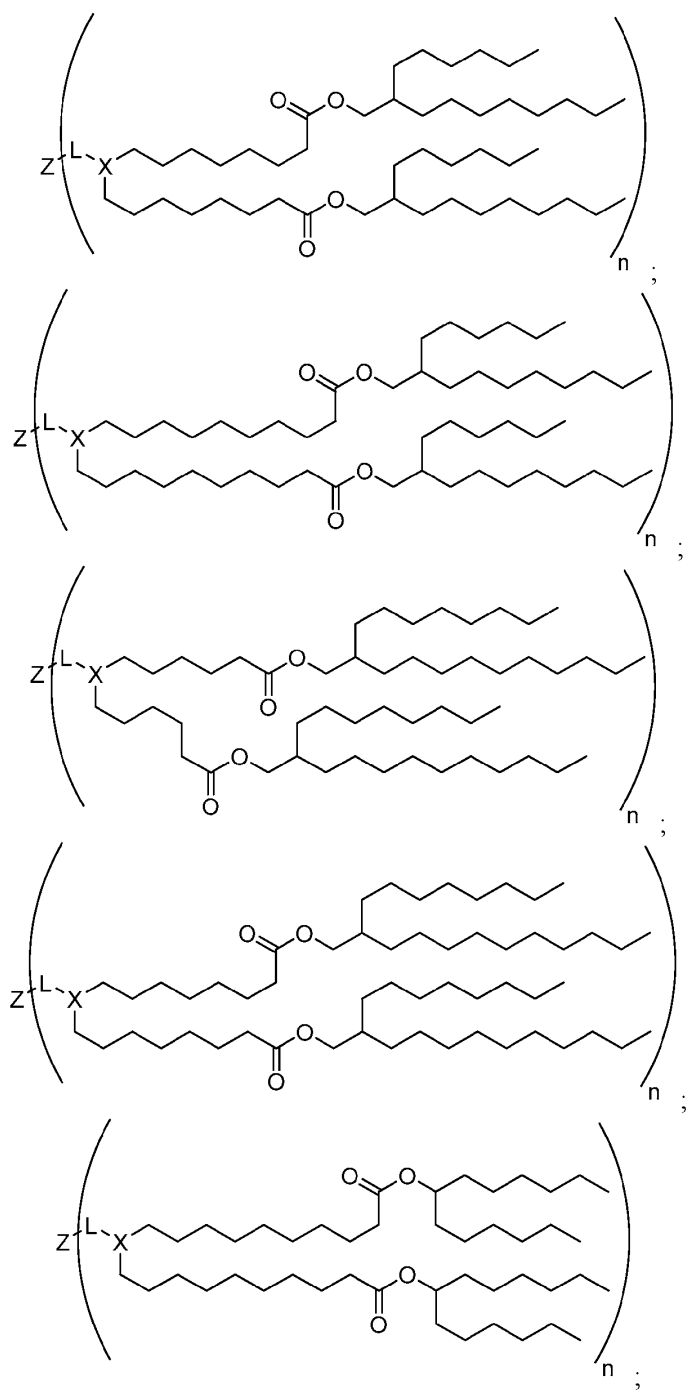
In certain embodiments of Formula (V), the sum of $a^1+c^1+d^1$ is an integer from 20 to 25, and the sum of $a^2+c^2+d^2$ is an integer from 20 to 25.

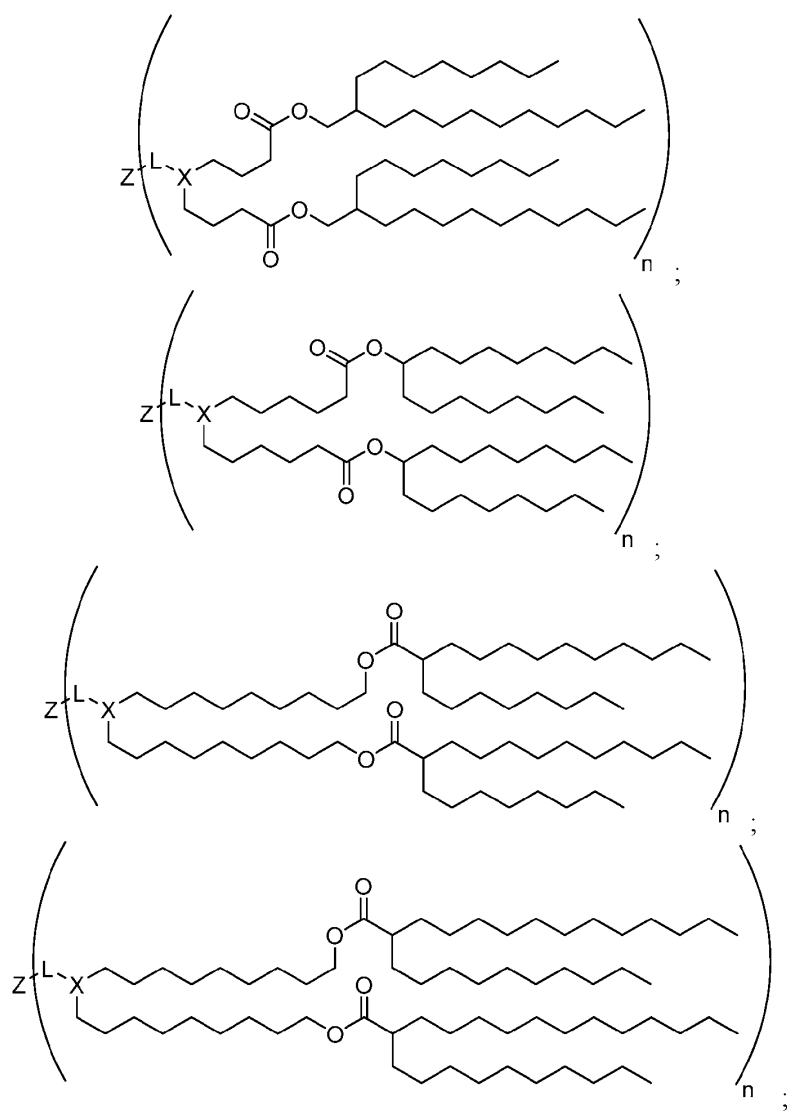
In other embodiments of Formula (V), R¹ and R² independently have one of the following structures:

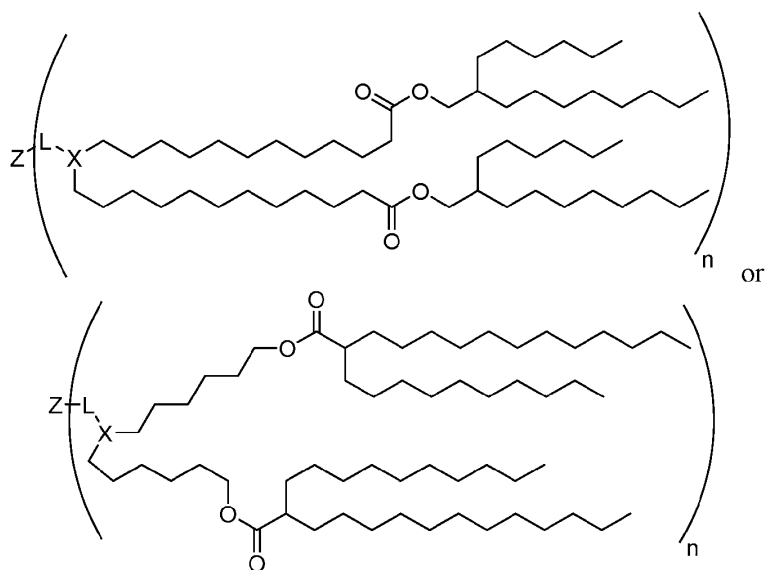


In more embodiments of Formula (V), the compound has one of the following structures:









In any of the foregoing embodiments of Formula (IV) or (V), n is 1. In other of the foregoing embodiments of Formula (IV) or (V), n is greater than 1.

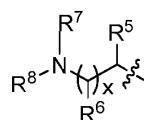
5 In more of any of the foregoing embodiments of Formula (IV) or (V), Z is a mono- or polyvalent moiety comprising at least one polar functional group. In some embodiments, Z is a monovalent moiety comprising at least one polar functional group. In other embodiments, Z is a polyvalent moiety comprising at least one polar functional group.

10 In more of any of the foregoing embodiments of Formula (IV) or (V), the polar functional group is a hydroxyl, alkoxy, ester, cyano, amide, amino, alkylaminyl, heterocyclyl or heteroaryl functional group.

In any of the foregoing embodiments of Formula (IV) or (V), Z is hydroxyl, hydroxylalkyl, alkoxyalkyl, amino, aminoalkyl, alkylaminyl,

15 alkylaminylalkyl, heterocyclyl or heterocyclalkyl.

In some other embodiments of Formula (IV) or (V), Z has the following structure:



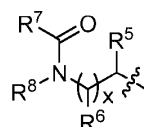
wherein:

R^5 and R^6 are independently H or C_1-C_6 alkyl;

R^7 and R^8 are independently H or C_1-C_6 alkyl or R^7 and R^8 , together with the nitrogen atom to which they are attached, join to form a 3-7 membered heterocyclic ring; and

5 x is an integer from 0 to 6.

In still different embodiments of Formula (IV) or (V), Z has the following structure:



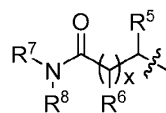
wherein:

10 R^5 and R^6 are independently H or C_1-C_6 alkyl;

R^7 and R^8 are independently H or C_1-C_6 alkyl or R^7 and R^8 , together with the nitrogen atom to which they are attached, join to form a 3-7 membered heterocyclic ring; and

x is an integer from 0 to 6.

15 In still different embodiments of formula (IV) or (V), Z has the following structure:



wherein:

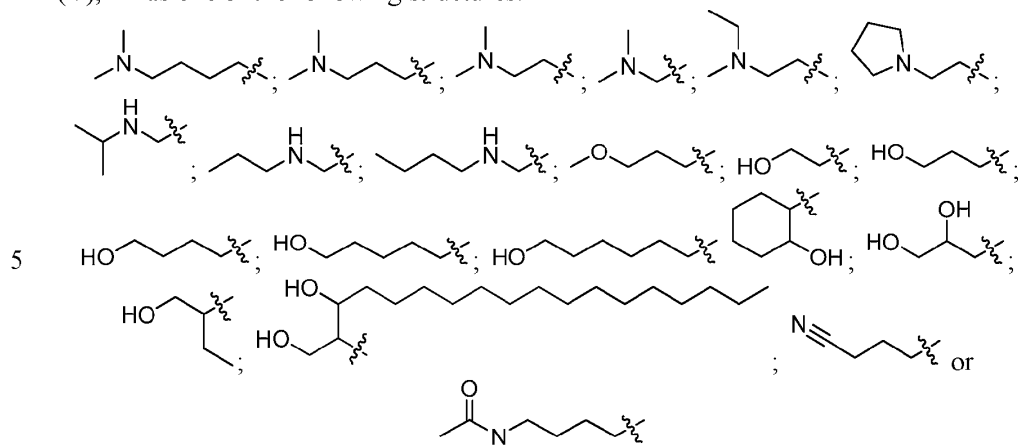
R^5 and R^6 are independently H or C_1-C_6 alkyl;

20 R^7 and R^8 are independently H or C_1-C_6 alkyl or R^7 and R^8 , together with the nitrogen atom to which they are attached, join to form a 3-7 membered heterocyclic ring; and

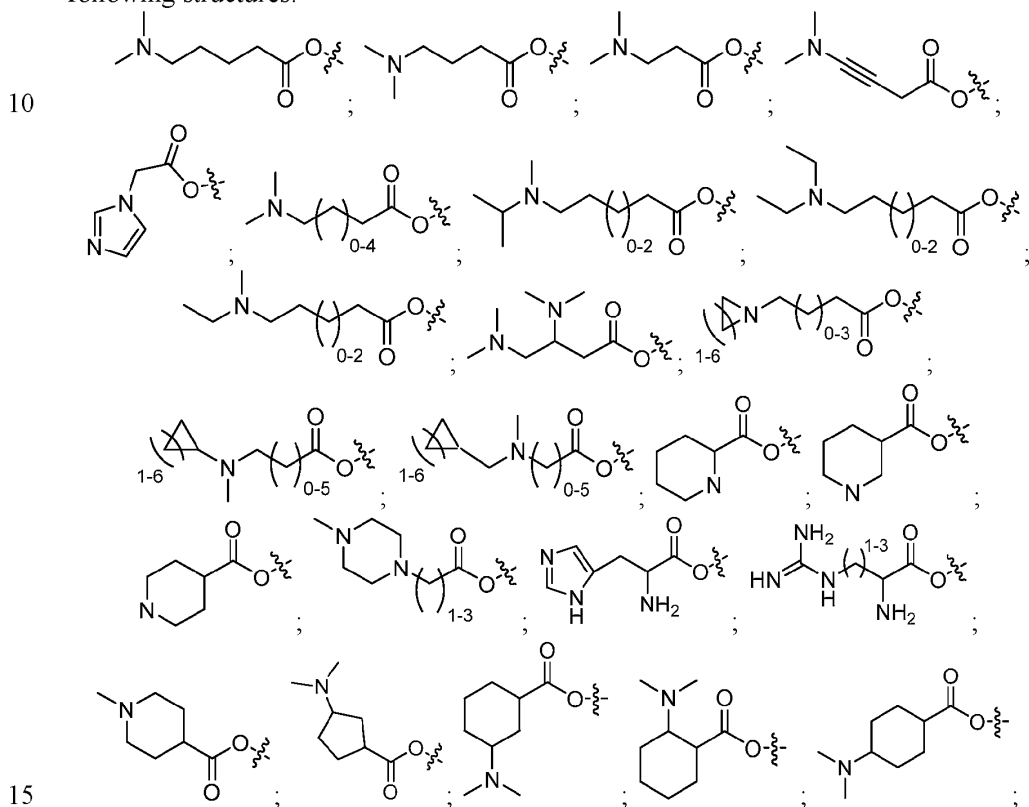
x is an integer from 0 to 6.

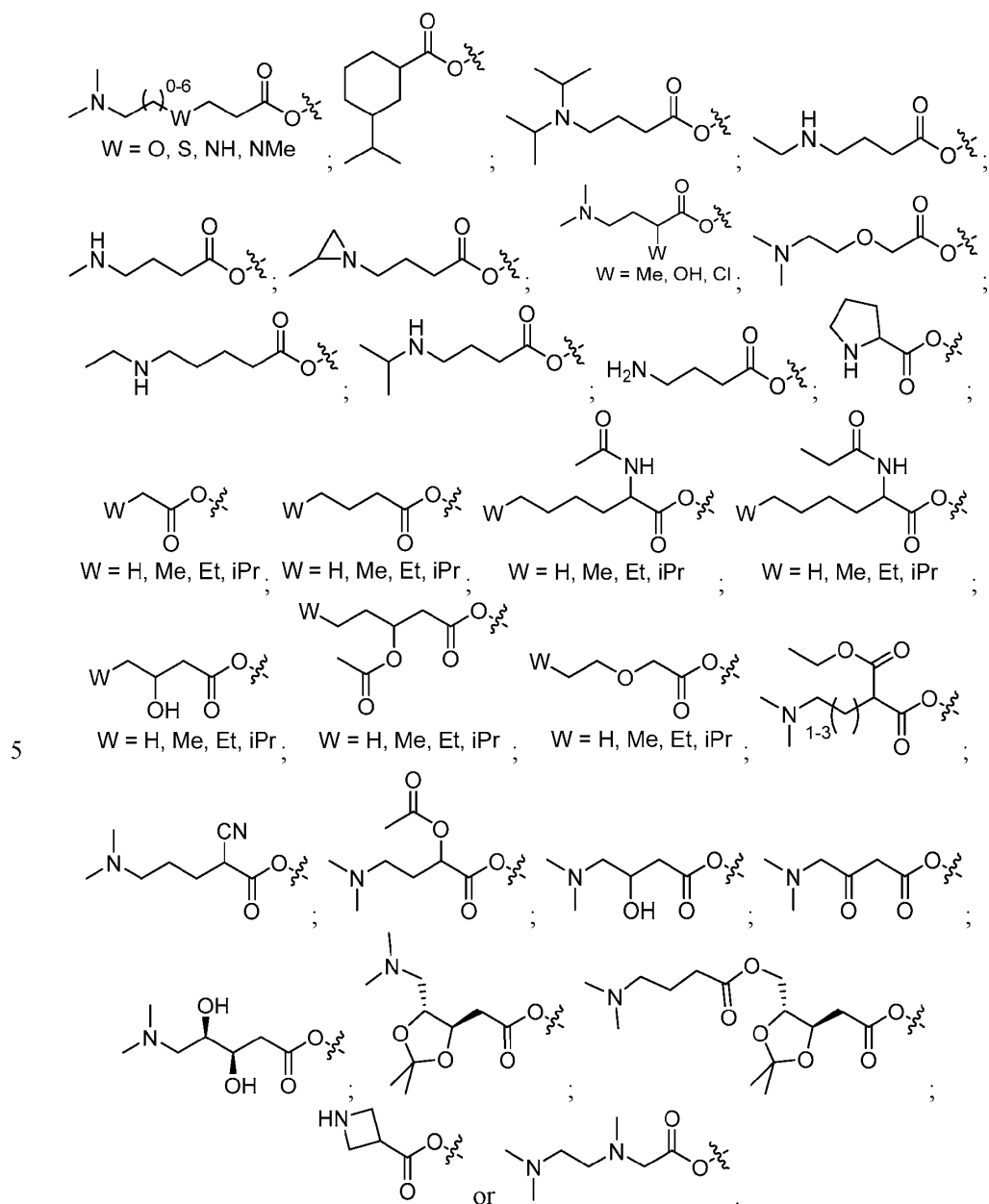
In some other embodiments of Formula (IV) or (V), Z is hydroxylalkyl, cyanoalkyl or an alkyl substituted with one or more ester or amide groups.

For example, in any of the foregoing embodiments of Formula (IV) or (V), Z has one of the following structures:

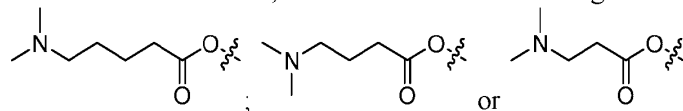


In other embodiments of Formula (IV) or (V), Z-L has one of the following structures:

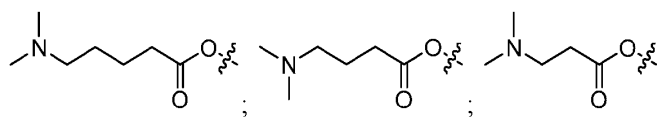




In other embodiments, Z-L has one of the following structures:



In still other embodiments, X is CH and Z-L has one of the following structures:

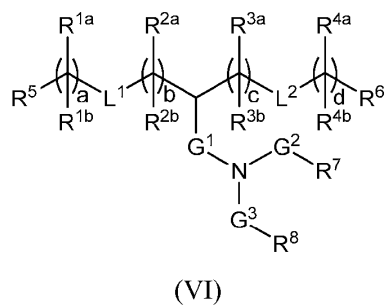


In various different embodiments, a cationic lipid of any one Embodiments 1, 2, 3, 4 or 5 has one of the structures set forth in Table 4 below.

Table 4: Representative Compounds of Formula (IV) or (V)

No.	Structure
IV-1	
IV-2	
IV-3	

5 In one embodiment, the cationic lipid is a compound having the following structure (VI):



or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof,
wherein:

- L^1 and L^2 are each independently $-\text{O}(\text{C}=\text{O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})-$, $-\text{O}-$,
 $-\text{S}(\text{O})_x-$, $-\text{S}-\text{S}-$, $-\text{C}(\text{=O})\text{S}-$, $-\text{SC}(\text{=O})-$, $-\text{NR}^a\text{C}(\text{=O})-$, $-\text{C}(\text{=O})\text{NR}^a-$, $-\text{NR}^a\text{C}(\text{=O})\text{NR}^a-$,
5 $-\text{OC}(\text{=O})\text{NR}^a-$, $-\text{NR}^a\text{C}(\text{=O})\text{O}-$ or a direct bond;
 G^1 is $\text{C}_1\text{-C}_2$ alkylene, $-(\text{C}=\text{O})-$, $-\text{O}(\text{C}=\text{O})-$, $-\text{SC}(\text{=O})-$, $-\text{NR}^a\text{C}(\text{=O})-$ or a
direct bond;
 G^2 is $-\text{C}(\text{=O})-$, $-(\text{C}=\text{O})\text{O}-$, $-\text{C}(\text{=O})\text{S}-$, $-\text{C}(\text{=O})\text{NR}^a-$ or a direct bond;
 G^3 is $\text{C}_1\text{-C}_6$ alkylene;
10 R^a is H or $\text{C}_1\text{-C}_{12}$ alkyl;
 R^{1a} and R^{1b} are, at each occurrence, independently either: (a) H or $\text{C}_1\text{-C}_{12}$
alkyl; or (b) R^{1a} is H or $\text{C}_1\text{-C}_{12}$ alkyl, and R^{1b} together with the carbon atom to which it
is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound
to form a carbon-carbon double bond;
15 R^{2a} and R^{2b} are, at each occurrence, independently either: (a) H or $\text{C}_1\text{-C}_{12}$
alkyl; or (b) R^{2a} is H or $\text{C}_1\text{-C}_{12}$ alkyl, and R^{2b} together with the carbon atom to which it
is bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound
to form a carbon-carbon double bond;
 R^{3a} and R^{3b} are, at each occurrence, independently either (a): H or $\text{C}_1\text{-C}_{12}$
20 alkyl; or (b) R^{3a} is H or $\text{C}_1\text{-C}_{12}$ alkyl, and R^{3b} together with the carbon atom to which it
is bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound
to form a carbon-carbon double bond;
 R^{4a} and R^{4b} are, at each occurrence, independently either: (a) H or $\text{C}_1\text{-C}_{12}$
alkyl; or (b) R^{4a} is H or $\text{C}_1\text{-C}_{12}$ alkyl, and R^{4b} together with the carbon atom to which it
25 is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound
to form a carbon-carbon double bond;
 R^5 and R^6 are each independently H or methyl;
 R^7 is H or $\text{C}_1\text{-C}_{20}$ alkyl;
 R^8 is OH, $-\text{N}(\text{R}^9)(\text{C}=\text{O})\text{R}^{10}$, $-(\text{C}=\text{O})\text{NR}^9\text{R}^{10}$, $-\text{NR}^9\text{R}^{10}$, $-(\text{C}=\text{O})\text{OR}^{11}$ or
30 $-\text{O}(\text{C}=\text{O})\text{R}^{11}$, provided that G^3 is $\text{C}_4\text{-C}_6$ alkylene when R^8 is $-\text{NR}^9\text{R}^{10}$,

R^9 and R^{10} are each independently H or C_1 - C_{12} alkyl;

R^{11} is aralkyl;

a, b, c and d are each independently an integer from 1 to 24; and

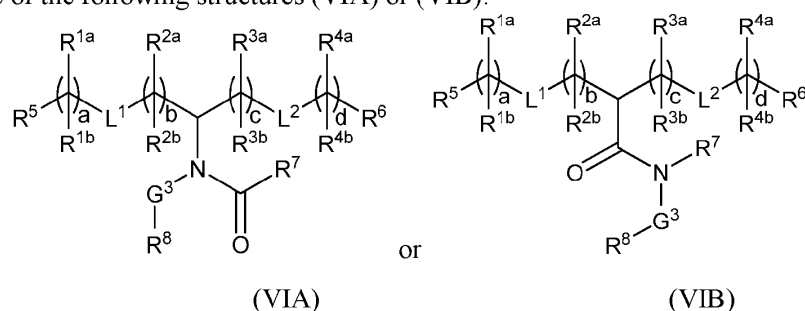
x is 0, 1 or 2,

5 wherein each alkyl, alkylene and aralkyl is optionally substituted.

In some embodiments of structure (VI), L^1 and L^2 are each independently $-O(C=O)-$, $-(C=O)O-$ or a direct bond. In other embodiments, G^1 and G^2 are each independently $-(C=O)-$ or a direct bond. In some different embodiments, L^1 and L^2 are each independently $-O(C=O)-$, $-(C=O)O-$ or a direct bond; and G^1 and G^2 are
10 each independently $-(C=O)-$ or a direct bond.

In some different embodiments of structure (VI), L^1 and L^2 are each independently $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $-SC(=O)-$, $-NR^a-$, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $-NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$, $-NR^aC(=O)O-$, $-NR^aS(O)_xNR^a-$, $-NR^aS(O)_x-$ or $-S(O)_xNR^a-$.

15 In other of the foregoing embodiments of structure (VI), the compound has one of the following structures (VIA) or (VIB):



20 In some embodiments, the compound has structure (VIA). In other embodiments, the compound has structure (VIB).

In any of the foregoing embodiments of structure (VI), one of L^1 or L^2 is $-O(C=O)-$. For example, in some embodiments each of L^1 and L^2 are $-O(C=O)-$.

In some different embodiments of any of the foregoing, one of L^1 or L^2 is $-(C=O)O-$. For example, in some embodiments each of L^1 and L^2 is $-(C=O)O-$.

In different embodiments of structure (VI), one of L^1 or L^2 is a direct bond. As used herein, a "direct bond" means the group (*e.g.*, L^1 or L^2) is absent. For example, in some embodiments each of L^1 and L^2 is a direct bond.

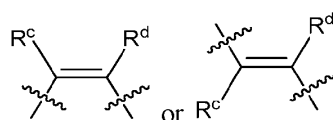
In other different embodiments of the foregoing, for at least one occurrence of R^{1a} and R^{1b} , R^{1a} is H or C_1 - C_{12} alkyl, and R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In still other different embodiments of structure (VI), for at least one occurrence of R^{4a} and R^{4b} , R^{4a} is H or C_1 - C_{12} alkyl, and R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

In more embodiments of structure (VI), for at least one occurrence of R^{2a} and R^{2b} , R^{2a} is H or C_1 - C_{12} alkyl, and R^{2b} together with the carbon atom to which it is bound is taken together with an adjacent R^{2b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

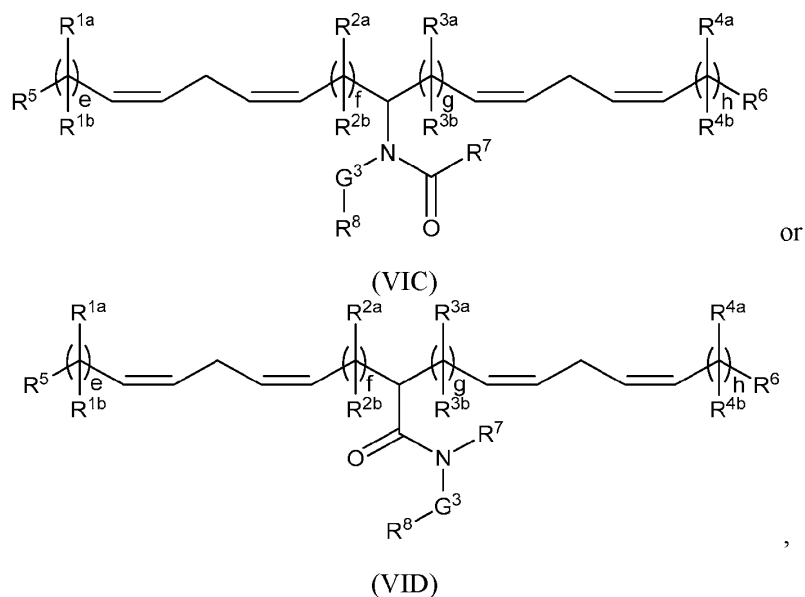
In other different embodiments of any of the foregoing, for at least one occurrence of R^{3a} and R^{3b} , R^{3a} is H or C_1 - C_{12} alkyl, and R^{3b} together with the carbon atom to which it is bound is taken together with an adjacent R^{3b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

It is understood that "carbon-carbon" double bond refers to one of the following structures:



wherein R^c and R^d are, at each occurrence, independently H or a substituent. For example, in some embodiments R^c and R^d are, at each occurrence, independently H, C_1 - C_{12} alkyl or cycloalkyl, for example H or C_1 - C_{12} alkyl.

In various other embodiments, the compound has one of the following structures (VIC) or (VID):

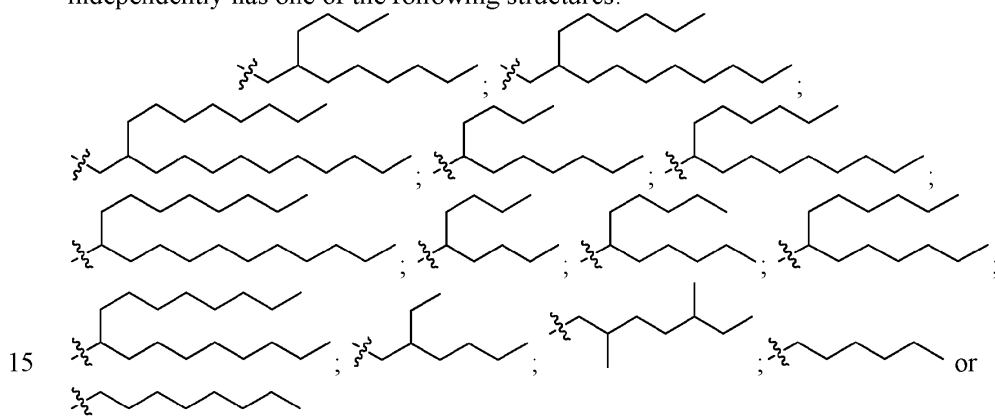


5 wherein e, f, g and h are each independently an integer from 1 to 12.

In some embodiments, the compound has structure (VIC). In other embodiments, the compound has structure (VID).

In various embodiments of the compounds of structures (VIC) or (VID), e, f, g and h are each independently an integer from 4 to 10.

10 In other different embodiments, or both, independently has one of the following structures:



In certain embodiments of the foregoing, a, b, c and d are each independently an integer from 2 to 12 or an integer from 4 to 12. In other embodiments, a, b, c and d are each independently an integer from 8 to 12 or 5 to 9. In some certain embodiments, a is 0. In some embodiments, a is 1. In other embodiments, a is 2. In more embodiments, a is 3. In yet other embodiments, a is 4. In some embodiments, a is 5. In other embodiments, a is 6. In more embodiments, a is 7. In yet other embodiments, a is 8. In some embodiments, a is 9. In other embodiments, a is 10. In more embodiments, a is 11. In yet other embodiments, a is 12. In some embodiments, a is 13. In other embodiments, a is 14. In more embodiments, a is 15. In yet other embodiments, a is 16.

In some embodiments of structure (VI), b is 1. In other embodiments, b is 2. In more embodiments, b is 3. In yet other embodiments, b is 4. In some embodiments, b is 5. In other embodiments, b is 6. In more embodiments, b is 7. In yet other embodiments, b is 8. In some embodiments, b is 9. In other embodiments, b is 10. In more embodiments, b is 11. In yet other embodiments, b is 12. In some embodiments, b is 13. In other embodiments, b is 14. In more embodiments, b is 15. In yet other embodiments, b is 16.

In some embodiments of structure (VI), c is 1. In other embodiments, c is 2. In more embodiments, c is 3. In yet other embodiments, c is 4. In some embodiments, c is 5. In other embodiments, c is 6. In more embodiments, c is 7. In yet other embodiments, c is 8. In some embodiments, c is 9. In other embodiments, c is 10. In more embodiments, c is 11. In yet other embodiments, c is 12. In some embodiments, c is 13. In other embodiments, c is 14. In more embodiments, c is 15. In yet other embodiments, c is 16.

In some certain embodiments of structure (VI), d is 0. In some embodiments, d is 1. In other embodiments, d is 2. In more embodiments, d is 3. In yet other embodiments, d is 4. In some embodiments, d is 5. In other embodiments, d is 6. In more embodiments, d is 7. In yet other embodiments, d is 8. In some embodiments, d is 9. In other embodiments, d is 10. In more embodiments, d is 11. In

yet other embodiments, d is 12. In some embodiments, d is 13. In other embodiments, d is 14. In more embodiments, d is 15. In yet other embodiments, d is 16.

In some embodiments of structure (VI), e is 1. In other embodiments, e is 2. In more embodiments, e is 3. In yet other embodiments, e is 4. In some
5 embodiments, e is 5. In other embodiments, e is 6. In more embodiments, e is 7. In yet other embodiments, e is 8. In some embodiments, e is 9. In other embodiments, e is 10. In more embodiments, e is 11. In yet other embodiments, e is 12.

In some embodiments of structure (VI), f is 1. In other embodiments, f is 2. In more embodiments, f is 3. In yet other embodiments, f is 4. In some
10 embodiments, f is 5. In other embodiments, f is 6. In more embodiments, f is 7. In yet other embodiments, f is 8. In some embodiments, f is 9. In other embodiments, f is 10. In more embodiments, f is 11. In yet other embodiments, f is 12.

In some embodiments of structure (VI), g is 1. In other embodiments, g is 2. In more embodiments, g is 3. In yet other embodiments, g is 4. In some
15 embodiments, g is 5. In other embodiments, g is 6. In more embodiments, g is 7. In yet other embodiments, g is 8. In some embodiments, g is 9. In other embodiments, g is 10. In more embodiments, g is 11. In yet other embodiments, g is 12.

In some embodiments of structure (VI), h is 1. In other embodiments, e is 2. In more embodiments, h is 3. In yet other embodiments, h is 4. In some
20 embodiments, e is 5. In other embodiments, h is 6. In more embodiments, h is 7. In yet other embodiments, h is 8. In some embodiments, h is 9. In other embodiments, h is 10. In more embodiments, h is 11. In yet other embodiments, h is 12.

In some other various embodiments of structure (VI), a and d are the same. In some other embodiments, b and c are the same. In some other specific
25 embodiments a and d are the same and b and c are the same.

The sum of a and b and the sum of c and d are factors which may be varied to obtain a lipid having the desired properties. In one embodiment, a and b are chosen such that their sum is an integer ranging from 14 to 24. In other embodiments, c and d are chosen such that their sum is an integer ranging from 14 to 24. In further
30 embodiment, the sum of a and b and the sum of c and d are the same. For example, in

some embodiments the sum of a and b and the sum of c and d are both the same integer which may range from 14 to 24. In still more embodiments, a, b, c and d are selected such that the sum of a and b and the sum of c and d is 12 or greater.

The substituents at R^{1a} , R^{2a} , R^{3a} and R^{4a} are not particularly limited. In some embodiments, at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is H. In certain embodiments R^{1a} , R^{2a} , R^{3a} and R^{4a} are H at each occurrence. In certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_{12} alkyl. In certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_8 alkyl. In certain other embodiments at least one of R^{1a} , R^{2a} , R^{3a} and R^{4a} is C_1 - C_6 alkyl. In some of the foregoing embodiments, the C_1 - C_8 alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In certain embodiments of the foregoing, R^{1a} , R^{1b} , R^{4a} and R^{4b} are C_1 - C_{12} alkyl at each occurrence.

In further embodiments of the foregoing, at least one of R^{1b} , R^{2b} , R^{3b} and R^{4b} is H or R^{1b} , R^{2b} , R^{3b} and R^{4b} are H at each occurrence.

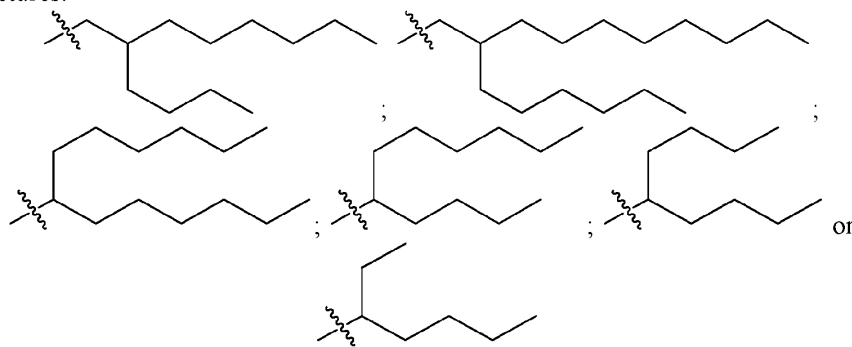
In certain embodiments of the foregoing, R^{1b} together with the carbon atom to which it is bound is taken together with an adjacent R^{1b} and the carbon atom to which it is bound to form a carbon-carbon double bond. In other embodiments of the foregoing R^{4b} together with the carbon atom to which it is bound is taken together with an adjacent R^{4b} and the carbon atom to which it is bound to form a carbon-carbon double bond.

The substituents at R^5 and R^6 are not particularly limited in the foregoing embodiments. In certain embodiments one of R^5 or R^6 is methyl. In other embodiments each of R^5 or R^6 is methyl.

The substituents at R^7 are not particularly limited in the foregoing embodiments. In certain embodiments R^7 is C_6 - C_{16} alkyl. In some other embodiments, R^7 is C_6 - C_9 alkyl. In some of these embodiments, R^7 is substituted with $-(C=O)OR^b$, $-O(C=O)R^b$, $-C(=O)R^b$, $-OR^b$, $-S(O)_xR^b$, $-S-SR^b$, $-C(=O)SR^b$, $-SC(=O)R^b$, $-NR^aR^b$, $-NR^aC(=O)R^b$, $-C(=O)NR^aR^b$, $-NR^aC(=O)NR^aR^b$, $-OC(=O)NR^aR^b$, $-NR^aC(=O)OR^b$,

-NR^aS(O)_xNR^aR^b, -NR^aS(O)_xR^b or -S(O)_xNR^aR^b, wherein: R^a is H or C₁-C₁₂ alkyl; R^b is C₁-C₁₅ alkyl; and x is 0, 1 or 2. For example, in some embodiments R⁷ is substituted with -(C=O)OR^b or -O(C=O)R^b.

In various of the foregoing embodiments of structure (VI), R^b is
5 branched C₃-C₁₅ alkyl. For example, in some embodiments R^b has one of the following
structures:

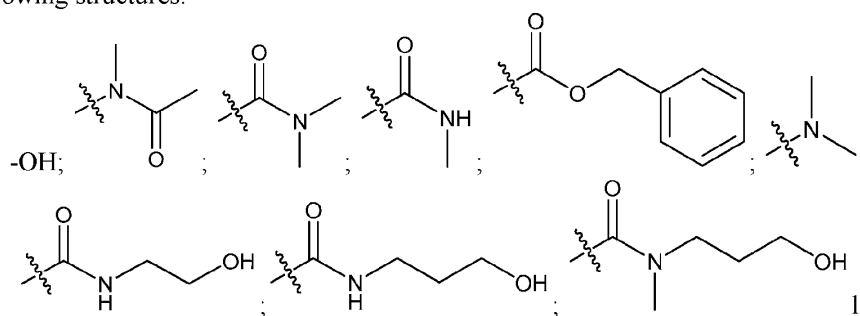


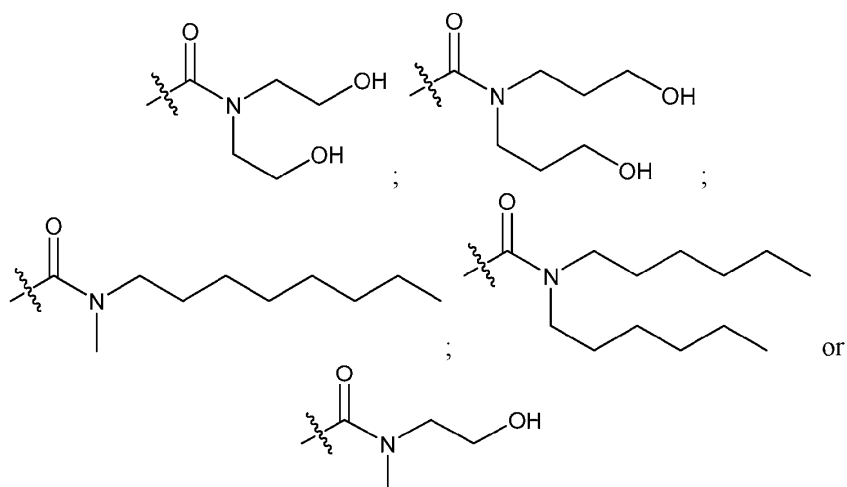
10 In certain embodiments, R⁸ is OH.

In other embodiments of structure (VI), R^8 is $-N(R^9)(C=O)R^{10}$. In some other embodiments, R^8 is $-(C=O)NR^9R^{10}$. In still more embodiments, R^8 is $-NR^9R^{10}$. In some of the foregoing embodiments, R^9 and R^{10} are each independently H or C_1 - C_8 alkyl, for example H or C_1 - C_3 alkyl. In more specific of these embodiments, the C_1 - C_8 alkyl or C_1 - C_3 alkyl is unsubstituted or substituted with hydroxyl. In other of these
15
embodiments, R^9 and R^{10} are each methyl.

In yet more embodiments of structure (VI), R⁸ is -(C=O)OR¹¹. In some of these embodiments R¹¹ is benzyl.

In yet more specific embodiments of structure (VI), R^8 has one of the
 20 following structures:





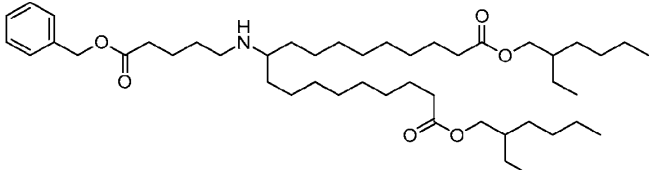
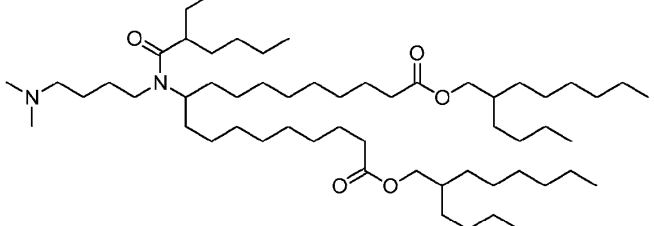
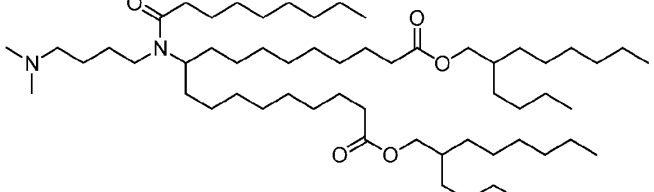
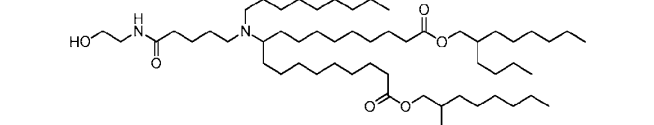
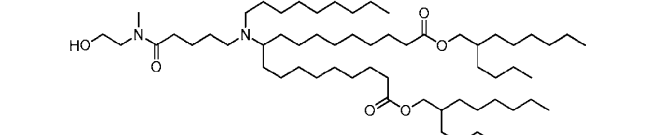
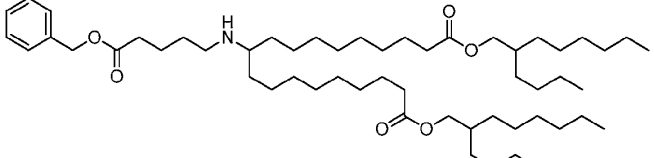
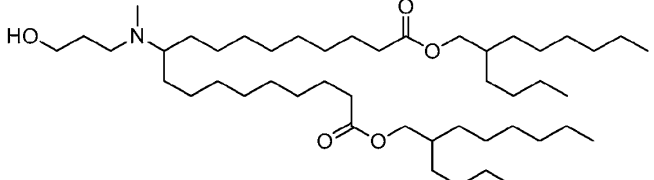
In still other embodiments of the foregoing compounds, G^3 is C_2 - C_5 alkylene, for example C_2 - C_4 alkylene, C_3 alkylene or C_4 alkylene. In some of these
 5 embodiments, R^8 is OH. In other embodiments, G^2 is absent and R^7 is C_1 - C_2 alkylene, such as methyl.

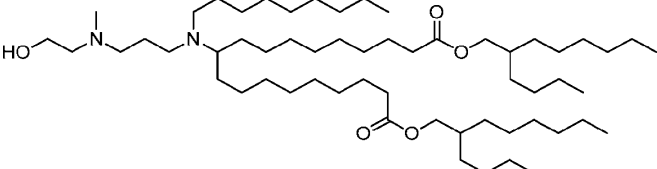
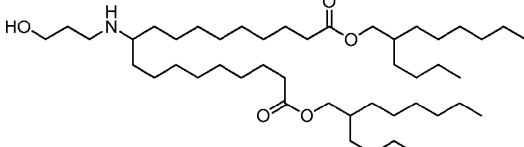
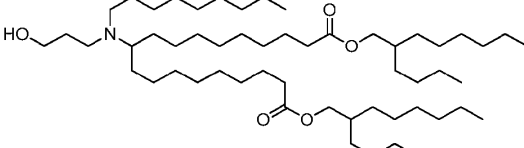
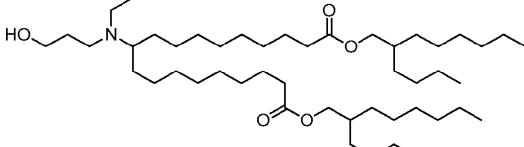
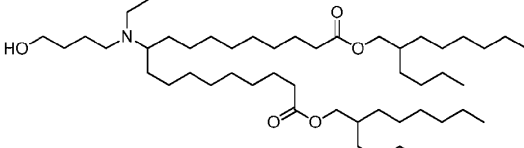
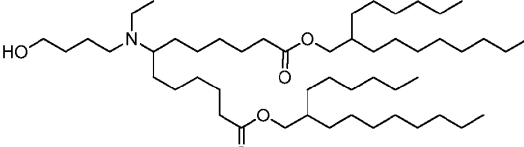
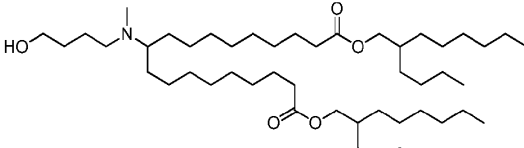
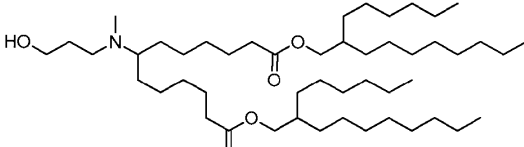
In various different embodiments, the compound has one of the structures set forth in Table 5 below.

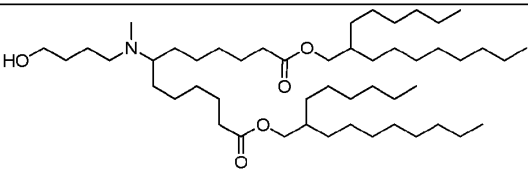
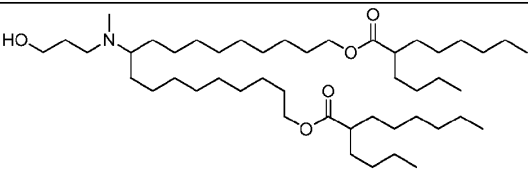
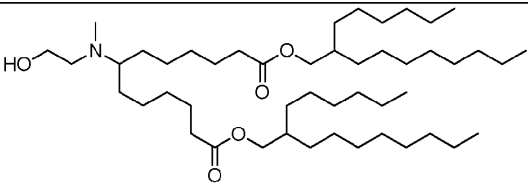
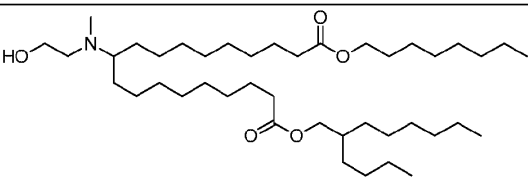
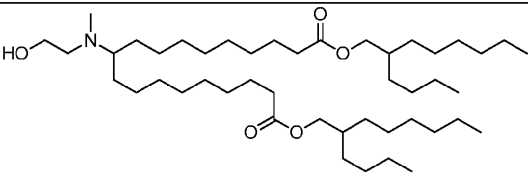
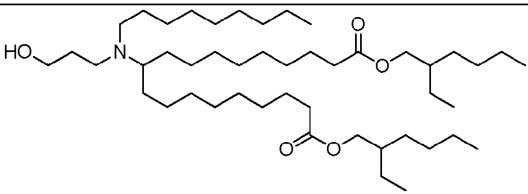
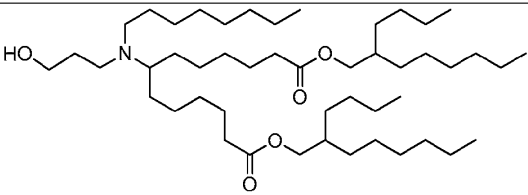
10

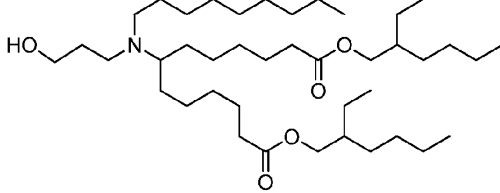
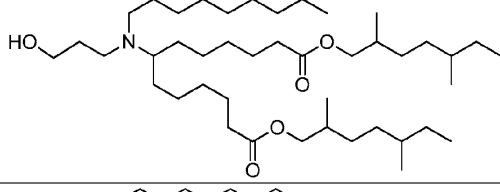
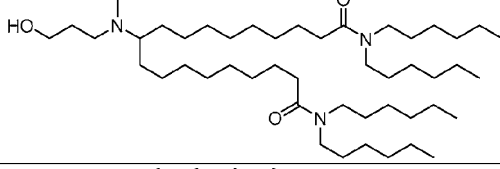
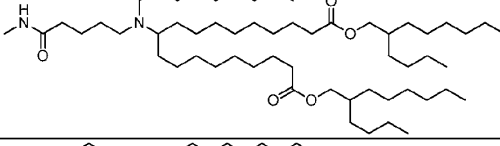
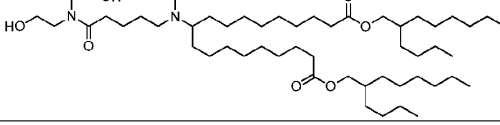
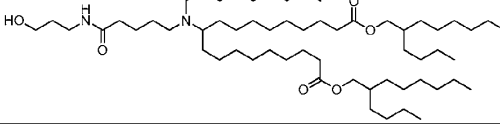
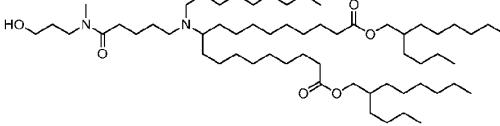
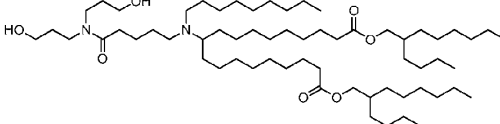
Table 5. Representative cationic lipids of structure (VI)

No.	Structure
VI-1	
VI-2	

No.	Structure
VI-3	
VI-4	
VI-5	
VI-6	
VI-7	
VI-8	
VI-9	

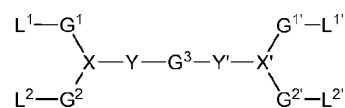
No.	Structure
VI-10	
VI-11	
VI-12	
VI-13	
VI-14	
VI-15	
VI-16	
VI-17	

No.	Structure
VI-18	
VI-19	
VI-20	
VI-21	
VI-22	
VI-23	
VI-24	

No.	Structure
VI-25	
VI-26	
VI-27	
VI-28	
VI-29	
VI-30	
VI-31	
VI-32	

No.	Structure
VI-33	
VI-34	
VI-35	
VI-36	
VI-37	

In one embodiment, the cationic lipid is a compound having the following structure (VII):



(VII)

5 or a pharmaceutically acceptable salt, prodrug or stereoisomer thereof, wherein:

X and X' are each independently N or CR;

Y and Y' are each independently absent, -O(C=O)-, -(C=O)O- or NR,

provided that:

a) Y is absent when X is N;

b) Y' is absent when X' is N;

c) Y is -O(C=O)-, -(C=O)O- or NR when X is CR; and

d) Y' is -O(C=O)-, -(C=O)O- or NR when X' is CR,

10

L^1 and $L^{1'}$ are each independently $-O(C=O)R^1$, $-(C=O)OR^1$, $-C(=O)R^1$, $-OR^1$, $-S(O)_zR^1$, $-S-SR^1$, $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ or $-NR^aC(=O)OR^1$;

L^2 and $L^{2'}$ are each independently $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_zR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(=O)NR^eR^f$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$, $-NR^dC(=O)OR^2$ or a direct bond to R^2 ;

G^1 , $G^{1'}$, G^2 and $G^{2'}$ are each independently C_2 - C_{12} alkenylene or C_2 - C_{12} alkenylene;

G^3 is C_2 - C_{24} heteroalkylene or C_2 - C_{24} heteroalkenylene;

R^a , R^b , R^d and R^e are, at each occurrence, independently H, C_1 - C_{12} alkyl or C_2 - C_{12} alkenyl;

R^c and R^f are, at each occurrence, independently C_1 - C_{12} alkyl or C_2 - C_{12} alkenyl;

R is, at each occurrence, independently H or C_1 - C_{12} alkyl;

R^1 and R^2 are, at each occurrence, independently branched C_6 - C_{24} alkyl or branched C_6 - C_{24} alkenyl;

z is 0, 1 or 2, and

wherein each alkyl, alkenyl, alkylene, alkenylene, heteroalkylene and heteroalkenylene is independently substituted or unsubstituted unless otherwise specified.

In other different embodiments of structure (VII):

X and X' are each independently N or CR;

Y and Y' are each independently absent or NR, provided that:

a) Y is absent when X is N;

b) Y' is absent when X' is N;

c) Y is NR when X is CR; and

d) Y' is NR when X' is CR,

L^1 and $L^{1'}$ are each independently $-O(C=O)R^1$, $-(C=O)OR^1$, $-C(=O)R^1$, $-OR^1$, $-S(O)_zR^1$, $-S-SR^1$, $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ or $-NR^aC(=O)OR^1$;

L^2 and $L^{2'}$ are each independently $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$,

-OR², -S(O)_zR², -S-SR², -C(=O)SR², -SC(=O)R², -NR^dC(=O)R², -C(=O)NR^eR^f,
-NR^dC(=O)NR^eR^f, -OC(=O)NR^eR^f, -NR^dC(=O)OR² or a direct bond to R²;

G¹, G^{1'}, G² and G^{2'} are each independently C₂-C₁₂ alkylene or C₂-C₁₂ alkenylene;

5 G³ is C₂-C₂₄ alkyleneoxide or C₂-C₂₄ alkenyleneoxide;

R^a, R^b, R^d and R^e are, at each occurrence, independently H, C₁-C₁₂ alkyl or C₂-C₁₂ alkenyl;

R^c and R^f are, at each occurrence, independently C₁-C₁₂ alkyl or C₂-C₁₂ alkenyl;

10 R is, at each occurrence, independently H or C₁-C₁₂ alkyl;

R¹ and R² are, at each occurrence, independently branched C₆-C₂₄ alkyl or branched C₆-C₂₄ alkenyl;

z is 0, 1 or 2, and

wherein each alkyl, alkenyl, alkylene, alkenylene, alkyleneoxide and alkenyleneoxide is

15 independently substituted or unsubstituted unless otherwise specified.

In some embodiments of structure (VII), G³ is C₂-C₂₄ alkyleneoxide or C₂-C₂₄ alkenyleneoxide. In certain embodiments, G³ is unsubstituted. In other embodiments, G³ is substituted, for example substituted with hydroxyl. In more specific embodiments G³ is C₂-C₁₂ alkyleneoxide, for example, in some embodiments

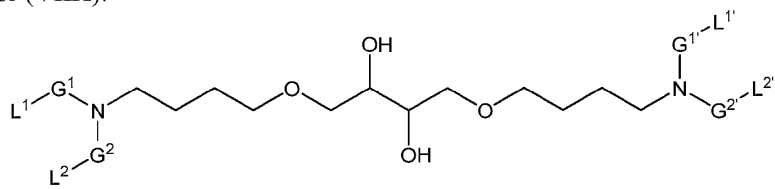
20 G³ is C₃-C₇ alkyleneoxide or in other embodiments G³ is C₃-C₁₂ alkyleneoxide.

In other embodiments of structure (VII), G³ is C₂-C₂₄ alkyleneaminyll or C₂-C₂₄ alkenyleneaminyll, for example C₆-C₁₂ alkyleneaminyll. In some of these embodiments, G³ is unsubstituted. In other of these embodiments, G³ is substituted with C₁-C₆ alkyl.

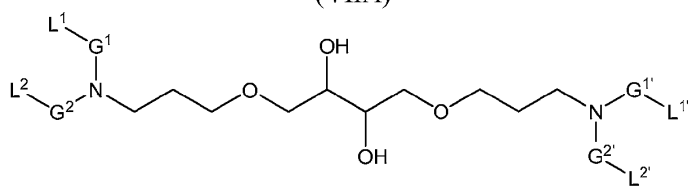
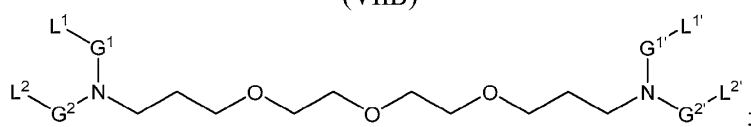
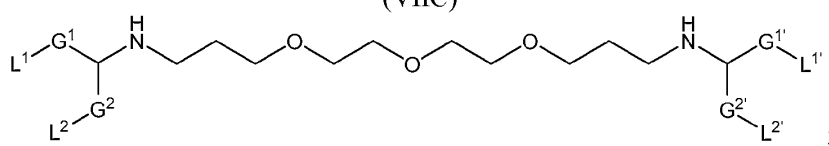
25 In some embodiments of structure (VII), X and X' are each N, and Y and Y' are each absent. In other embodiments, X and X' are each CR, and Y and Y' are each NR. In some of these embodiments, R is H.

In certain embodiments of structure (VII), X and X' are each CR, and Y and Y' are each independently -O(C=O)- or -(C=O)O-.

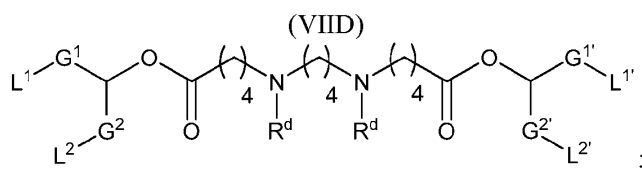
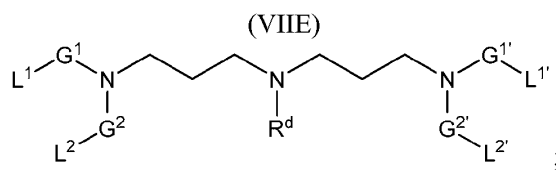
In some of the foregoing embodiments of structure (VII), the compound has one of the following structures (VIIA), (VIIB), (VIIC), (VIID), (VIIE), (VIIF), (VIIG) or (VIIH):



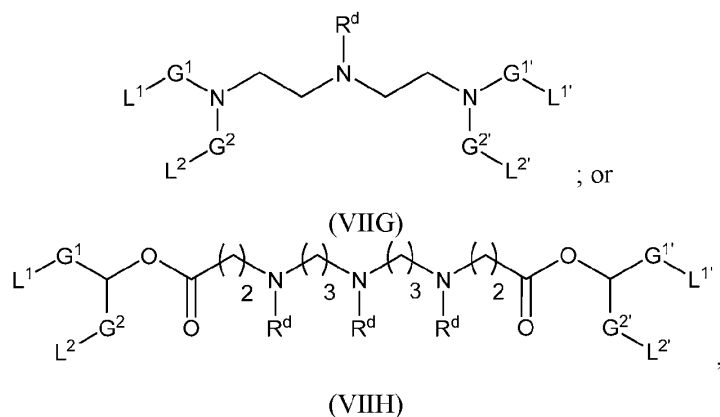
5



10



15

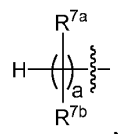


5 wherein R^d is, at each occurrence, independently H or optionally substituted C_1 - C_6 alkyl. For example, in some embodiments R^d is H. In other embodiments, R^d is C_1 - C_6 alkyl, such as methyl. In other embodiments, R^d is substituted C_1 - C_6 alkyl, such as C_1 - C_6 alkyl substituted with $-\text{O}(\text{C}=\text{O})\text{R}$, $-(\text{C}=\text{O})\text{OR}$, $-\text{NRC}(\text{O})\text{R}$ or $-\text{C}(\text{O})\text{N}(\text{R})_2$, wherein R is, at each occurrence, independently H or C_1 - C_{12} alkyl.

10 In some of the foregoing embodiments of structure (VII), L^1 and $L^{1'}$ are each independently $-\text{O}(\text{C}=\text{O})\text{R}^1$, $-(\text{C}=\text{O})\text{OR}^1$ or $-\text{C}(\text{O})\text{NR}^b\text{R}^c$, and L^2 and $L^{2'}$ are each independently $-\text{O}(\text{C}=\text{O})\text{R}^2$, $-(\text{C}=\text{O})\text{OR}^2$ or $-\text{C}(\text{O})\text{NR}^e\text{R}^f$. For example, in some embodiments L^1 and $L^{1'}$ are each $-(\text{C}=\text{O})\text{OR}^1$, and L^2 and $L^{2'}$ are each $-(\text{C}=\text{O})\text{OR}^2$. In other embodiments L^1 and $L^{1'}$ are each $-(\text{C}=\text{O})\text{OR}^1$, and L^2 and $L^{2'}$ are each $-\text{C}(\text{O})\text{NR}^e\text{R}^f$. In other embodiments L^1 and $L^{1'}$ are each $-\text{C}(\text{O})\text{NR}^b\text{R}^c$, and L^2 and $L^{2'}$ are each $-\text{C}(\text{O})\text{NR}^e\text{R}^f$.

In some embodiments of the foregoing, G^1 , $G^{1'}$, G^2 and $G^{2'}$ are each independently C_2 - C_8 alkylene, for example C_4 - C_8 alkylene.

20 In some of the foregoing embodiments of structure (VII), R^1 or R^2 , are each, at each occurrence, independently branched C_6 - C_{24} alkyl. For example, in some embodiments, R^1 and R^2 at each occurrence, independently have the following structure:



wherein:

R^{7a} and R^{7b} are, at each occurrence, independently H or C_1 - C_{12} alkyl;

and

a is an integer from 2 to 12,

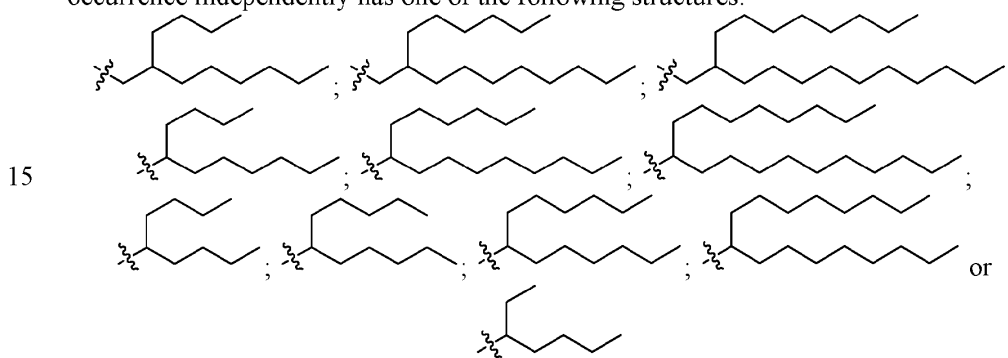
wherein R^{7a} , R^{7b} and a are each selected such that R^1 and R^2 each independently

5 comprise from 6 to 20 carbon atoms. For example, in some embodiments a is an integer ranging from 5 to 9 or from 8 to 12.

In some of the foregoing embodiments of structure (VII), at least one occurrence of R^{7a} is H. For example, in some embodiments, R^{7a} is H at each occurrence.

In other different embodiments of the foregoing, at least one occurrence of R^{7b} is C_1 - C_8
10 alkyl. For example, in some embodiments, C_1 - C_8 alkyl is methyl, ethyl, n-propyl, isopropyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

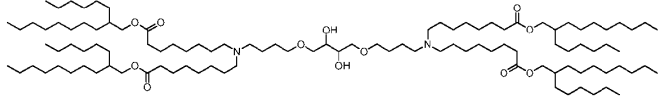
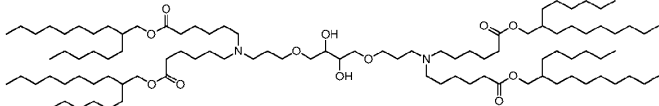
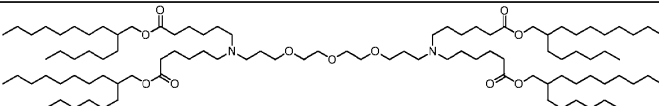
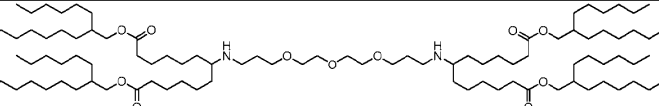
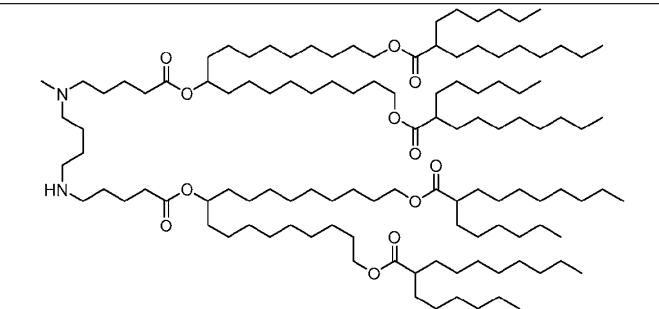
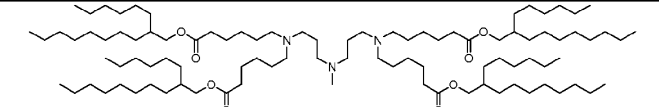
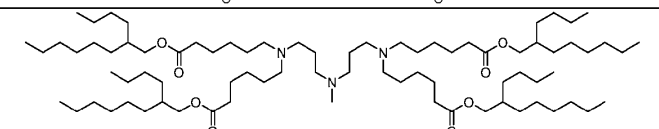
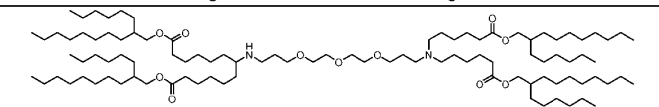
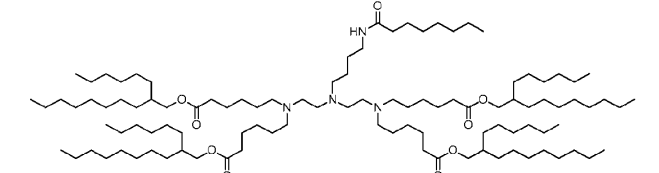
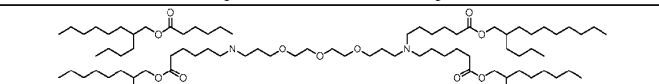
In different embodiments of structure (VII), R^1 or R^2 , or both, at each occurrence independently has one of the following structures:



In some of the foregoing embodiments of structure (VII), R^b , R^c , R^e and R^f , when present, are each independently C_3 - C_{12} alkyl. For example, in some
20 embodiments R^b , R^c , R^e and R^f , when present, are n-hexyl and in other embodiments R^b , R^c , R^e and R^f , when present, are n-octyl.

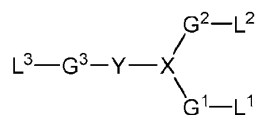
In various different embodiments of structure (VII), the cationic lipid has one of the structures set forth in Table 6 below.

Table 6. Representative cationic lipids of structure (VII)

No.	Structure
VII-1	
VII-2	
VII-3	
VII-4	
VII-5	
VII-6	
VII-7	
VII-8	
VII-9	
VII-10	

No.	Structure
VII-11	

In one embodiment, the cationic lipid is a compound having the following structure (VIII):



(VIII)

- 5 or a pharmaceutically acceptable salt, prodrug or stereoisomer thereof, wherein:
- X is N, and Y is absent; or X is CR, and Y is NR;
- L^1 is $-\text{O}(\text{C}=\text{O})\text{R}^1$, $-(\text{C}=\text{O})\text{OR}^1$, $-\text{C}(\text{O})\text{R}^1$, $-\text{OR}^1$, $-\text{S}(\text{O})_x\text{R}^1$, $-\text{S}-\text{SR}^1$, $-\text{C}(\text{O})\text{SR}^1$, $-\text{SC}(\text{O})\text{R}^1$, $-\text{NR}^a\text{C}(\text{O})\text{R}^1$, $-\text{C}(\text{O})\text{NR}^b\text{R}^c$, $-\text{NR}^a\text{C}(\text{O})\text{NR}^b\text{R}^c$, $-\text{OC}(\text{O})\text{NR}^b\text{R}^c$ or $-\text{NR}^a\text{C}(\text{O})\text{OR}^1$;
- 10 L^2 is $-\text{O}(\text{C}=\text{O})\text{R}^2$, $-(\text{C}=\text{O})\text{OR}^2$, $-\text{C}(\text{O})\text{R}^2$, $-\text{OR}^2$, $-\text{S}(\text{O})_x\text{R}^2$, $-\text{S}-\text{SR}^2$, $-\text{C}(\text{O})\text{SR}^2$, $-\text{SC}(\text{O})\text{R}^2$, $-\text{NR}^d\text{C}(\text{O})\text{R}^2$, $-\text{C}(\text{O})\text{NR}^e\text{R}^f$, $-\text{NR}^d\text{C}(\text{O})\text{NR}^e\text{R}^f$, $-\text{OC}(\text{O})\text{NR}^e\text{R}^f$, $-\text{NR}^d\text{C}(\text{O})\text{OR}^2$ or a direct bond to R^2 ;
- L^3 is $-\text{O}(\text{C}=\text{O})\text{R}^3$ or $-(\text{C}=\text{O})\text{OR}^3$;
- G^1 and G^2 are each independently C_2 - C_{12} alkylene or C_2 - C_{12} alkenylene;
- 15 G^3 is C_1 - C_{24} alkylene, C_2 - C_{24} alkenylene, C_1 - C_{24} heteroalkylene or C_2 - C_{24} heteroalkenylene;
- R^a , R^b , R^d and R^e are each independently H or C_1 - C_{12} alkyl or C_1 - C_{12} alkenyl;
- R^c and R^f are each independently C_1 - C_{12} alkyl or C_2 - C_{12} alkenyl;
- 20 each R is independently H or C_1 - C_{12} alkyl;

R^1 , R^2 and R^3 are each independently C_1 - C_{24} alkyl or C_2 - C_{24} alkenyl; and
 x is 0, 1 or 2, and

wherein each alkyl, alkenyl, alkylene, alkenylene, heteroalkylene and
heteroalkenylene is independently substituted or unsubstituted unless otherwise

5 specified.

In more embodiments of structure (I):

X is N, and Y is absent; or X is CR, and Y is NR;

L^1 is $-O(C=O)R^1$, $-(C=O)OR^1$, $-C(=O)R^1$, $-OR^1$, $-S(O)_xR^1$, $-S-SR^1$,
 $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$,
10 $-OC(=O)NR^bR^c$ or $-NR^aC(=O)OR^1$;

L^2 is $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_xR^2$, $-S-SR^2$,
 $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(=O)NR^eR^f$, $-NR^dC(=O)NR^eR^f$,
 $-OC(=O)NR^eR^f$; $-NR^dC(=O)OR^2$ or a direct bond to R^2 ;

L^3 is $-O(C=O)R^3$ or $-(C=O)OR^3$;

15 G^1 and G^2 are each independently C_2 - C_{12} alkylene or C_2 - C_{12} alkenylene;

G^3 is C_1 - C_{24} alkylene, C_2 - C_{24} alkenylene, C_1 - C_{24} heteroalkylene or C_2 -
 C_{24} heteroalkenylene when X is CR, and Y is NR; and G^3 is C_1 - C_{24} heteroalkylene or
 C_2 - C_{24} heteroalkenylene when X is N, and Y is absent;

R^a , R^b , R^d and R^e are each independently H or C_1 - C_{12} alkyl or C_1 - C_{12}
20 alkenyl;

R^c and R^f are each independently C_1 - C_{12} alkyl or C_2 - C_{12} alkenyl;

each R is independently H or C_1 - C_{12} alkyl;

R^1 , R^2 and R^3 are each independently C_1 - C_{24} alkyl or C_2 - C_{24} alkenyl; and
 x is 0, 1 or 2, and

25 wherein each alkyl, alkenyl, alkylene, alkenylene, heteroalkylene and
heteroalkenylene is independently substituted or unsubstituted unless otherwise
specified.

In other embodiments of structure (I):

X is N and Y is absent, or X is CR and Y is NR;

30 L^1 is $-O(C=O)R^1$, $-(C=O)OR^1$, $-C(=O)R^1$, $-OR^1$, $-S(O)_xR^1$, $-S-SR^1$,

-C(=O)SR¹, -SC(=O)R¹, -NR^aC(=O)R¹, -C(=O)NR^bR^c, -NR^aC(=O)NR^bR^c,
-OC(=O)NR^bR^c or -NR^aC(=O)OR¹;

L² is -O(C=O)R², -(C=O)OR², -C(=O)R², -OR², -S(O)_xR², -S-SR²,

-C(=O)SR², -SC(=O)R², -NR^dC(=O)R², -C(=O)NR^eR^f, -NR^dC(=O)NR^eR^f,

5 -OC(=O)NR^eR^f, -NR^dC(=O)OR² or a direct bond to R²;

L³ is -O(C=O)R³ or -(C=O)OR³;

G¹ and G² are each independently C₂-C₁₂ alkylene or C₂-C₁₂ alkenylene;

G³ is C₁-C₂₄ alkylene, C₂-C₂₄ alkenylene, C₁-C₂₄ heteroalkylene or C₂-

C₂₄ heteroalkenylene;

10 R^a, R^b, R^d and R^e are each independently H or C₁-C₁₂ alkyl or C₁-C₁₂
alkenyl;

R^c and R^f are each independently C₁-C₁₂ alkyl or C₂-C₁₂ alkenyl;

each R is independently H or C₁-C₁₂ alkyl;

R¹, R² and R³ are each independently branched C₆-C₂₄ alkyl or branched

15 C₆-C₂₄ alkenyl; and

x is 0, 1 or 2, and

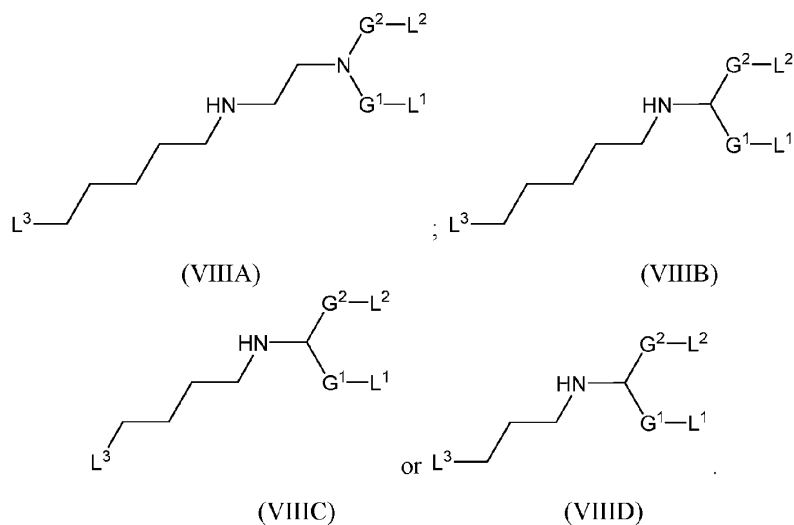
wherein each alkyl, alkenyl, alkylene, alkenylene, heteroalkylene and heteroalkenylene
is independently substituted or unsubstituted unless otherwise specified.

In certain embodiments of structure (VIII), G³ is unsubstituted. In more
20 specific embodiments G³ is C₂-C₁₂ alkylene, for example, in some embodiments G³ is
C₃-C₇ alkylene or in other embodiments G³ is C₃-C₁₂ alkylene. In some embodiments,
G³ is C₂ or C₃ alkylene.

In other embodiments of structure (VIII), G³ is C₁-C₁₂ heteroalkylene,
for example C₁-C₁₂ aminylalkylene.

25 In certain embodiments of structure (VIII), X is N and Y is absent. In
other embodiments, X is CR and Y is NR, for example in some of these embodiments R
is H.

In some of the foregoing embodiments of structure (VIII), the compound
has one of the following structures (VIII A), (VIII B), (VIII C) or (VIII D):



5 In some of the foregoing embodiments of structure (VIII), L¹ is -

O(C=O)R¹, -(C=O)OR¹ or

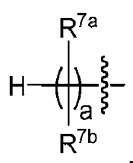
-C(=O)NR^bR^c, and L² is -O(C=O)R², -(C=O)OR² or -C(=O)NR^eR^f. In other specific

embodiments, L¹ is -(C=O)OR¹ and L² is -(C=O)OR². In any of the foregoing

embodiments, L³ is -(C=O)OR³.

10 In some of the foregoing embodiments of structure (VIII), G¹ and G² are each independently C₂-C₁₂ alkylene, for example C₄-C₁₀ alkylene.

In some of the foregoing embodiments of structure (VIII), R¹, R² and R³ are each, independently branched C₆-C₂₄ alkyl. For example, in some embodiments, R¹, R² and R³ each, independently have the following structure:



wherein:

R^{7a} and R^{7b} are, at each occurrence, independently H or C₁-C₁₂ alkyl;

and

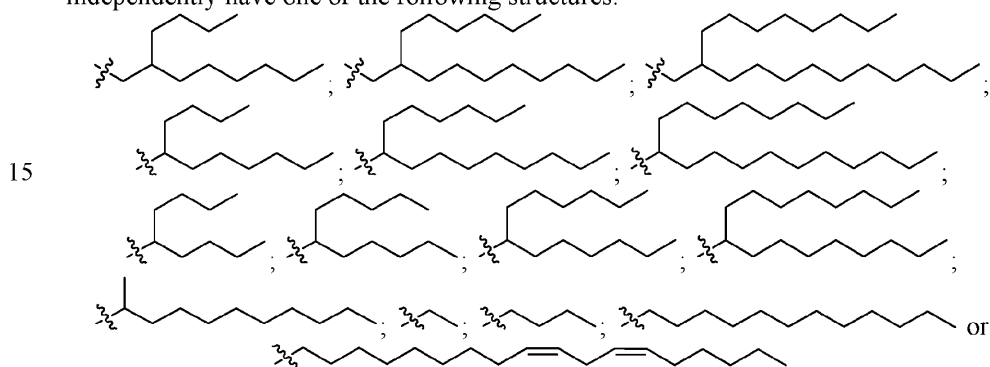
a is an integer from 2 to 12,

wherein R^{7a} , R^{7b} and a are each selected such that R^1 and R^2 each independently comprise from 6 to 20 carbon atoms. For example, in some embodiments a is an integer ranging from 5 to 9 or from 8 to 12.

In some of the foregoing embodiments of structure (VIII), at least one occurrence of R^{7a} is H. For example, in some embodiments, R^{7a} is H at each occurrence. In other different embodiments of the foregoing, at least one occurrence of R^{7b} is C_1 - C_8 alkyl. For example, in some embodiments, C_1 - C_8 alkyl is methyl, ethyl, n-propyl, isopropyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In some of the foregoing embodiments of structure (VIII), X is CR, Y is NR and R^3 is C_1 - C_{12} alkyl, such as ethyl, propyl or butyl. In some of these embodiments, R^1 and R^2 are each independently branched C_6 - C_{24} alkyl.

In different embodiments of structure (VIII), R^1 , R^2 and R^3 each, independently have one of the following structures:



In certain embodiments of structure (VIII), R^1 and R^2 and R^3 are each, independently, branched C_6 - C_{24} alkyl and R^3 is C_1 - C_{24} alkyl or C_2 - C_{24} alkenyl.

In some of the foregoing embodiments of structure (VIII), R^b , R^c , R^e and R^f are each independently C_3 - C_{12} alkyl. For example, in some embodiments R^b , R^c , R^e and R^f are n-hexyl and in other embodiments R^b , R^c , R^e and R^f are n-octyl.

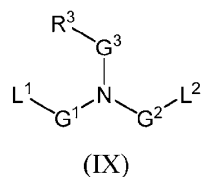
In various different embodiments of structure (VIII), the compound has one of the structures set forth in Table 7 below.

Table 7. Representative cationic lipids of structure (VIII)

No.	Structure
VIII-1	
VIII-2	
VIII-3	
VIII-4	
VIII-5	
VIII-6	
VIII-7	
VIII-8	

No.	Structure
VIII-9	
VIII-10	
VIII-11	
VIII-12	

In one embodiment, the cationic lipid is a compound having the following structure (IX):



- 5 or a pharmaceutically acceptable salt, prodrug or stereoisomer thereof, wherein:
- L^1 is $-O(C=O)R^1$, $-(C=O)OR^1$, $-C(=O)R^1$, $-OR^1$, $-S(O)_xR^1$, $-S-SR^1$, $-C(=O)SR^1$, $-SC(=O)R^1$, $-NR^aC(=O)R^1$, $-C(=O)NR^bR^c$, $-NR^aC(=O)NR^bR^c$, $-OC(=O)NR^bR^c$ or $-NR^aC(=O)OR^1$;
- L^2 is $-O(C=O)R^2$, $-(C=O)OR^2$, $-C(=O)R^2$, $-OR^2$, $-S(O)_xR^2$, $-S-SR^2$, $-C(=O)SR^2$, $-SC(=O)R^2$, $-NR^dC(=O)R^2$, $-C(=O)NR^eR^f$, $-NR^dC(=O)NR^eR^f$, $-OC(=O)NR^eR^f$, $-NR^dC(=O)OR^2$ or a direct bond to R^2 ;
- 10 G^1 and G^2 are each independently C_2 - C_{12} alkylene or C_2 - C_{12} alkenylene;
- G^3 is C_1 - C_{24} alkylene, C_2 - C_{24} alkenylene, C_3 - C_8 cycloalkylene or C_3 - C_8 cycloalkenylene;

R^a , R^b , R^d and R^e are each independently H or C_1 - C_{12} alkyl or C_1 - C_{12} alkenyl;

R^c and R^f are each independently C_1 - C_{12} alkyl or C_2 - C_{12} alkenyl;

R^1 and R^2 are each independently branched C_6 - C_{24} alkyl or branched C_6 -
5 C_{24} alkenyl;

R^3 is $-N(R^4)R^5$;

R^4 is C_1 - C_{12} alkyl;

R^5 is substituted C_1 - C_{12} alkyl; and

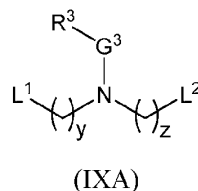
x is 0, 1 or 2, and

10 wherein each alkyl, alkenyl, alkylene, alkenylene, cycloalkylene, cycloalkenylene, aryl and aralkyl is independently substituted or unsubstituted unless otherwise specified.

In certain embodiments of structure (XI), G^3 is unsubstituted. In more specific embodiments G^3 is C_2 - C_{12} alkylene, for example, in some embodiments G^3 is C_3 - C_7 alkylene or in other embodiments G^3 is C_3 - C_{12} alkylene. In some embodiments,

15 G^3 is C_2 or C_3 alkylene.

In some of the foregoing embodiments of structure (IX), the compound has the following structure (IX A):



20 wherein y and z are each independently integers ranging from 2 to 12, for example an integer from 2 to 6, from 4 to 10, or for example 4 or 5. In certain embodiments, y and z are each the same and selected from 4, 5, 6, 7, 8 and 9.

In some of the foregoing embodiments of structure (IX), L^1 is -

$O(C=O)R^1$, $-(C=O)OR^1$ or $-C(=O)NR^bR^c$, and L^2 is $-O(C=O)R^2$, $-(C=O)OR^2$ or -

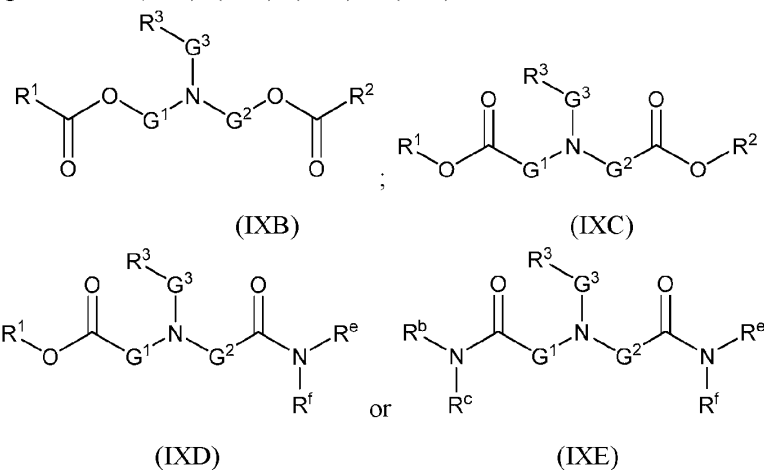
25 $C(=O)NR^eR^f$. For example, in some embodiments L^1 and L^2 are $-(C=O)OR^1$ and -

$(C=O)OR^2$, respectively. In other embodiments L^1 is $-(C=O)OR^1$ and L^2 is -

$C(=O)NR^eR^f$. In other embodiments L^1 is

$-C(=O)NR^bR^c$ and L^2 is $-C(=O)NR^eR^f$.

In other embodiments of the foregoing, the compound has one of the following structures (IXB), (IXC), (IXD) or (IXE):

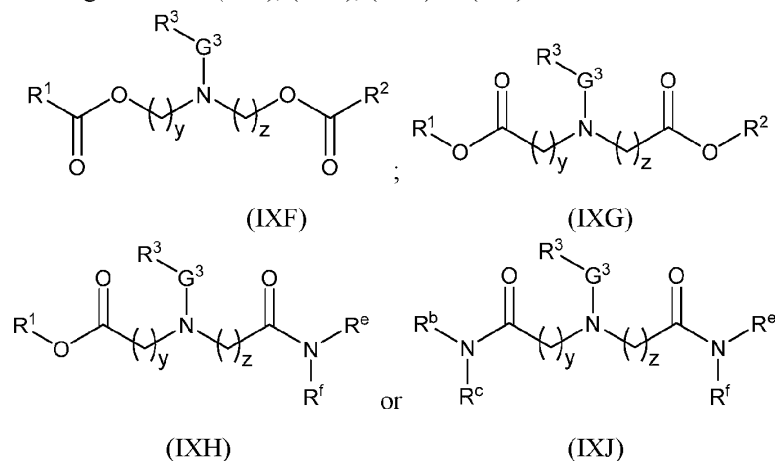


5

In some of the foregoing embodiments, the compound has structure (IXB), in other embodiments, the compound has structure (IXC) and in still other embodiments the compound has the structure (IXD). In other embodiments, the compound has structure (IXE).

10

In some different embodiments of the foregoing, the compound has one of the following structures (IXF), (IXG), (IXH) or (IXJ):

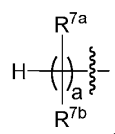


15

wherein y and z are each independently integers ranging from 2 to 12, for example an integer from 2 to 6, for example 4.

In some of the foregoing embodiments of structure (IX), y and z are each independently an integer ranging from 2 to 10, 2 to 8, from 4 to 10 or from 4 to 7. For example, in some embodiments, y is 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, z is 4, 5, 6, 7, 8, 9, 10, 11 or 12. In some embodiments, y and z are the same, while in other embodiments y and z are different.

In some of the foregoing embodiments of structure (IX), R^1 or R^2 , or both is branched C_6 - C_{24} alkyl. For example, in some embodiments, R^1 and R^2 each, independently have the following structure:



wherein:

R^{7a} and R^{7b} are, at each occurrence, independently H or C_1 - C_{12} alkyl;

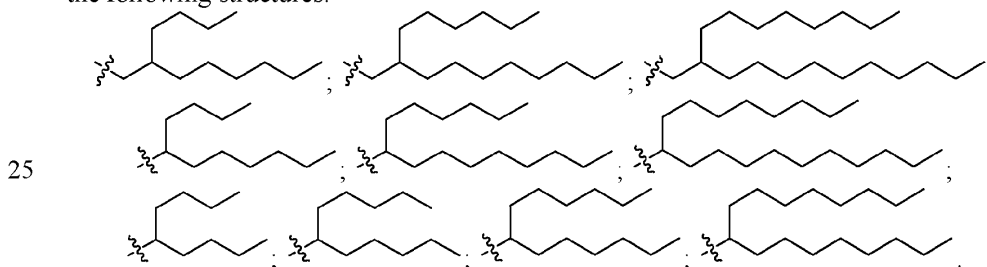
and

a is an integer from 2 to 12,

wherein R^{7a} , R^{7b} and a are each selected such that R^1 and R^2 each independently comprise from 6 to 20 carbon atoms. For example, in some embodiments a is an integer ranging from 5 to 9 or from 8 to 12.

In some of the foregoing embodiments of structure (IX), at least one occurrence of R^{7a} is H. For example, in some embodiments, R^{7a} is H at each occurrence. In other different embodiments of the foregoing, at least one occurrence of R^{7b} is C_1 - C_8 alkyl. For example, in some embodiments, C_1 - C_8 alkyl is methyl, ethyl, n-propyl, isopropyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

In different embodiments of structure (IX), R^1 or R^2 , or both, has one of the following structures:



In some of the foregoing embodiments of structure (IX), R^b, R^c, R^e and R^f are each independently C₃-C₁₂ alkyl. For example, in some embodiments R^b, R^c, R^e and R^f are n-hexyl and in other embodiments R^b, R^c, R^e and R^f are n-octyl.

In any of the foregoing embodiments of structure (IX), R⁴ is substituted or unsubstituted: methyl, ethyl, propyl, n-butyl, n-hexyl, n-octyl or n-nonyl. For example, in some embodiments R⁴ is unsubstituted. In other R⁴ is substituted with one or more substituents selected from the group consisting of -OR^g, -NR^gC(=O)R^h, -C(=O)NR^gR^h, -C(=O)R^h, -OC(=O)R^h, -C(=O)OR^h and -ORⁱOH, wherein:

R^g is, at each occurrence independently H or C₁-C₆ alkyl;

10 R^h is at each occurrence independently C₁-C₆ alkyl; and

Rⁱ is, at each occurrence independently C₁-C₆ alkylene.

In other of the foregoing embodiments of structure (IX), R⁵ is substituted: methyl, ethyl, propyl, n-butyl, n-hexyl, n-octyl or n-nonyl. In some embodiments, R⁵ is substituted ethyl or substituted propyl. In other different
15 embodiments, R⁵ is substituted with hydroxyl. In still more embodiments, R⁵ is substituted with one or more substituents selected from the group consisting of -OR^g, -NR^gC(=O)R^h, -C(=O)NR^gR^h, -C(=O)R^h, -OC(=O)R^h, -C(=O)OR^h and -ORⁱOH, wherein:

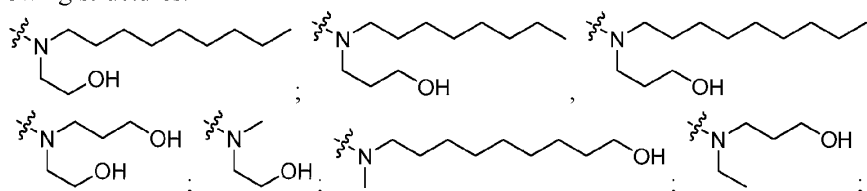
R^g is, at each occurrence independently H or C₁-C₆ alkyl;

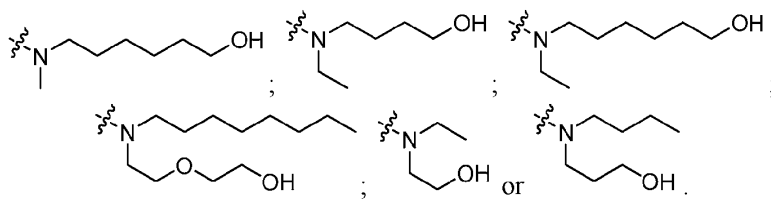
20 R^h is at each occurrence independently C₁-C₆ alkyl; and

Rⁱ is, at each occurrence independently C₁-C₆ alkylene.

In other embodiments of structure (IX), R⁴ is unsubstituted methyl, and R⁵ is substituted: methyl, ethyl, propyl, n-butyl, n-hexyl, n-octyl or n-nonyl. In some of these embodiments, R⁵ is substituted with hydroxyl.

25 In some other specific embodiments of structure (IX), R^3 has one of the following structures:



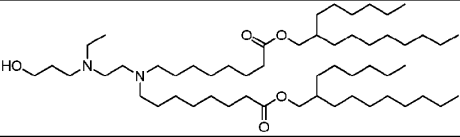
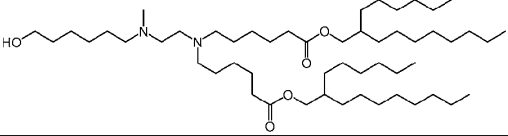
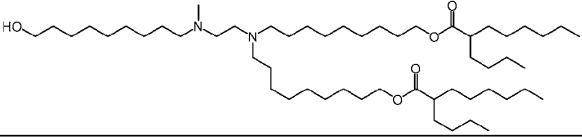
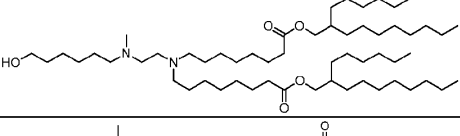
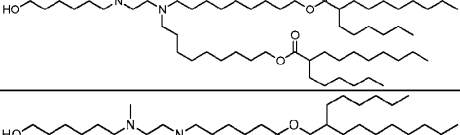
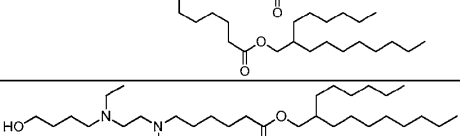
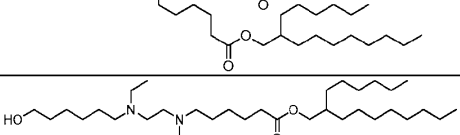
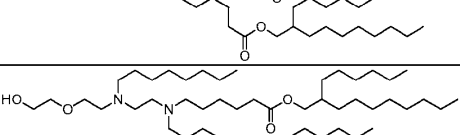
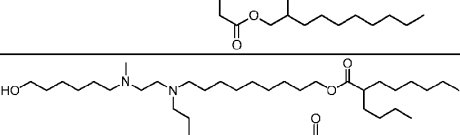
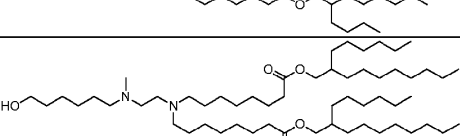
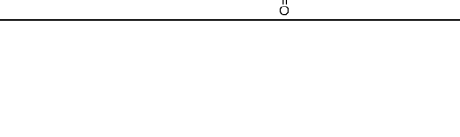


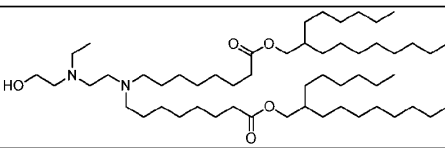
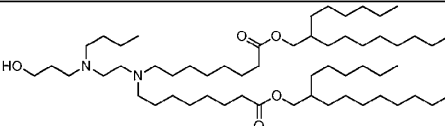
In various different embodiments of structure (IX), the cationic lipid has one of the structures set forth in Table 8 below.

5

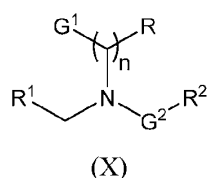
Table 8. Representative cationic lipids of structure (IX)

No.	Structure
IX-1	<p>Chemical structure IX-1: A bis-amine compound. It features a central nitrogen atom bonded to a 2-hydroxyethyl group and a 6-octyl group. This central nitrogen is also bonded to a second nitrogen atom, which is further bonded to a 6-octyl group and an ester-linked long alkyl chain (12-octadecyl). The second nitrogen atom is also bonded to another ester-linked long alkyl chain (12-octadecyl).</p>
IX-2	<p>Chemical structure IX-2: A bis-amine compound. It features a central nitrogen atom bonded to a 2-hydroxyethyl group and a 6-octyl group. This central nitrogen is also bonded to a second nitrogen atom, which is further bonded to a 6-octyl group and an ester-linked long alkyl chain (12-octadecyl). The second nitrogen atom is also bonded to another ester-linked long alkyl chain (12-octadecyl).</p>
IX-3	<p>Chemical structure IX-3: A bis-amine compound. It features a central nitrogen atom bonded to a 2-hydroxyethyl group and a 6-octyl group. This central nitrogen is also bonded to a second nitrogen atom, which is further bonded to a 6-octyl group and an ester-linked long alkyl chain (12-octadecyl). The second nitrogen atom is also bonded to another ester-linked long alkyl chain (12-octadecyl).</p>
IX-4	<p>Chemical structure IX-4: A bis-amine compound. It features a central nitrogen atom bonded to a 2-hydroxyethyl group and a 6-octyl group. This central nitrogen is also bonded to a second nitrogen atom, which is further bonded to a 6-octyl group and an ester-linked long alkyl chain (12-octadecyl). The second nitrogen atom is also bonded to another ester-linked long alkyl chain (12-octadecyl).</p>
IX-5	<p>Chemical structure IX-5: A bis-amine compound. It features a central nitrogen atom bonded to a 2-hydroxyethyl group and a 6-octyl group. This central nitrogen is also bonded to a second nitrogen atom, which is further bonded to a 6-octyl group and an ester-linked long alkyl chain (12-octadecyl). The second nitrogen atom is also bonded to another ester-linked long alkyl chain (12-octadecyl).</p>

No.	Structure
IX-6	
IX-7	
IX-8	
IX-9	
IX-10	
IX-11	
IX-12	
IX-13	
IX-14	
IX-15	
IX-16	

No.	Structure
IX-17	
IX-18	

In one embodiment, the cationic lipid is a compound having the following structure (X):



- 5 or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

G^1 is $-\text{OH}$, $-\text{NR}^3\text{R}^4$, $-(\text{C}=\text{O})\text{NR}^5$ or $-\text{NR}^3(\text{C}=\text{O})\text{R}^5$;

G^2 is $-\text{CH}_2-$ or $-(\text{C}=\text{O})-$;

R is, at each occurrence, independently H or OH;

- 10 R^1 and R^2 are each independently branched, saturated or unsaturated C_{12} - C_{36} alkyl;

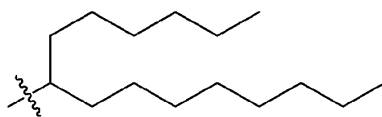
R^3 and R^4 are each independently H or straight or branched, saturated or unsaturated C_1 - C_6 alkyl;

R^5 is straight or branched, saturated or unsaturated C_1 - C_6 alkyl; and

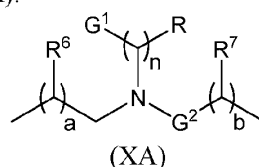
- 15 n is an integer from 2 to 6.

In some embodiments, R^1 and R^2 are each independently branched, saturated or unsaturated C_{12} - C_{30} alkyl, C_{12} - C_{20} alkyl, or C_{15} - C_{20} alkyl. In some specific embodiments, R^1 and R^2 are each saturated. In certain embodiments, at least one of R^1 and R^2 is unsaturated.

- 20 In some of the foregoing embodiments of structure (X), R^1 and R^2 have the following structure:



In some of the foregoing embodiments of structure (X), the compound has the following structure (XA):



5

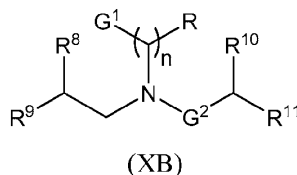
wherein:

R^6 and R^7 are, at each occurrence, independently H or straight or branched, saturated or unsaturated C_1 - C_{14} alkyl;

a and b are each independently an integer ranging from 1 to 15,

10 provided that R^6 and a, and R^7 and b, are each independently selected such that R^1 and R^2 , respectively, are each independently branched, saturated or unsaturated C_{12} - C_{36} alkyl.

In some of the foregoing embodiments, the compound has the following structure (XB):



15

wherein:

R^8 , R^9 , R^{10} and R^{11} are each independently straight or branched, saturated or unsaturated C_4 - C_{12} alkyl, provided that R^8 and R^9 , and R^{10} and R^{11} , are each
 20 independently selected such that R^1 and R^2 , respectively, are each independently branched, saturated or unsaturated C_{12} - C_{36} alkyl. In some embodiments of (XB), R^8 , R^9 , R^{10} and R^{11} are each independently straight or branched, saturated or unsaturated C_6 - C_{10} alkyl. In certain embodiments of (XB), at least one of R^8 , R^9 , R^{10} and R^{11} is unsaturated. In other certain specific embodiments of (XB), each of R^8 , R^9 , R^{10} and R^{11}
 25 is saturated.

In some of the foregoing embodiments, the compound has structure (XA), and in other embodiments, the compound has structure (XB).

In some of the foregoing embodiments, G^1 is $-OH$, and in some embodiments G^1 is $-NR^3R^4$. For example, in some embodiments, G^1 is $-NH_2$, $-NHCH_3$,
 5 or $-N(CH_3)_2$. In certain embodiments, G^1 is $-(C=O)NR^5$. In certain other embodiments, G^1 is $-NR^3(C=O)R^5$. For example, in some embodiments G^1 is $-NH(C=O)CH_3$ or $-NH(C=O)CH_2CH_2CH_3$.

In some of the foregoing embodiments of structure (X), G^2 is $-CH_2-$. In some different embodiments, G^2 is $-(C=O)-$.

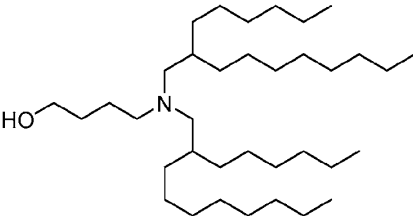
10 In some of the foregoing embodiments of structure (X), n is an integer ranging from 2 to 6, for example, in some embodiments n is 2, 3, 4, 5 or 6. In some embodiments, n is 2. In some embodiments, n is 3. In some embodiments, n is 4.

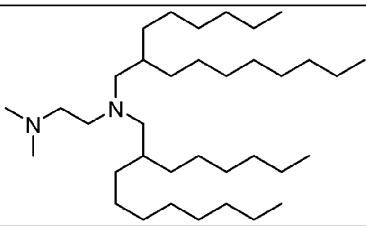
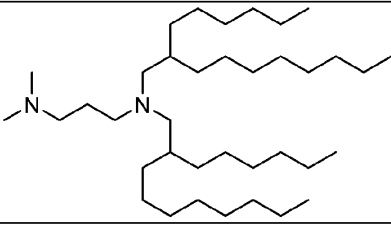
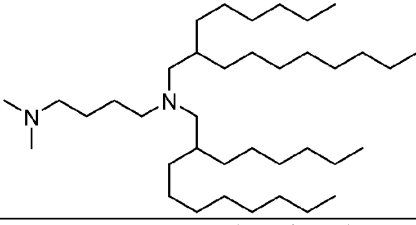
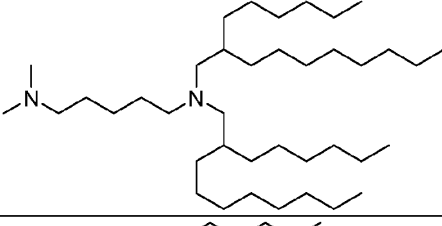
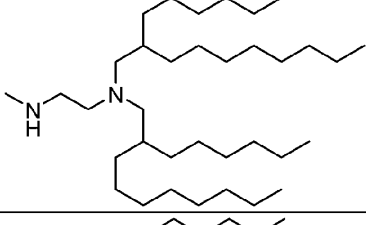
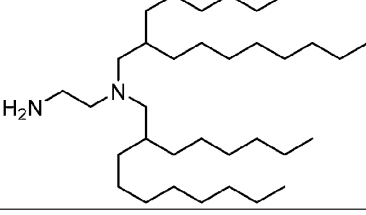
In certain of the foregoing embodiments of structure (X), at least one of R^1 , R^2 , R^3 , R^4 and R^5 is unsubstituted. For example, in some embodiments, R^1 , R^2 , R^3 ,
 15 R^4 and R^5 are each unsubstituted. In some embodiments, R^3 is substituted. In other embodiments R^4 is substituted. In still more embodiments, R^5 is substituted. In certain specific embodiments, each of R^3 and R^4 are substituted. In some embodiments, a substituent on R^3 , R^4 or R^5 is hydroxyl. In certain embodiments, R^3 and R^4 are each substituted with hydroxyl.

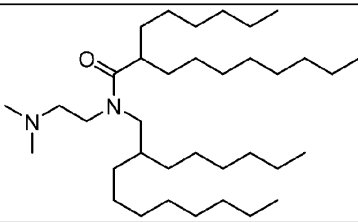
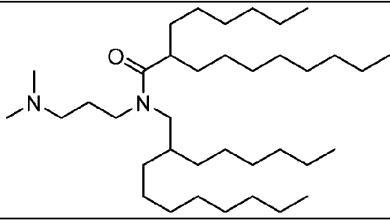
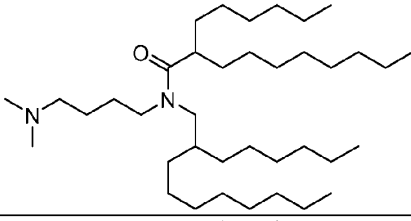
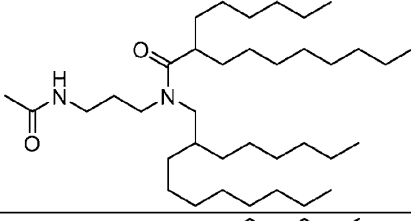
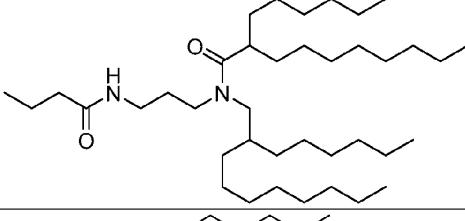
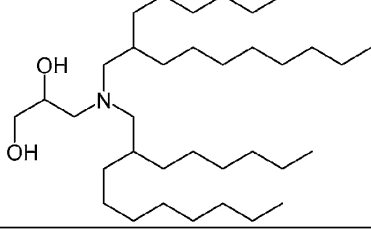
20 In some of the foregoing embodiments of structure (X), at least one R is OH . In other embodiments, each R is H .

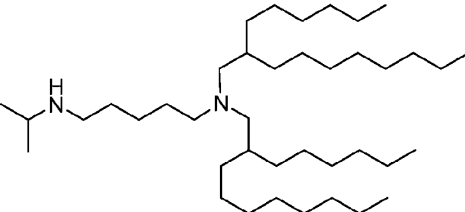
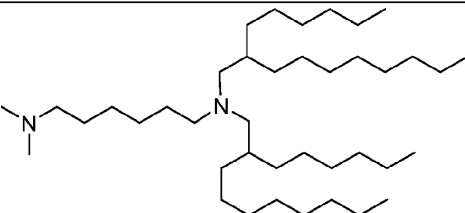
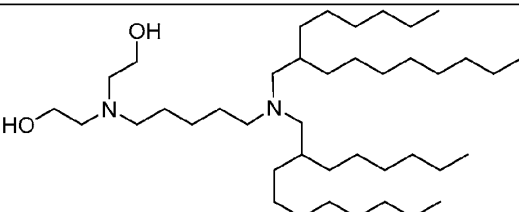
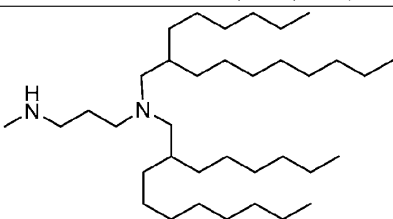
In various different embodiments of structure (X), the compound has one of the structures set forth in Table 9 below.

Table 9. Representative cationic lipids of structure (X)

No.	Structure
X-1	

No.	Structure
X-2	
X-3	
X-4	
X-5	
X-6	
X-7	

No.	Structure
X-8	
X-9	
X-10	
X-11	
X-12	
X-13	

No.	Structure
X-14	
X-15	
X-16	
X-17	

In any of Embodiments 1, 2, 3, 4 or 5, the LNPs further comprise a neutral lipid. In various embodiments, the molar ratio of the cationic lipid to the neutral lipid ranges from about 2:1 to about 8:1. In certain embodiments, the neutral lipid is present in any of the foregoing LNPs in a concentration ranging from 5 to 10 mol percent, from 5 to 15 mol percent, 7 to 13 mol percent, or 9 to 11 mol percent. In certain specific embodiments, the neutral lipid is present in a concentration of about 9.5, 10 or 10.5 mol percent. In some embodiments, the molar ratio of cationic lipid to the neutral lipid ranges from about 4.1:1.0 to about 4.9:1.0, from about 4.5:1.0 to about 4.8:1.0, or from about 4.7:1.0 to 4.8:1.0. In some embodiments, the molar ratio of total

cationic lipid to the neutral lipid ranges from about 4.1:1.0 to about 4.9:1.0, from about 4.5:1.0 to about 4.8:1.0, or from about 4.7:1.0 to 4.8:1.0.

Exemplary neutral lipids for use in any of Embodiments 1, 2, 3, 4 or 5 include, for example, distearoylphosphatidylcholine (DSPC),
 5 dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC),
 dioleoylphosphatidylglycerol (DOPG), dipalmitoylphosphatidylglycerol (DPPG),
 dioleoyl-phosphatidylethanolamine (DOPE), palmitoyl-oleoylphosphatidylcholine
 (POPC), palmitoyl-oleoyl-phosphatidylethanolamine (POPE) and dioleoyl-
 phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (DOPE-
 10 mal), dipalmitoyl phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine
 (DMPE), distearoyl-phosphatidylethanolamine (DSPE), 16-O-monomethyl PE, 16-O-
 dimethyl PE, 18-1-trans PE, 1-stearoyl-2-oleoylphosphatidylethanol amine (SOPE), and
 1,2-dielaidoyl-sn-glycero-3-phosphoethanolamine (transDOPE). In one embodiment, the
 neutral lipid is 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC). In some
 15 embodiments, the neutral lipid is selected from DSPC, DPPC, DMPC, DOPC, POPC,
 DOPE and SM. In some embodiments, the neutral lipid is DSPC.

In various embodiments of Embodiments 1, 2, 3, 4 or 5, any of the disclosed lipid nanoparticles comprise a steroid or steroid analogue. In certain
 embodiments, the steroid or steroid analogue is cholesterol. In some embodiments, the
 20 steroid is present in a concentration ranging from 39 to 49 molar percent, 40 to 46
 molar percent, from 40 to 44 molar percent, from 40 to 42 molar percent, from 42 to 44
 molar percent, or from 44 to 46 molar percent. In certain specific embodiments, the
 steroid is present in a concentration of 40, 41, 42, 43, 44, 45, or 46 molar percent.

In certain embodiments, the molar ratio of cationic lipid to the steroid
 25 ranges from 1.0:0.9 to 1.0:1.2, or from 1.0:1.0 to 1.0:1.2. In some of these
 embodiments, the molar ratio of cationic lipid to cholesterol ranges from about 5:1 to
 1:1. In certain embodiments, the steroid is present in a concentration ranging from 32
 to 40 mol percent of the steroid.

In certain embodiments, the molar ratio of total cationic to the steroid
 30 ranges from 1.0:0.9 to 1.0:1.2, or from 1.0:1.0 to 1.0:1.2. In some of these

embodiments, the molar ratio of total cationic lipid to cholesterol ranges from about 5:1 to 1:1. In certain embodiments, the steroid is present in a concentration ranging from 32 to 40 mol percent of the steroid.

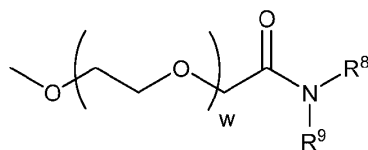
In some embodiments of Embodiments 1, 2, 3 4 or 5, the LNPs further
 5 comprise a polymer conjugated lipid. In various other embodiments of Embodiments 1, 2, 3 4 or 5, the polymer conjugated lipid is a pegylated lipid. For example, some embodiments include a pegylated diacylglycerol (PEG-DAG) such as 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG-DMG), a pegylated phosphatidylethanolamine (PEG-PE), a PEG succinate diacylglycerol (PEG-S-DAG) such as 4-O-(2',3'-di(tetradecanoyloxy)propyl-1-O-(ω -
 10 methoxy(polyethoxy)ethyl)butanedioate (PEG-S-DMG), a pegylated ceramide (PEG-cer), or a PEG dialkoxypentylcarbamate such as ω -methoxy(polyethoxy)ethyl-N-(2,3-di(tetradecanoyloxy)propyl)carbamate or 2,3-di(tetradecanoyloxy)propyl-N-(ω -methoxy(polyethoxy)ethyl)carbamate.

15 In various embodiments, the polymer conjugated lipid is present in a concentration ranging from 1.0 to 2.5 molar percent. In certain specific embodiments, the polymer conjugated lipid is present in a concentration of about 1.7 molar percent. In some embodiments, the polymer conjugated lipid is present in a concentration of about 1.5 molar percent.

20 In certain embodiments, the molar ratio of cationic lipid to the polymer conjugated lipid ranges from about 35:1 to about 25:1. In some embodiments, the molar ratio of cationic lipid to polymer conjugated lipid ranges from about 100:1 to about 20:1.

In certain embodiments, the molar ratio of total cationic lipid (*i.e.*, the
 25 sum of the first and second cationic lipid) to the polymer conjugated lipid ranges from about 35:1 to about 25:1. In some embodiments, the molar ratio of total cationic lipid to polymer conjugated lipid ranges from about 100:1 to about 20:1.

In some embodiments of Embodiments 1, 2, 3 4 or 5, the pegylated lipid, when present, has the following Formula (XI):



(XI)

or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof,

wherein:

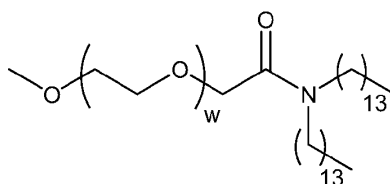
5 R¹² and R¹³ are each independently a straight or branched, saturated or unsaturated alkyl chain containing from 10 to 30 carbon atoms, wherein the alkyl chain is optionally interrupted by one or more ester bonds; and

w has a mean value ranging from 30 to 60.

In some embodiments, R^{12} and R^{13} are each independently straight, saturated alkyl chains containing from 12 to 16 carbon atoms. In other embodiments, the average w ranges from 42 to 55, for example, the average w is 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54 or 55. In some specific embodiments, the average w is about 49.

In some embodiments, the pegylated lipid has the following Formula

15 (XIa):



(XIa)

wherein the average w is about 49.

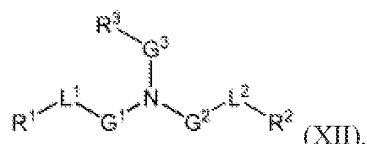
In some embodiments of Embodiments 1, 2, 3 4 or 5, the nucleic acid is
20 selected from antisense and messenger RNA. For example, messenger RNA may be
used to induce an immune response (*e.g.*, as a vaccine), for example by translation of
immunogenic proteins.

In other embodiments of Embodiments 1, 2, 3 4 or 5, the nucleic acid is mRNA, and the mRNA to lipid ratio in the LNP (*i.e.*, N/P, where N represents the moles of cationic lipid and P represents the moles of phosphate present as part of the nucleic

[0414] In an embodiment, the transfer vehicle comprises a lipid or an ionizable lipid described in US patent publication number 20190314524.

[0415] Some embodiments of the present invention provide nucleic acid-lipid nanoparticle compositions comprising one or more of the novel cationic lipids described herein as structures listed in Table 10, that provide increased activity of the nucleic acid and improved tolerability of the compositions *in vivo*.

[0416] In one embodiment, an ionizable lipid has the following structure (XII):



or a pharmaceutically acceptable salt, tautomer, prodrug or stereoisomer thereof, wherein:

one of L¹ or L² is —O(C=O)—, —(C=O)O—, —C(=O)—, —O—, —S(O)_x—, —S—, —S—, —C(=O)S—, SC(=O)—, —NR^aC(=O)—, —C(=O)NR^a—, NR^aC(=O)NR^a—, —OC(=O)NR^a— or —NR^aC(=O)O—, and the other of L¹ or L² is —O(C=O)—, —(C=O)O—, —C(=O)—, —O—, —S(O)_x—, —S—S—, —C(=O)S—, SC(=O)—, —NR^aC(=O)—, —C(=O)NR^a—, NR^aC(=O)NR^a—, —OC(=O)NR^a— or —NR^aC(=O)O— or a direct bond;

G¹ and G² are each independently unsubstituted C₁-C₁₂ alkylene or C₁-C₁₂ alkenylene;

G³ is C₁-C₂₄ alkylene, C₁-C₂₄ alkenylene, C₃-C₈ cycloalkylene, C₃-C₈ cycloalkenylene;

R^a is H or C₁-C₁₂ alkyl;

R¹ and R² are each independently C₆-C₂₄ alkyl or C₆-C₂₄ alkenyl;

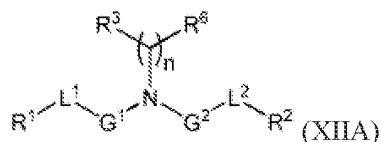
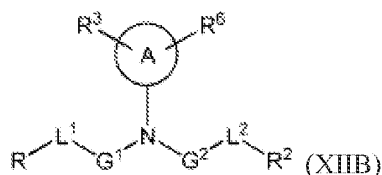
R^3 is H, OR^5 , CN, $-C(=O)OR^4$, $-OC(=O)R^4$ or $-NR^5C(=O)R^4$.

R⁴ is C₁-C₁₂ alkyl;

R⁵ is H or C₁-C₆ alkyl; and

x is 0, 1 or 2.

[0417] In some embodiments, an ionizable lipid has one of the following structures (XIIA) or (XIIB):



wherein:

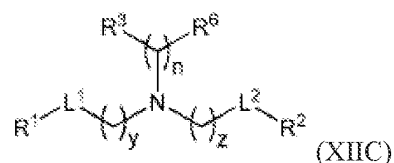
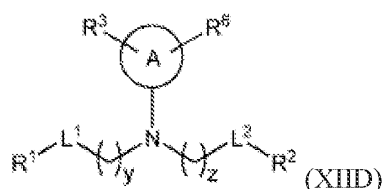
A is a 3 to 8-membered cycloalkyl or cycloalkylene ring;

R⁶ is, at each occurrence, independently H, OH or C₁-C₂₄ alkyl; and

n is an integer ranging from 1 to 15.

[0418] In some embodiments, the ionizable lipid has structure (XIIA), and in other embodiments, the ionizable lipid has structure (XIIB).

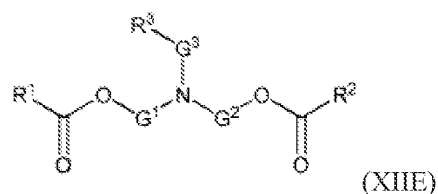
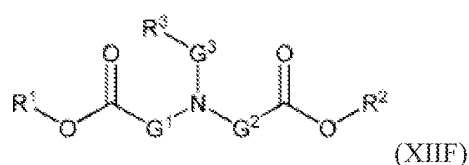
[0419] In other embodiments, an ionizable lipid has one of the following structures (XIIC) or (XIID):



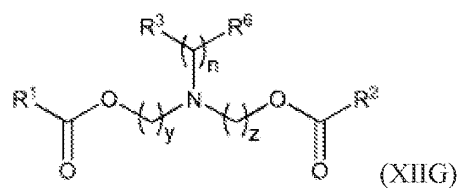
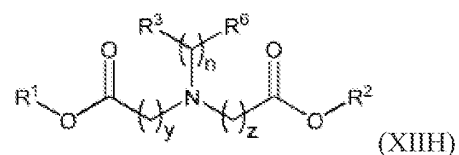
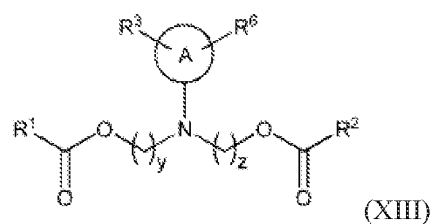
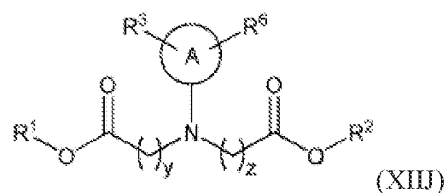
wherein y and z are each independently integers ranging from 1 to 12.

[0420] In some embodiments, one of L¹ or L² is —O(C=O)—. For example, in some embodiments each of L¹ and L² are —O(C=O)—. In some different embodiments of any of the foregoing, L¹ and L² are each independently —(C=O)O— or —O(C=O)—. For example, in some embodiments each of L¹ and L² is —(C=O)O—.

[0421] In some embodiments, an ionizable lipid has one of the following structures (XIIIE) or (XIIIF):



[0422] In some embodiments, an ionizable lipid has one of the following structures (XIIIG), (XIIH), (XIIII), or (XIIJ):



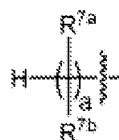
[0423] In some embodiments, n is an integer ranging from 2 to 12, for example from 2 to 8 or from 2 to 4. For example, in some embodiments, n is 3, 4, 5 or 6. In some embodiments, n is 3. In some embodiments, n is 4. In some embodiments, n is 5. In some embodiments, n is 6.

[0424] In some embodiments, y and z are each independently an integer ranging from 2 to 10. For example, in some embodiments, y and z are each independently an integer ranging from 4 to 9 or from 4 to 6.

[0425] In some embodiments, R^6 is H. In other embodiments, R^6 is C_1 - C_{24} alkyl. In other embodiments, R^6 is OH.

[0426] In some embodiments, G^3 is unsubstituted. In other embodiments, G^3 is substituted. In various different embodiments, G^3 is linear C_1 - C_{24} alkylene or linear C_1 - C_{24} alkenylene.

[0427] In some embodiments, R^1 or R^2 , or both, is C_6 - C_{24} alkenyl. For example, in some embodiments, R^1 and R^2 each, independently have the following structure:



wherein:

R^{7a} and R^{7b} are, at each occurrence, independently H or C₁-C₁₂ alkyl; and

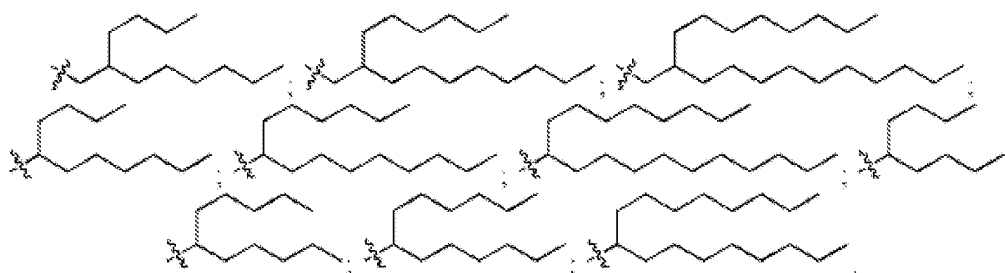
a is an integer from 2 to 12,

wherein R^{7a} , R^{7b} and a are each selected such that R^1 and R^2 each independently comprise from 6 to 20 carbon atoms.

[0428] In some embodiments, a is an integer ranging from 5 to 9 or from 8 to 12.

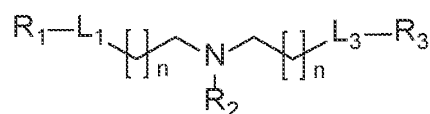
[0429] In some embodiments, at least one occurrence of R^{7a} is H. For example, in some embodiments, R^{7a} is H at each occurrence. In other different embodiments, at least one occurrence of R^{7b} is C₁-C₈ alkyl. For example, in some embodiments, C₁-C₈ alkyl is methyl, ethyl, n-propyl, iso-propyl, n-butyl, iso-butyl, tert-butyl, n-hexyl or n-octyl.

[0430] In different embodiments, R^1 or R^2 , or both, has one of the following structures:



[0431] In some embodiments, R^3 is —OH, —CN, —C(=O)OR⁴, —OC(=O)R⁴ or —NHC(=O)R⁴. In some embodiments, R^4 is methyl or ethyl.

[0432] In some embodiments, an ionizable lipid is a compound of Formula (1):



Formula (1),

wherein:

each n is independently 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15; and

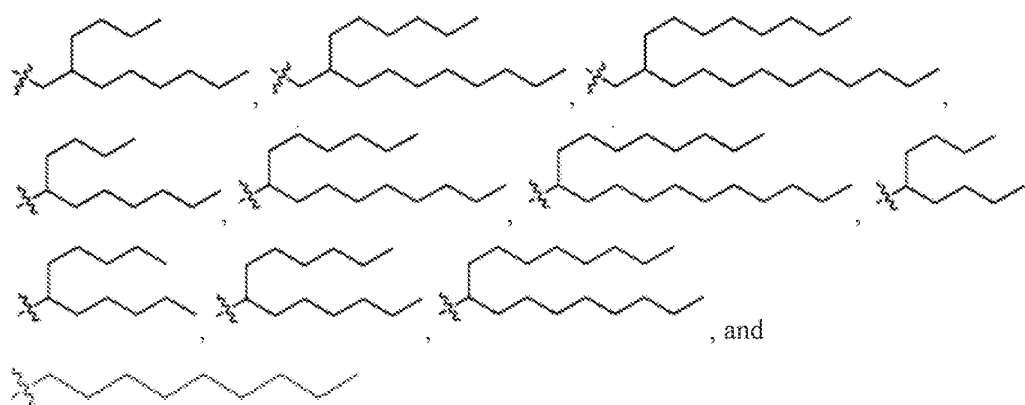
L_1 and L_3 are each independently —OC(O)—* or —C(O)O—*, wherein “*” indicates the attachment point to R_1 or R_3 ;

R_1 and R_3 are each independently a linear or branched C₉-C₂₀ alkyl or C₉-C₂₀ alkenyl,

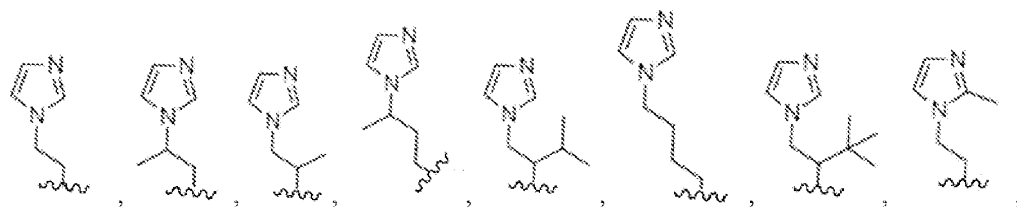
optionally substituted by one or more substituents selected from oxo, halo, hydroxy, cyano, alkyl, alkenyl, aldehyde, heterocyclalkyl, hydroxyalkyl, dihydroxyalkyl, hydroxyalkylaminoalkyl, aminoalkyl, alkylaminoalkyl, dialkylaminoalkyl, (heterocycl)(alkyl)aminoalkyl, heterocycl, heteroaryl, alkylheteroaryl, alkynyl, alkoxy, amino, dialkylamino, aminoalkylcarbonylamino, aminocarbonylalkylamino, (aminocarbonylalkyl)(alkyl)amino, alkenylcarbonylamino, hydroxycarbonyl, alkyloxy, aminocarbonyl, aminoalkylaminocarbonyl, alkylaminoalkylaminocarbonyl, dialkylaminoalkylaminocarbonyl, heterocyclalkylaminocarbonyl, (alkylaminoalkyl)(alkyl)aminocarbonyl, alkylaminoalkylcarbonyl, dialkylaminoalkylcarbonyl, heterocyclcarbonyl, alkenylcarbonyl, alkynylcarbonyl, alkylsulfoxide, alkylsulfoxidealkyl, alkylsulfonyl, and alkylsulfonealkyl.

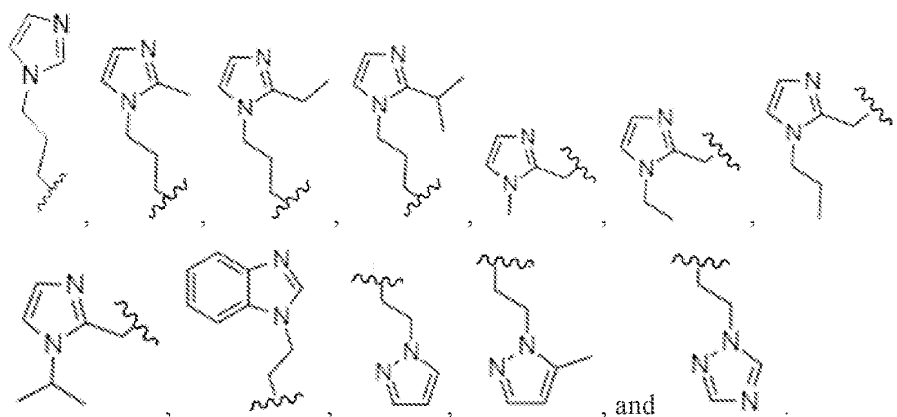
[0433] In some embodiments, R_1 and R_3 are the same. In some embodiments, R_1 and R_3 are different.

[0434] In some embodiments, R_1 and R_3 are each independently a branched saturated C_9 - C_{20} alkyl. In some embodiments, one of R_1 and R_3 is a branched saturated C_9 - C_{20} alkyl, and the other is an unbranched saturated C_9 - C_{20} alkyl. In some embodiments, R_1 and R_3 are each independently selected from a group consisting of:



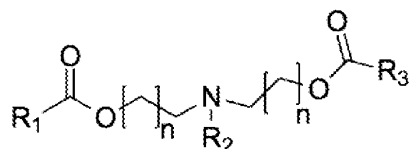
[0435] In various embodiments, R_2 is selected from a group consisting of:



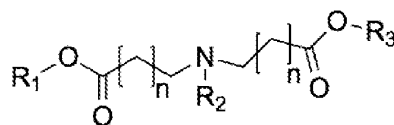


[0436] In some embodiments, R₂ may be as described in International Pat. Pub. No. WO2019/152848 A1, which is incorporated herein by reference in its entirety.

[0437] In some embodiments, an ionizable lipid is a compound of Formula (1-1) or Formula (1-2):



Formula (1-1)



Formula (1-2)

wherein n , R_1 , R_2 , and R_3 are as defined in Formula (1).

[0438] Preparation methods for the above compounds and compositions are described herein below and/or known in the art.

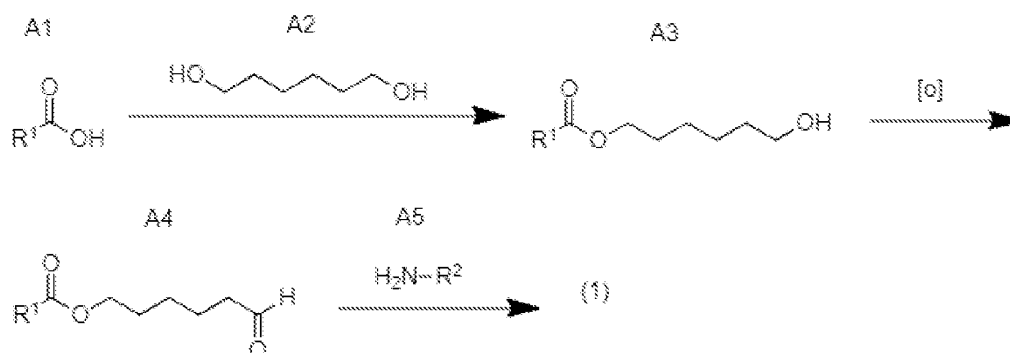
[0439] It will be appreciated by those skilled in the art that in the process described herein the functional groups of intermediate compounds may need to be protected by suitable protecting groups. Such functional groups include, *e.g.*, hydroxyl, amino, mercapto, and carboxylic acid. Suitable protecting groups for hydroxyl include, *e.g.*, trialkylsilyl or diarylalkylsilyl (for example, *t*-butyldimethylsilyl, *t*-butyldiphenylsilyl or trimethylsilyl), tetrahydropyranyl, benzyl, and the like. Suitable protecting groups for amino, amidino, and guanidino include, *e.g.*, *t*-butoxycarbonyl, benzyloxycarbonyl, and the like. Suitable protecting groups for mercapto include, *e.g.*, $-C(O)-R''$ (where R'' is alkyl, aryl, or arylalkyl), *p*-methoxybenzyl, trityl, and the like. Suitable protecting groups for carboxylic

acid include, *e.g.*, alkyl, aryl, or arylalkyl esters. Protecting groups may be added or removed in accordance with standard techniques, which are known to one skilled in the art and as described herein. The use of protecting groups is described in detail in, *e.g.*, Green, T. W. and P. G. M. Wutz, *Protective Groups in Organic Synthesis* (1999), 3rd Ed., Wiley. As one of skill in the art would appreciate, the protecting group may also be a polymer resin such as a Wang resin, Rink resin, or a 2-chlorotrityl-chloride resin.

[0440] It will also be appreciated by those skilled in the art, although such protected derivatives of compounds of this invention may not possess pharmacological activity as such, they may be administered to a mammal and thereafter metabolized in the body to form compounds of the invention which are pharmacologically active. Such derivatives may therefore be described as prodrugs. All prodrugs of compounds of this invention are included within the scope of the invention.

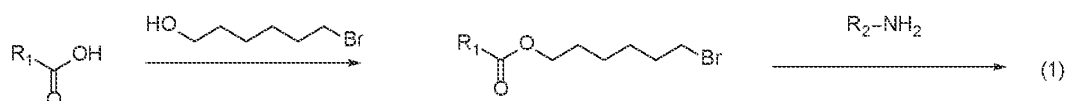
[0441] Furthermore, all compounds of the invention which exist in free base or acid form can be converted to their pharmaceutically acceptable salts by treatment with the appropriate inorganic or organic base or acid by methods known to one skilled in the art. Salts of the compounds of the invention can also be converted to their free base or acid form by standard techniques.

[0442] The following reaction scheme illustrates an exemplary method to make compounds of Formula (1):



A1 are purchased or prepared according to methods known in the art. Reaction of A1 with diol A2 under appropriate condensation conditions (*e.g.*, DCC) yields ester/alcohol A3, which can then be oxidized (*e.g.*, with PCC) to aldehyde A4. Reaction of A4 with amine A5 under reductive amination conditions yields a compound of Formula (1).

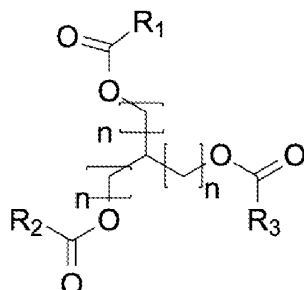
[0443] The following reaction scheme illustrates a second exemplary method to make compounds of Formula (1), wherein R_1 and R_3 are the same:



[0444] Modifications to the above reaction scheme, such as using protecting groups, may yield compounds wherein R₁ and R₃ are different. The use of protecting groups, as well as other modification methods, to the above reaction scheme will be readily apparent to one of ordinary skill in the art.

[0445] It is understood that one skilled in the art may be able to make these compounds by similar methods or by combining other methods known to one skilled in the art. It is also understood that one skilled in the art would be able to make other compounds of Formula (1) not specifically illustrated herein by using the appropriate starting materials and modifying the parameters of the synthesis. In general, starting materials may be obtained from sources such as Sigma Aldrich, Lancaster Synthesis, Inc., Maybridge, Matrix Scientific, TCI, and Fluorochem USA, *etc.* or synthesized according to sources known to those skilled in the art (*see, e.g.*, Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, 5th edition (Wiley, December 2000)) or prepared as described in this invention.

[0446] In some embodiments, an ionizable lipid is a compound of Formula (2):

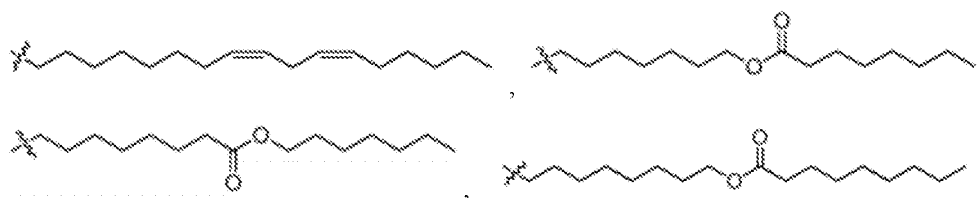


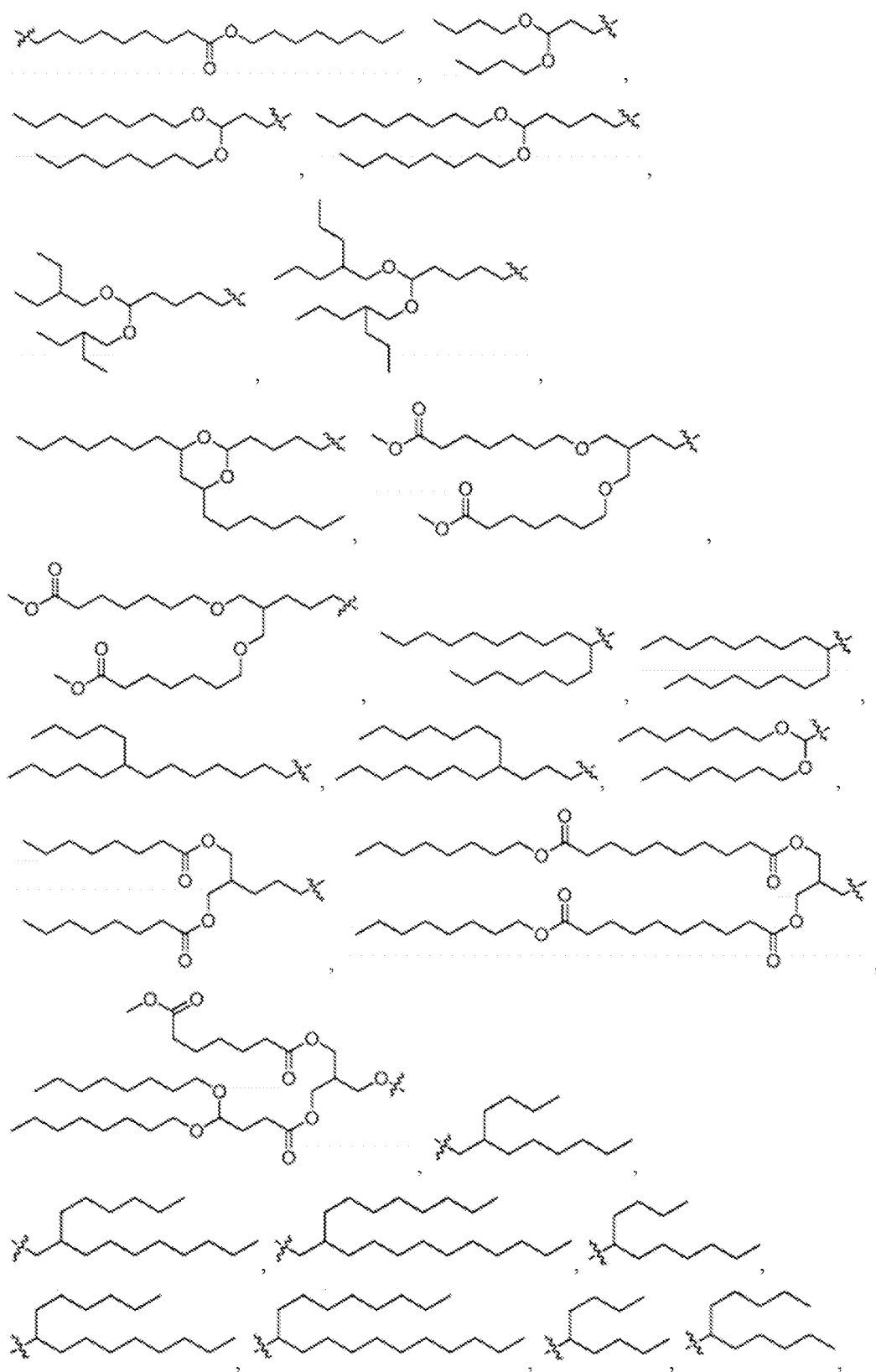
Formula (2).

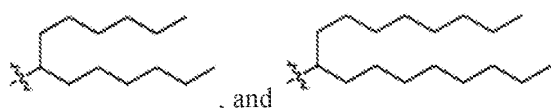
wherein each n is independently 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15.

[0447] In some embodiments, as used in Formula (2), R₁ and R₂ are as defined in Formula (1).

[0448] In some embodiments, as used in Formula (2), R₁ and R₂ are each independently selected from a group consisting of:

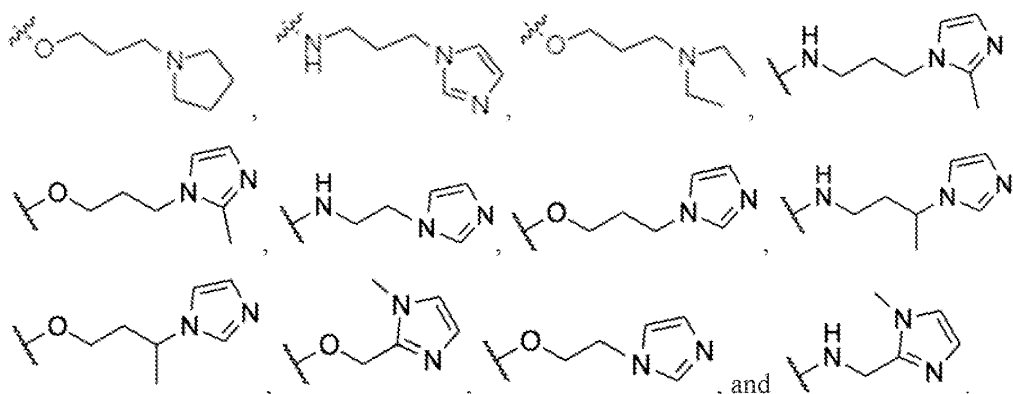




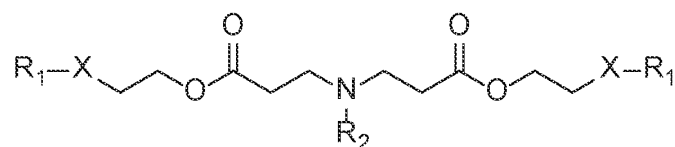


[0449] In some embodiments, R_1 and/or R_2 as used in Formula (2) may be as described in International Pat. Pub. No. WO2015/095340 A1, which is incorporated herein by reference in its entirety. In some embodiments, R_1 as used in Formula (2) may be as described in International Pat. Pub. No. WO2019/152557 A1, which is incorporated herein by reference in its entirety.

[0450] In some embodiments, as used in Formula (2), R_3 is selected from a group consisting of:



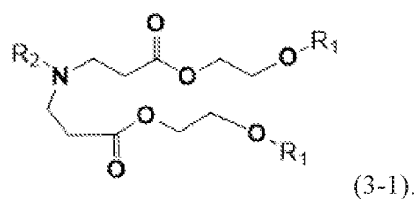
[0451] In some embodiments, an ionizable lipid is a compound of Formula (3)



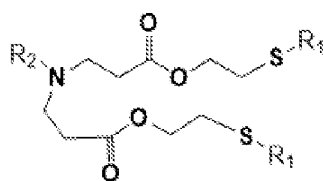
wherein X is selected from $-O-$,

$-S-$, or $-OC(O)-*$, wherein * indicates the attachment point to R_1 .

[0452] In some embodiments, an ionizable lipid is a compound of Formula (3-1):

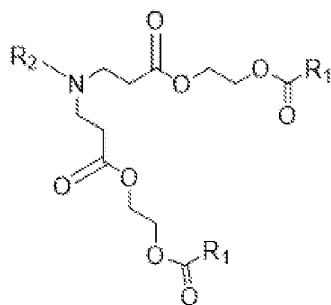


[0453] In some embodiments, an ionizable lipid is a compound of Formula (3-2):



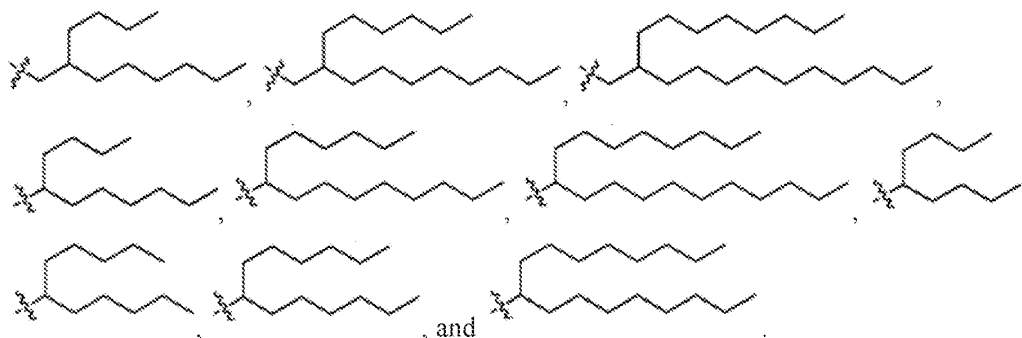
(3-2).

[0454] In some embodiments, an ionizable lipid is a compound of Formula (3-3):



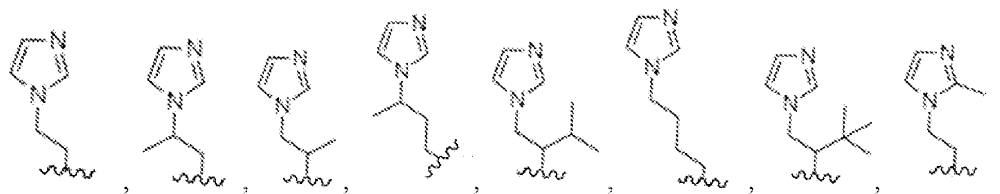
(3-3).

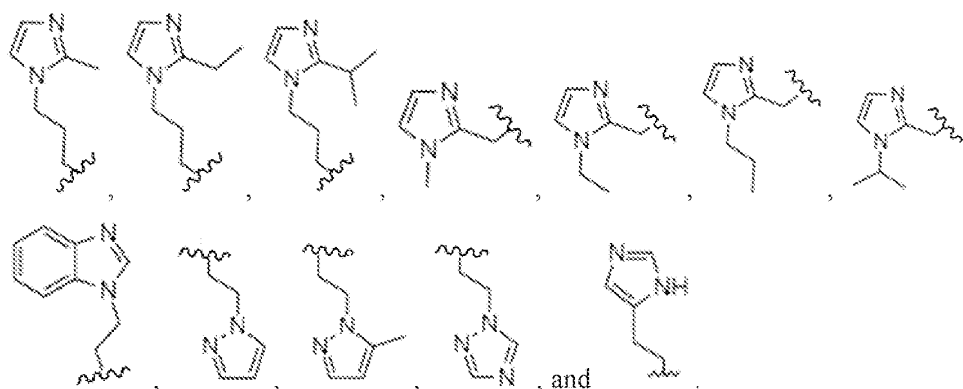
[0455] In some embodiments, as used in Formula (3-1), (3-2), or (3-3), each R₁ is independently a branched saturated C₉-C₂₀ alkyl. In some embodiments, each R₁ is independently selected from a group consisting of:



[0456] In some embodiments, each R₁ in Formula (3-1), (3-2), or (3-3) are the same.

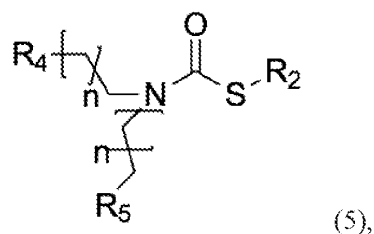
[0457] In some embodiments, as used in Formula (3-1), (3-2), or (3-3), R₂ is selected from a group consisting of:





[0458] In some embodiments, R_2 as used in Formula (3-1), (3-2), or (3-3) may be as described in International Pat. Pub. No. WO2019/152848A1, which is incorporated herein by reference in its entirety.

[0459] In some embodiments, an ionizable lipid is a compound of Formula (5):



wherein:

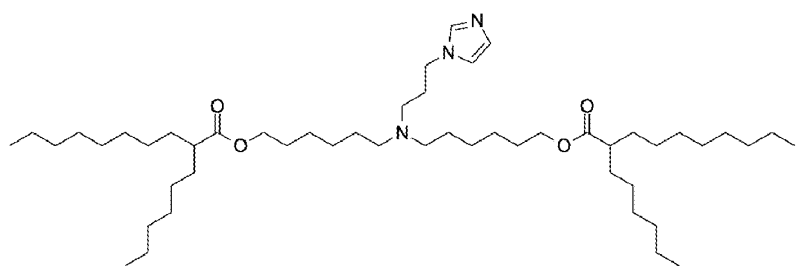
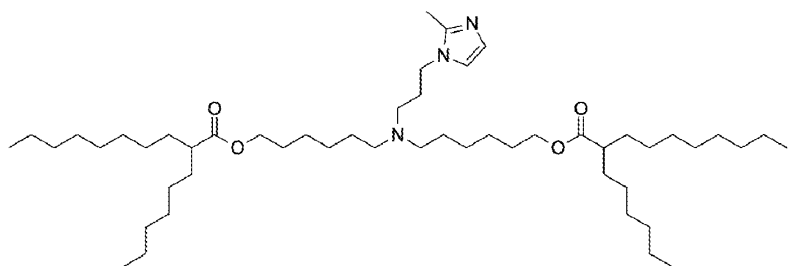
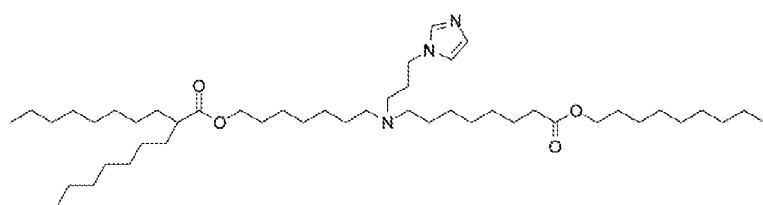
each n is independently 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15; and

R_2 is as defined in Formula (1).

[0460] In some embodiments, as used in Formula (5), R_4 and R_5 are defined as R_1 and R_3 , respectively, in Formula (1). In some embodiments, as used in Formula (5), R_4 and R_5 may be as described in International Pat. Pub. No. WO2019/191780 A1, which is incorporated herein by reference in its entirety.

[0461] In some embodiments, an ionizable lipid of the disclosure is selected from Table 10a. In some embodiments, the ionizable lipid is Lipid 26 in Table 10a. In some embodiments, the ionizable lipid is Lipid 27 in Table 10a. In some embodiments, the ionizable lipid is Lipid 53 in Table 10a. In some embodiments, the ionizable lipid is Lipid 54 in Table 10a.

[0462] In some embodiments, an ionizable lipid of the disclosure is selected from the group consisting of:



, and

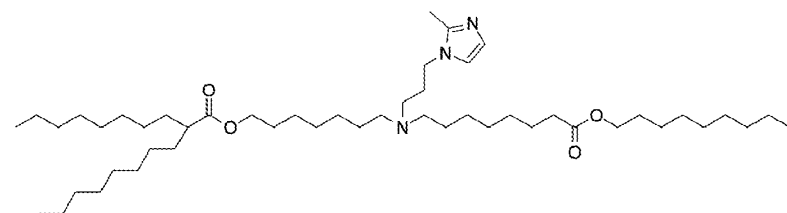
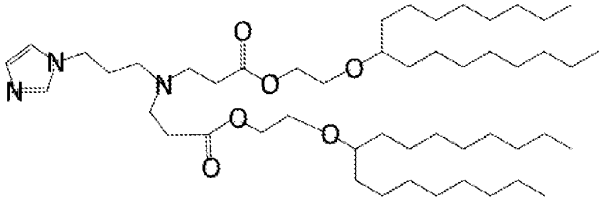
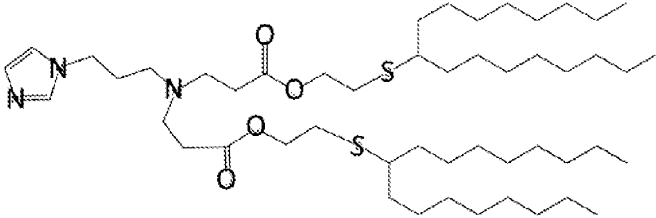
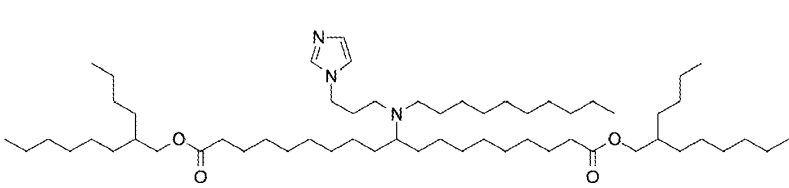
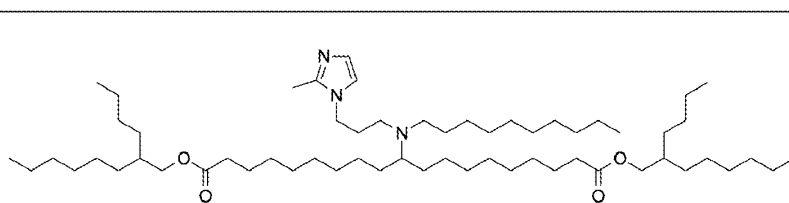
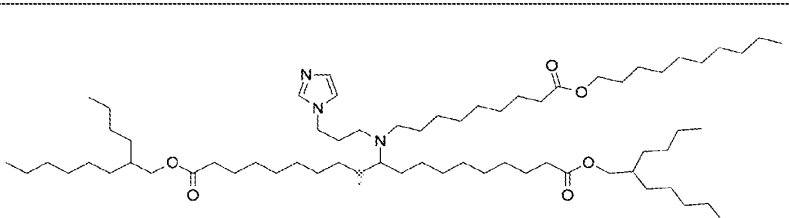
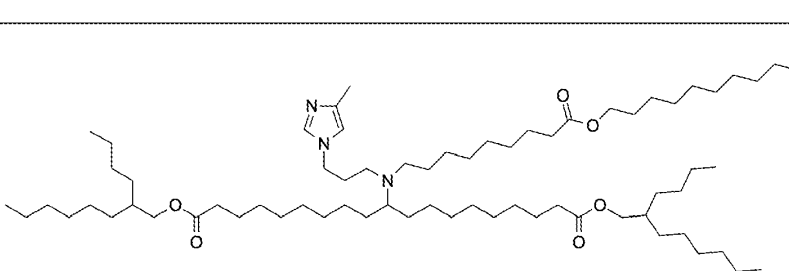
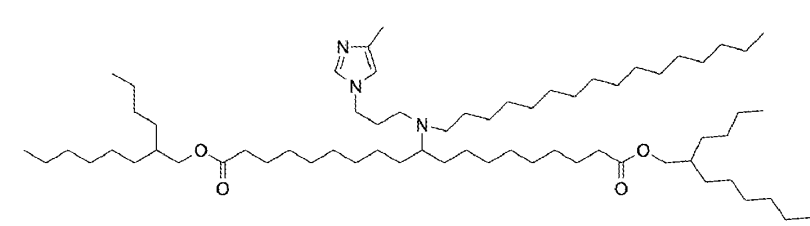
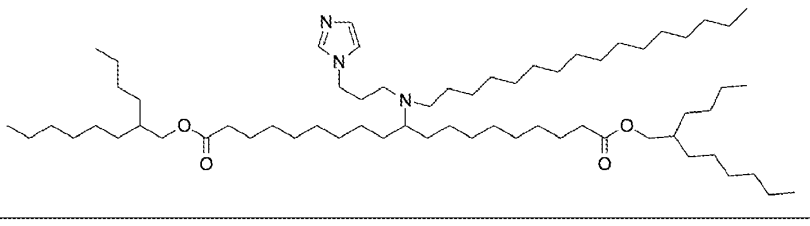
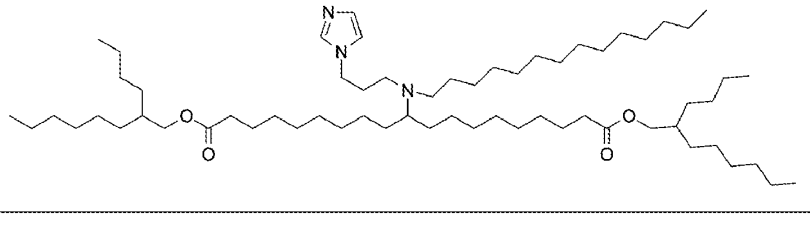
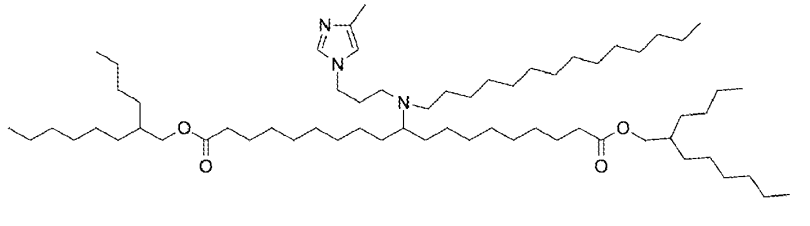
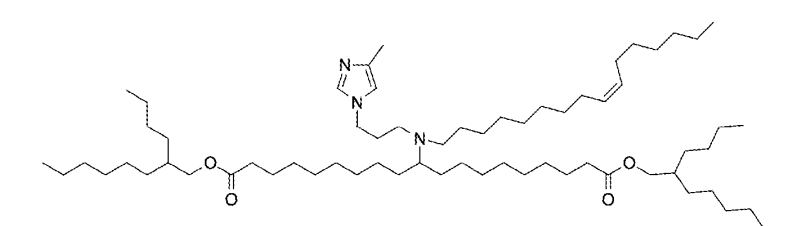
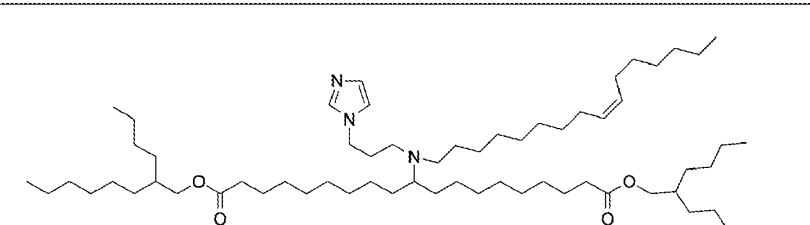
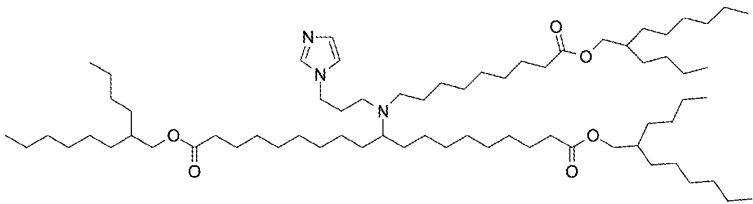
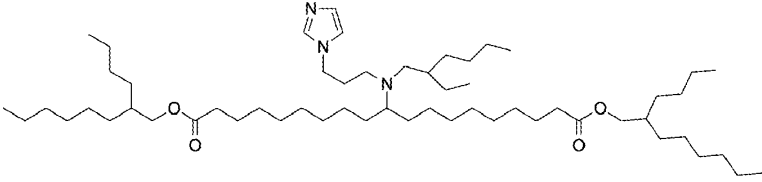
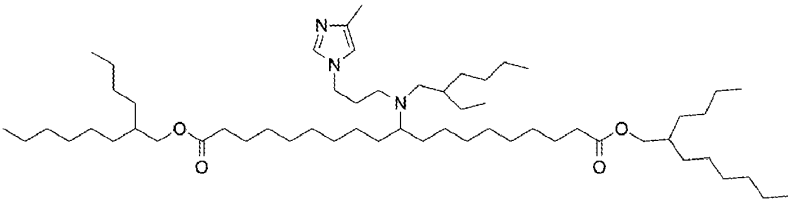
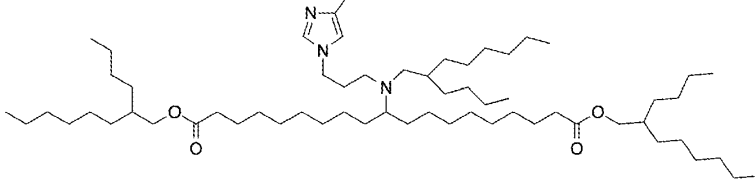
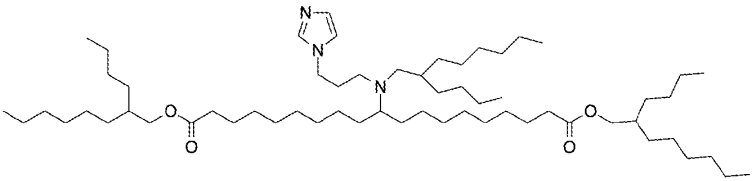
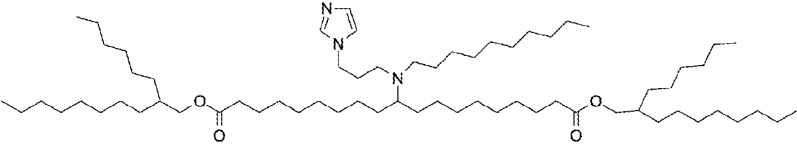


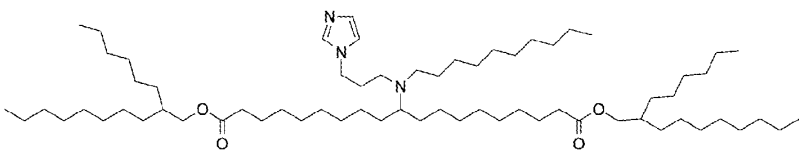
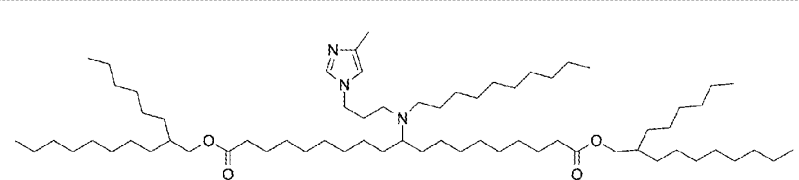
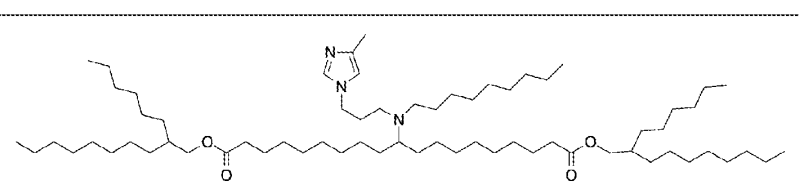
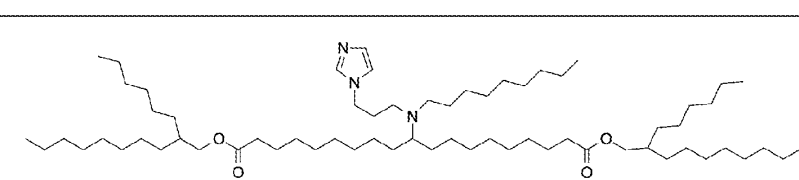
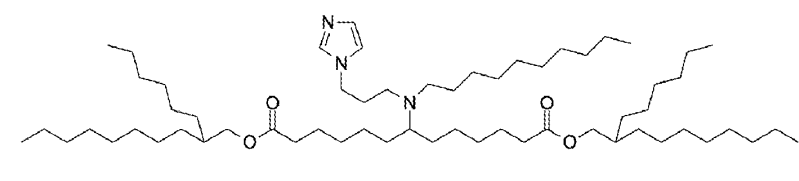
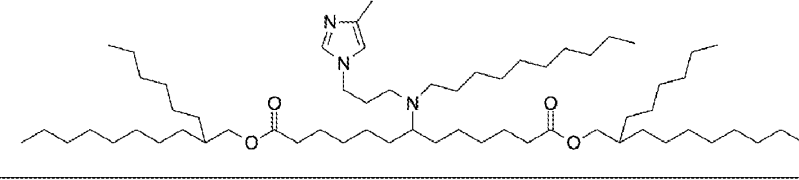
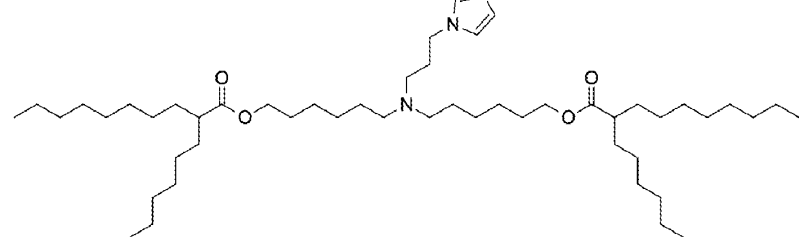
Table 10a

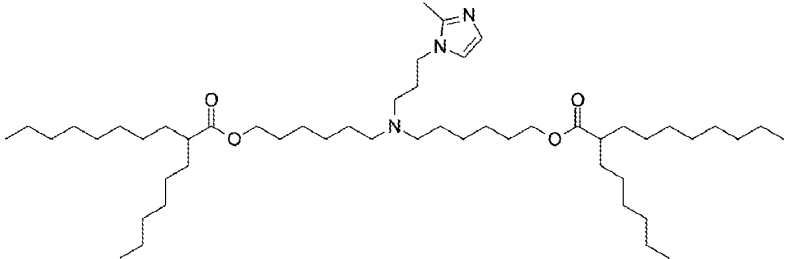
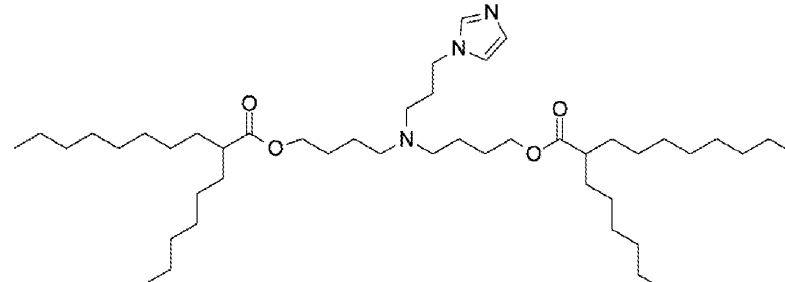
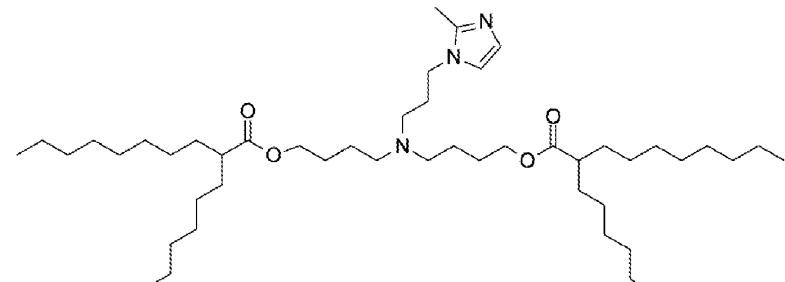
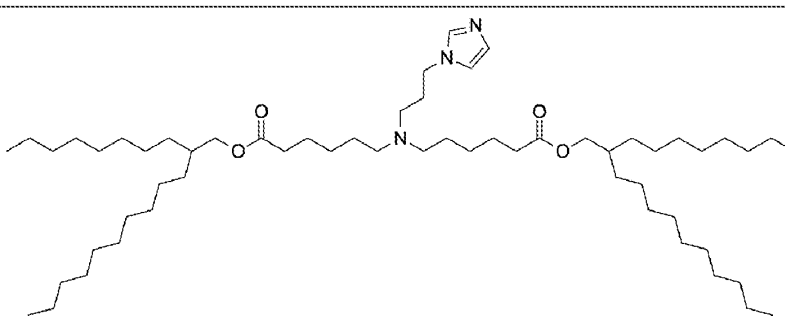
Ionizable lipid number	Structure
1	<p>Chemical structure of lipid 1. It features a central nitrogen atom connected to two long alkyl chains (one 10-carbon and one 12-carbon) and a 1,2,4-triazole ring. The molecule is shown in a simplified representation with wavy lines for the alkyl chains.</p>

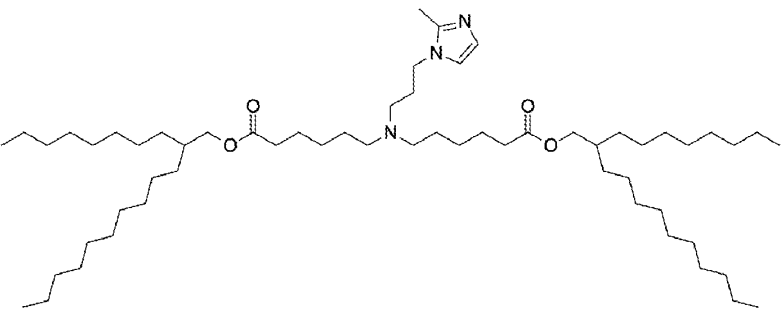
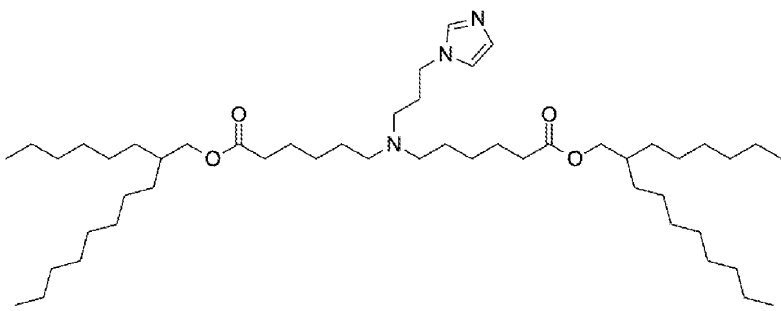
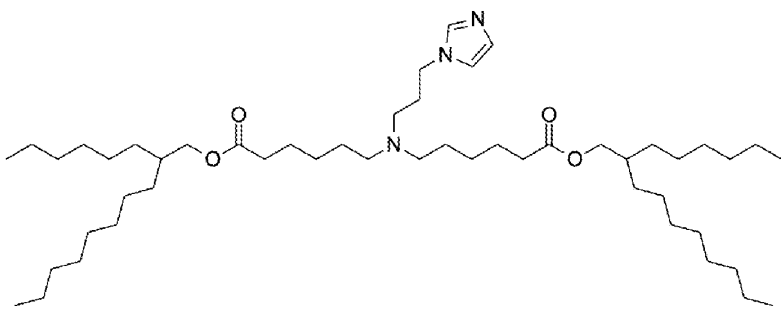
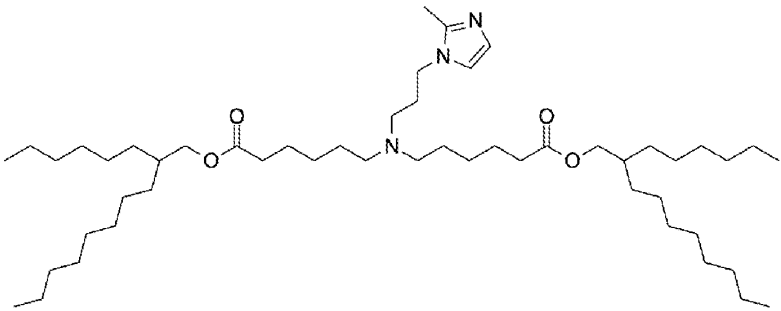
2	
3	
4	
5	
6	
7	

8	
9	
10	
11	
12	
13	

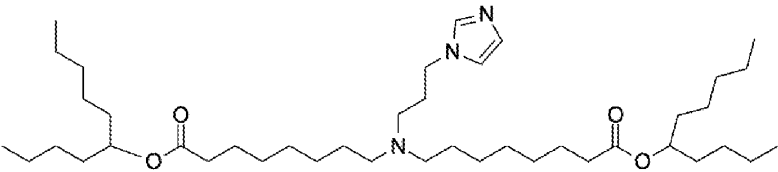
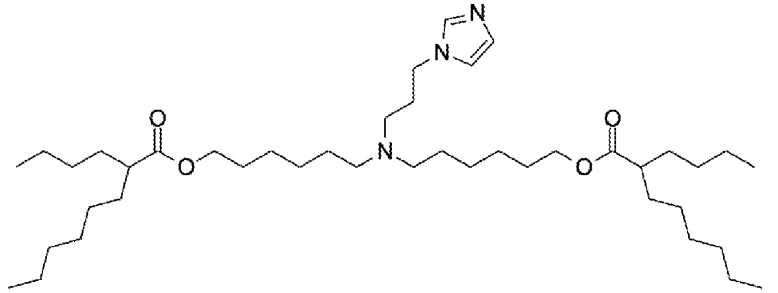
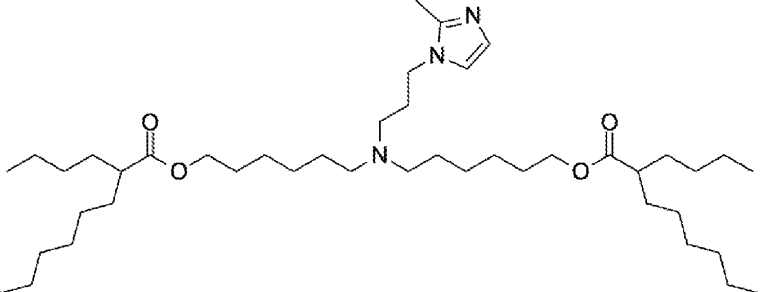
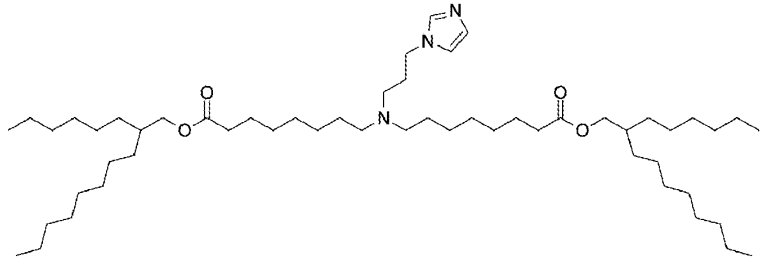
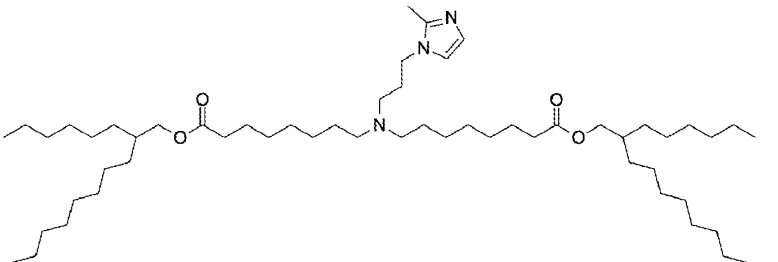
14	
15	
16	
17	
18	
19	

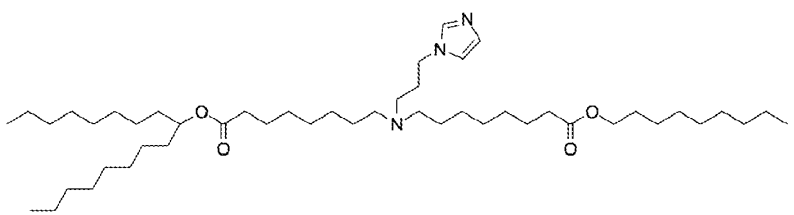
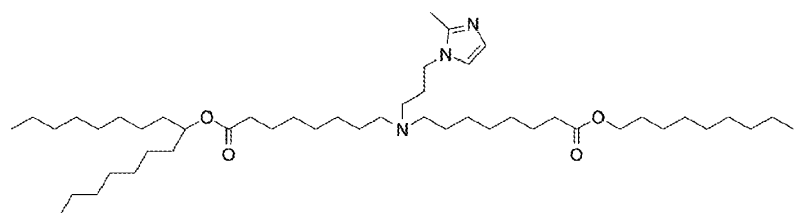
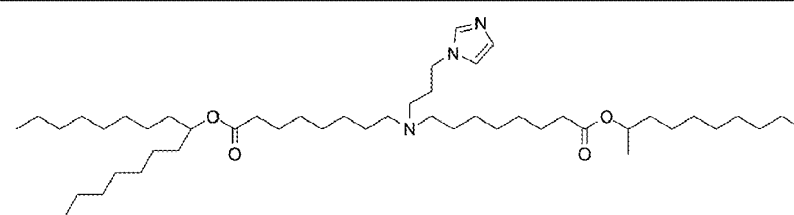
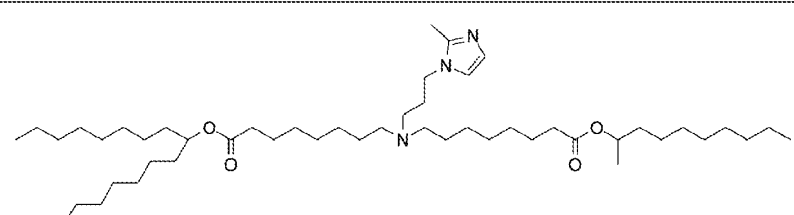
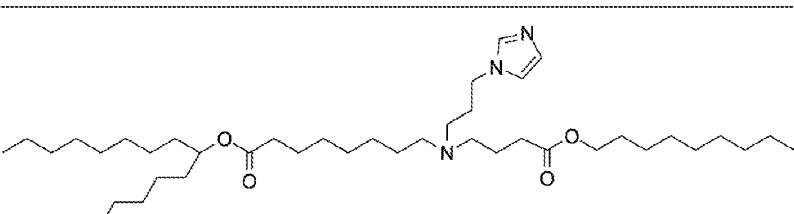
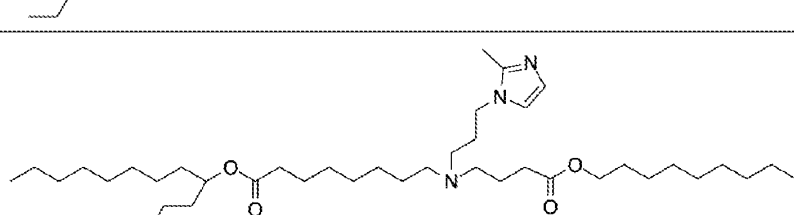
20	
21	
22	
23	
24	
25	
26	

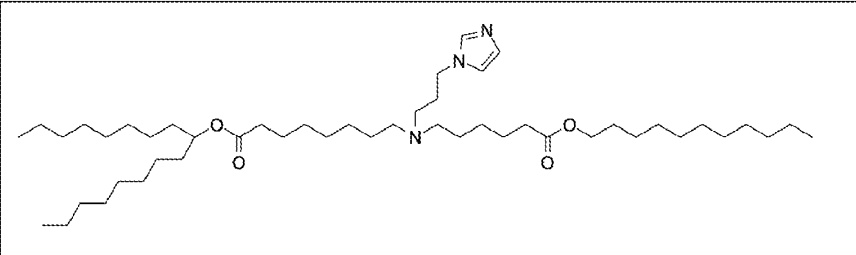
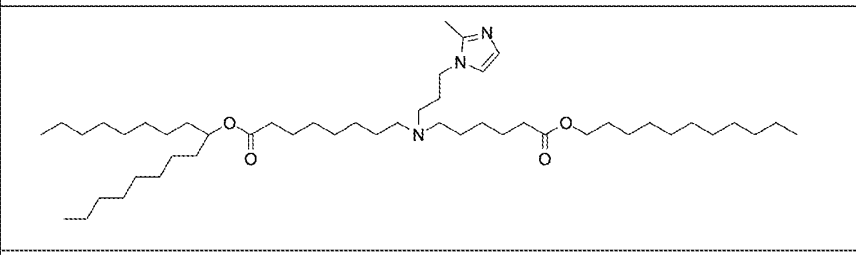
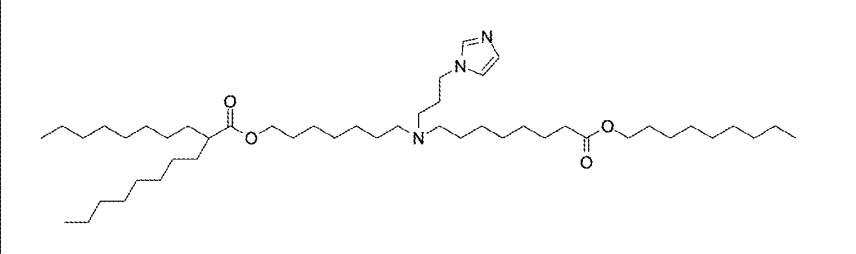
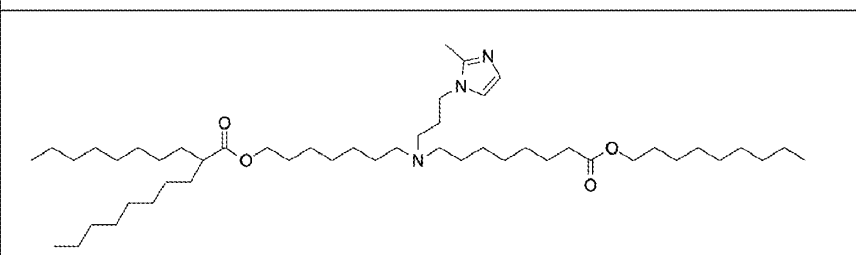
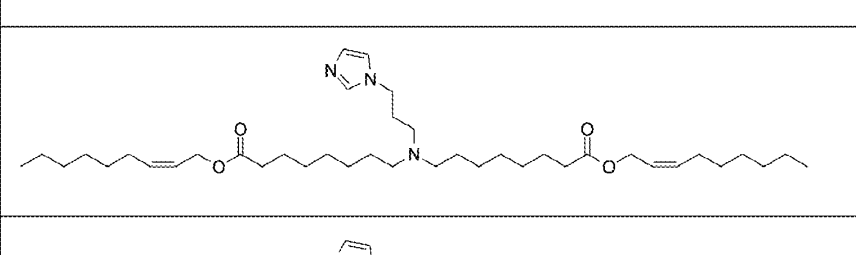
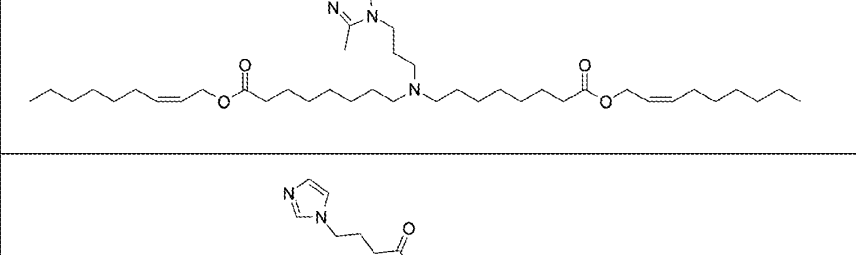
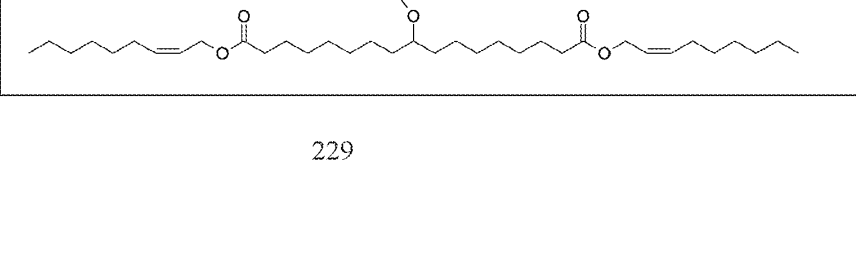
27	
28	
29	
30	

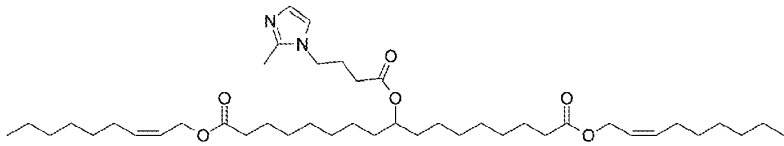
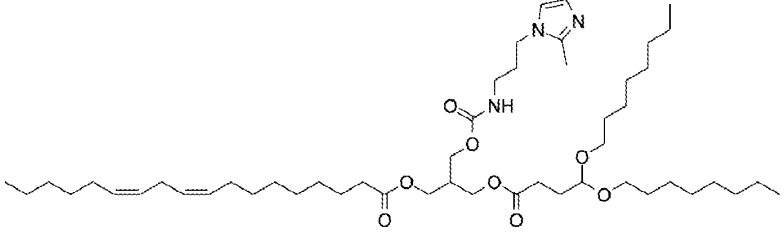
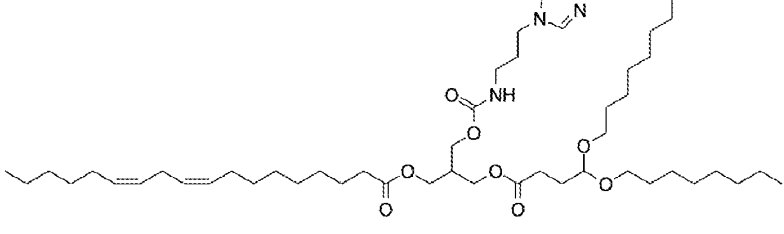
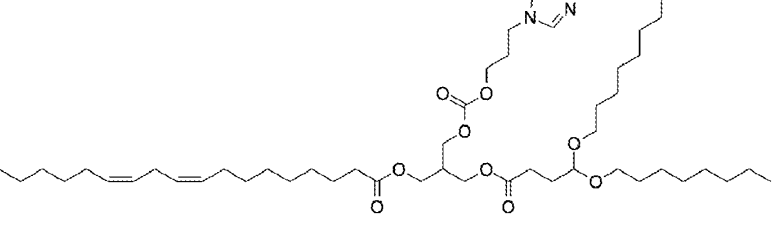
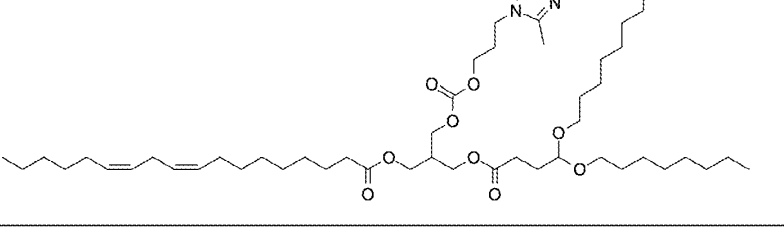
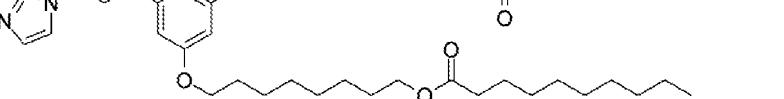
31	
32	
33	
34	

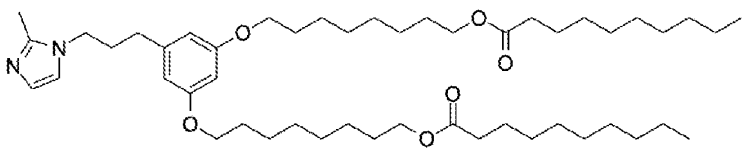
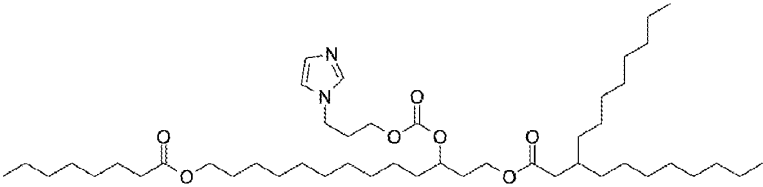
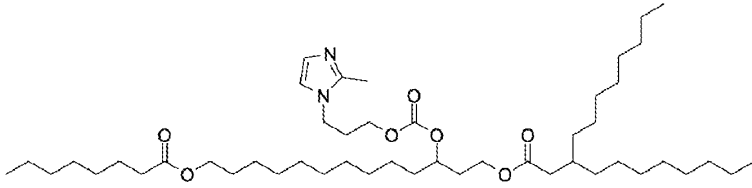
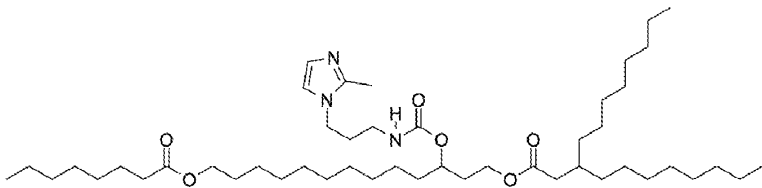
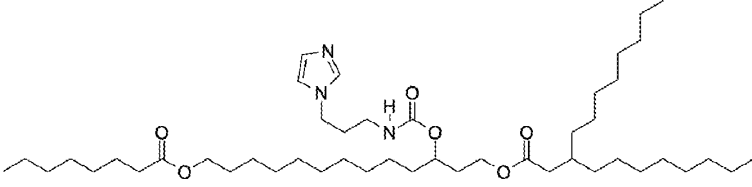
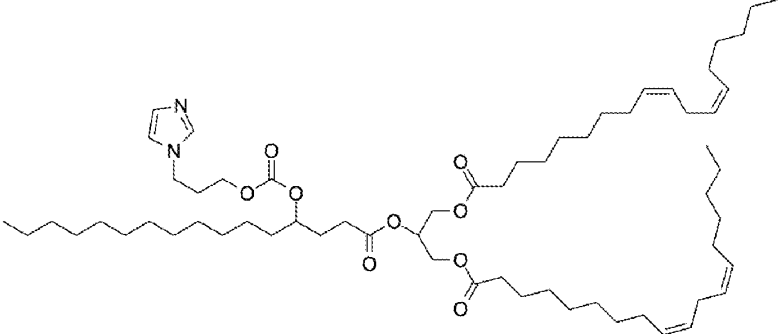
[illegible]

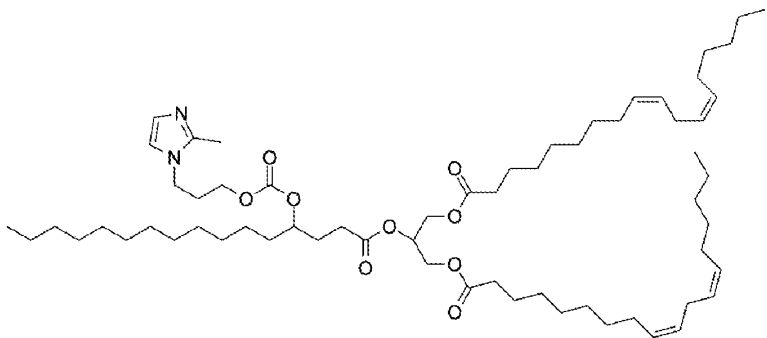
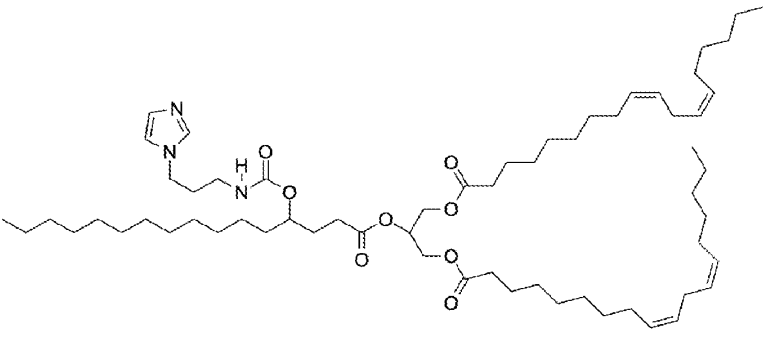
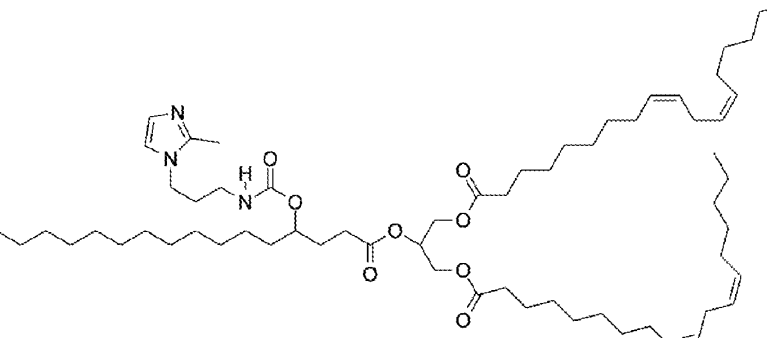
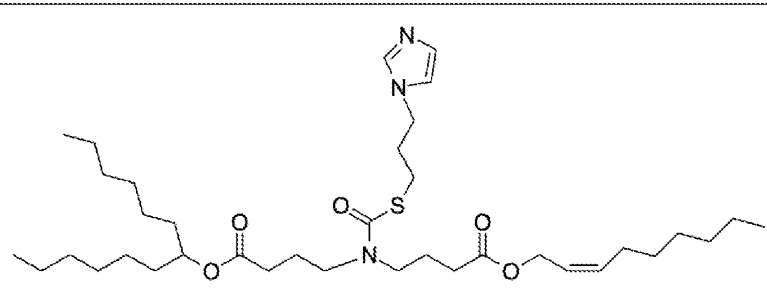
40	
41	
42	
43	
44	

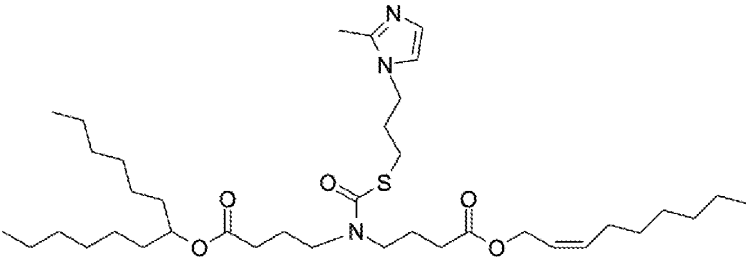
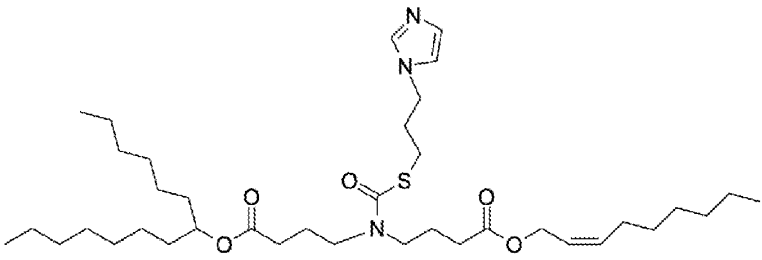
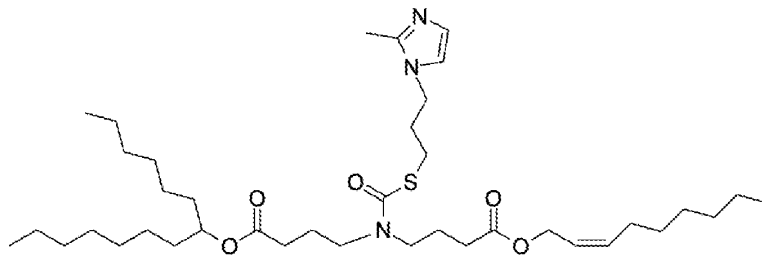
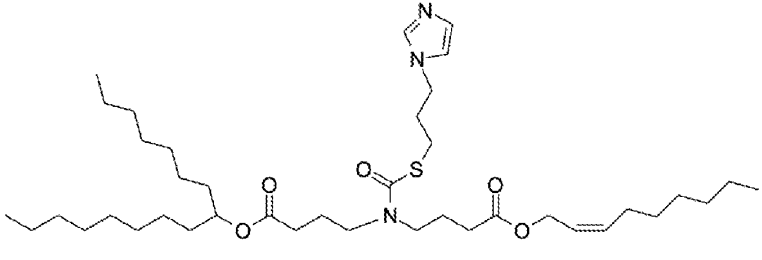
45	
46	
47	
48	
49	
50	

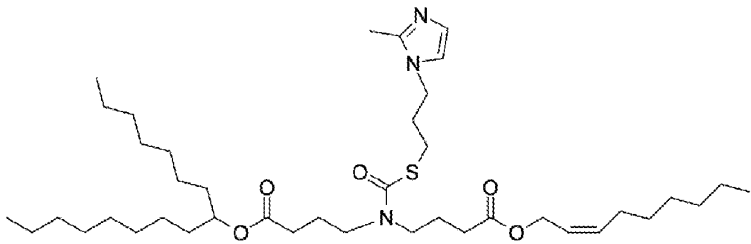
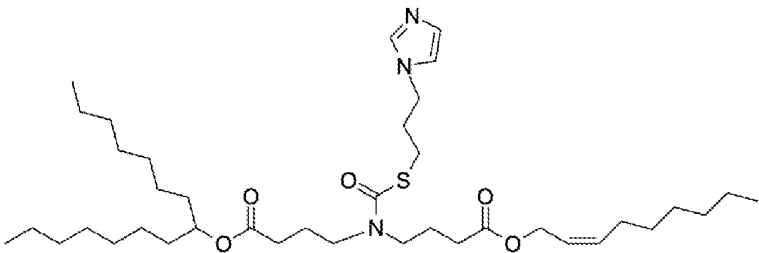
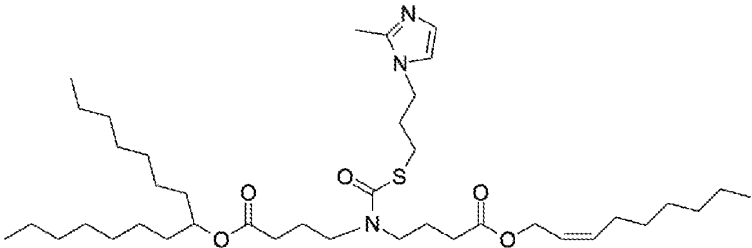
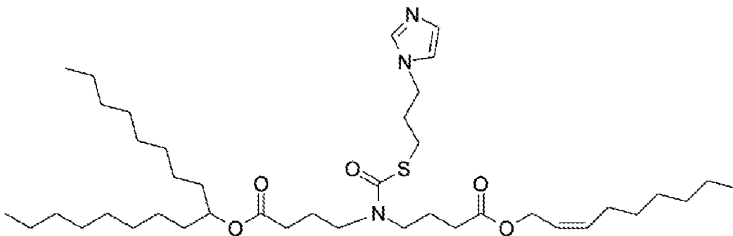
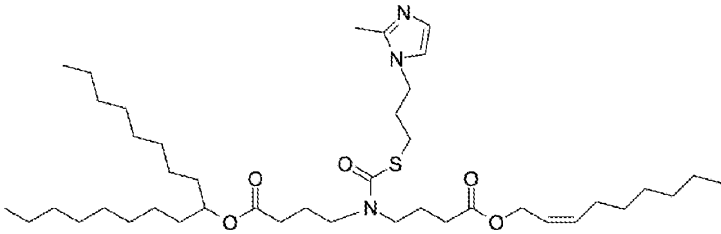
51	
52	
53	
54	
55	
56	
57	

58	
59	
60	
61	
62	
63	

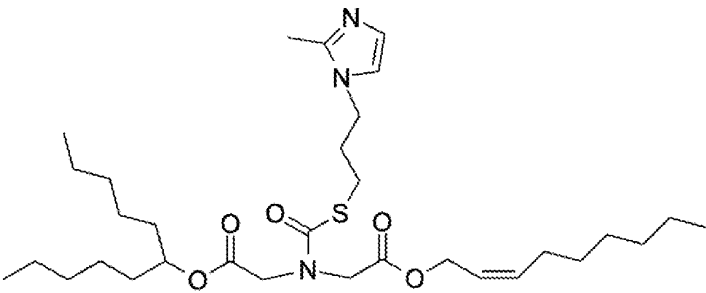
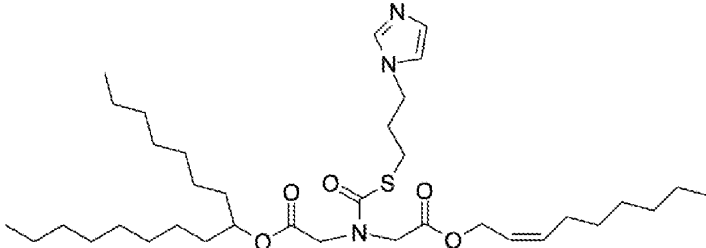
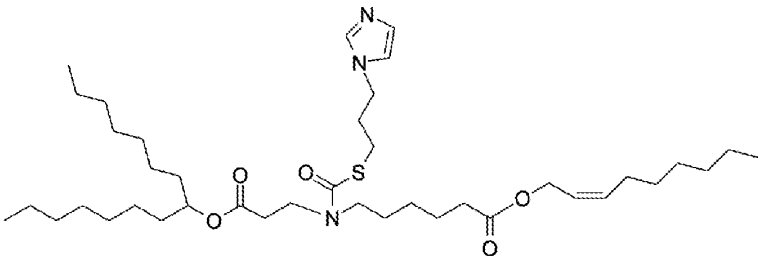
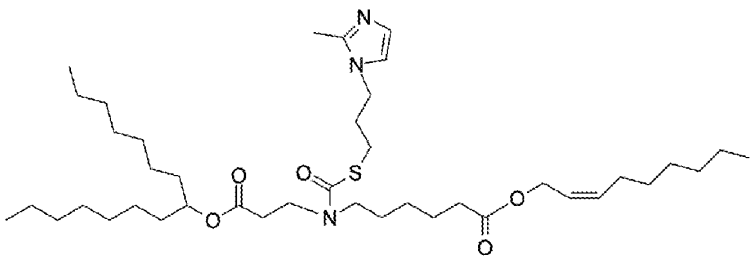
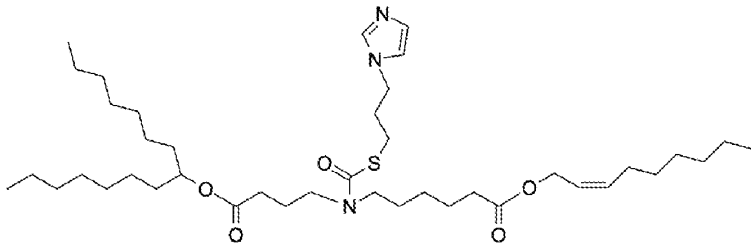
64	
65	
66	
67	
68	
69	

70	
71	
72	
73	

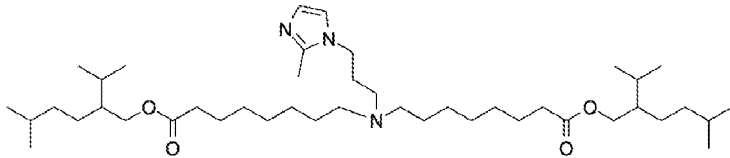
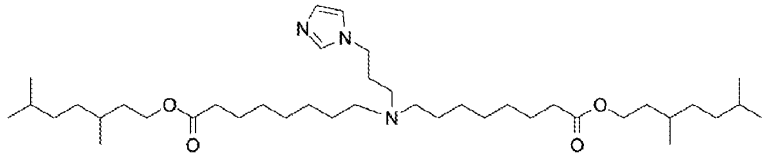
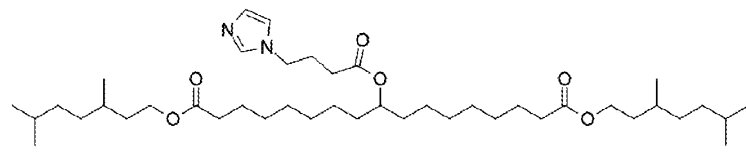
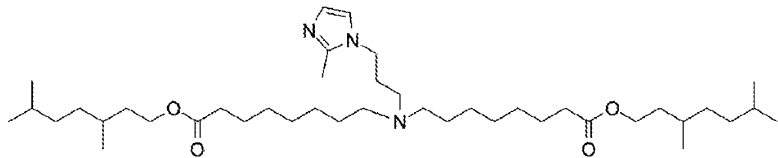
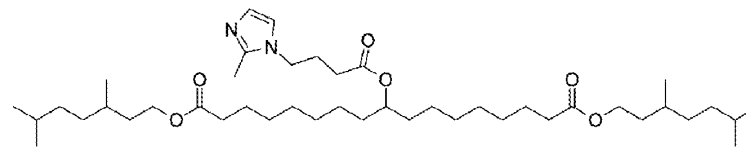
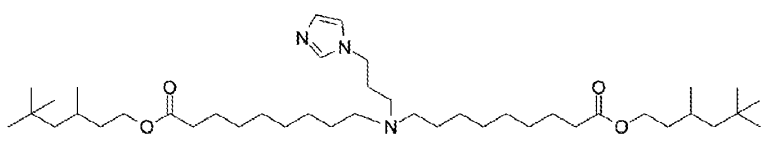
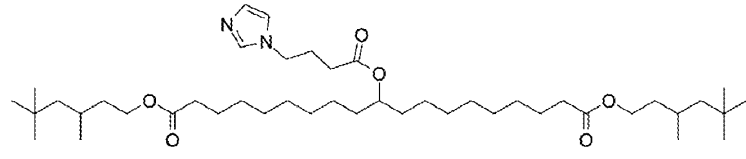
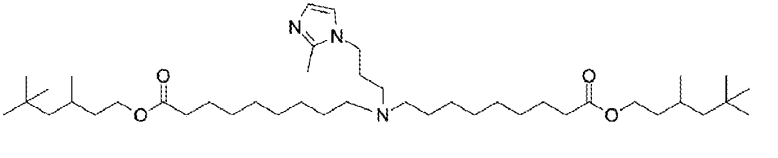
74	
75	
76	
77	

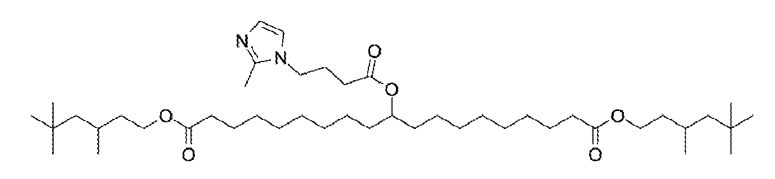
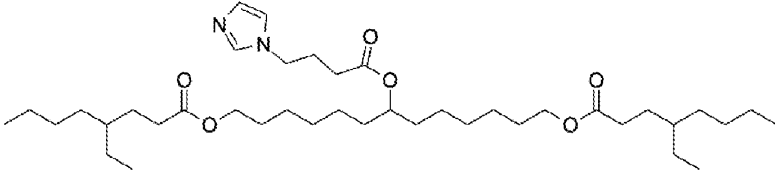
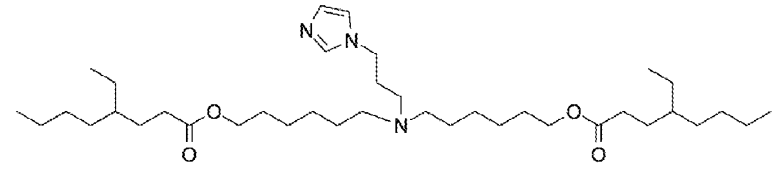
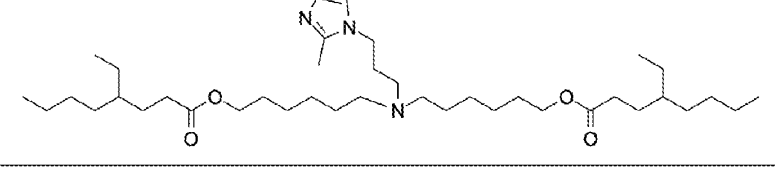
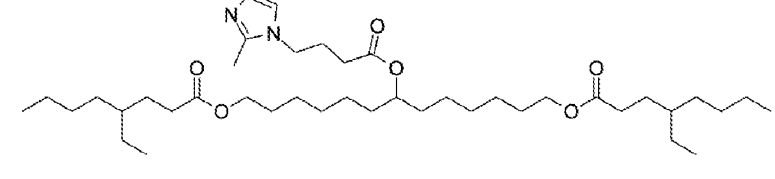
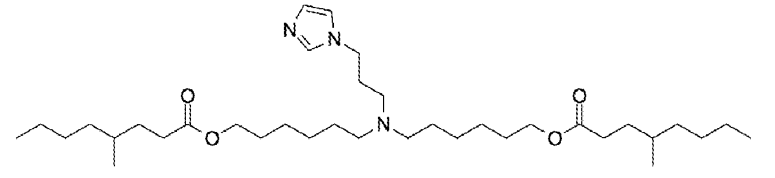
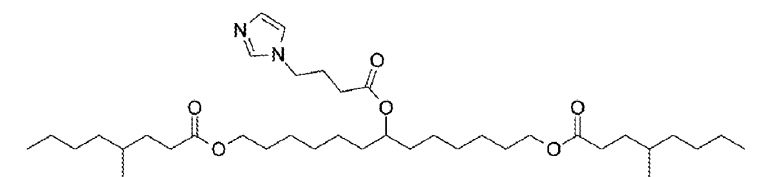
78	
79	
80	
81	
82	

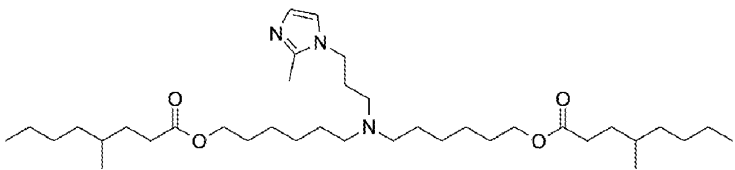
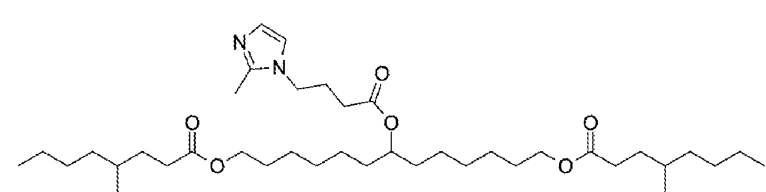
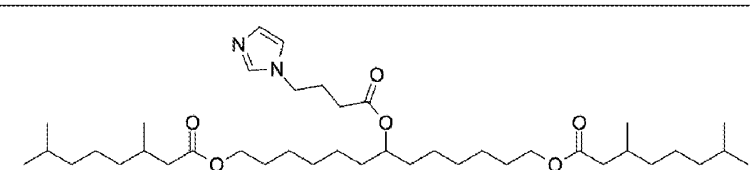
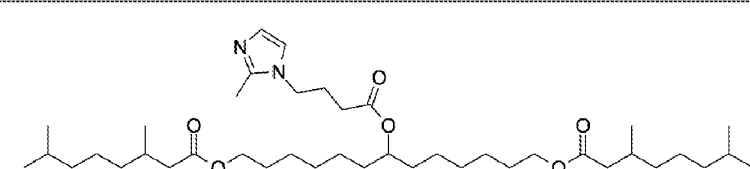
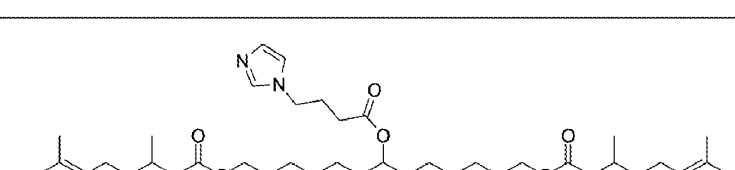
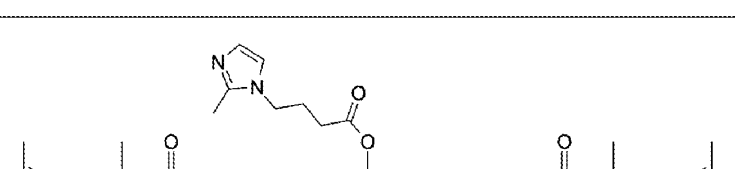
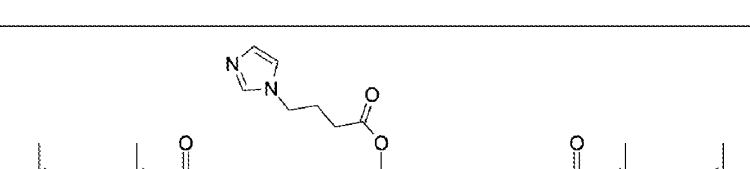
83	
84	
85	
86	
87	

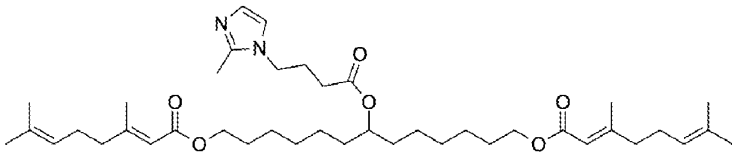
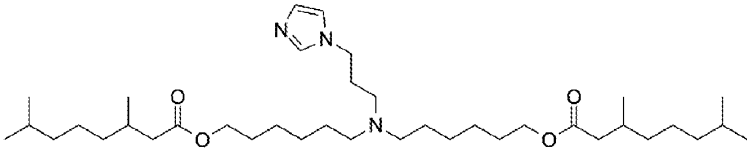
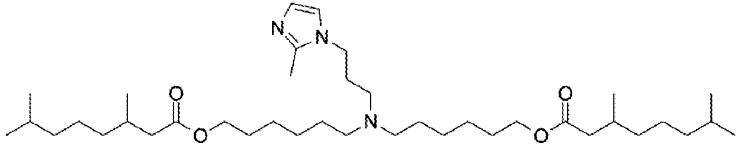
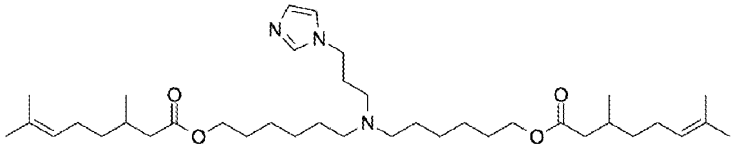
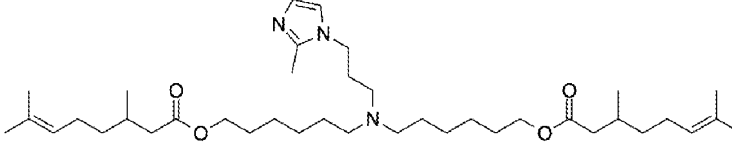
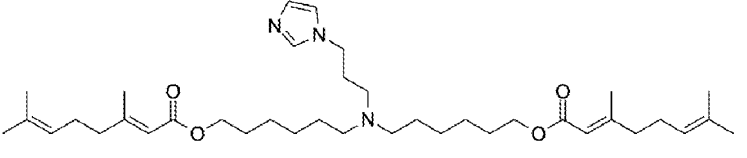
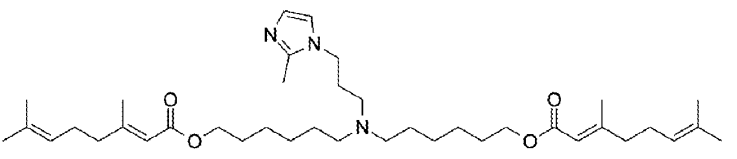
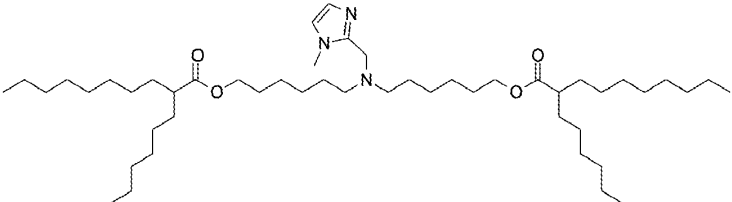
88	
89	
90	
91	
92	

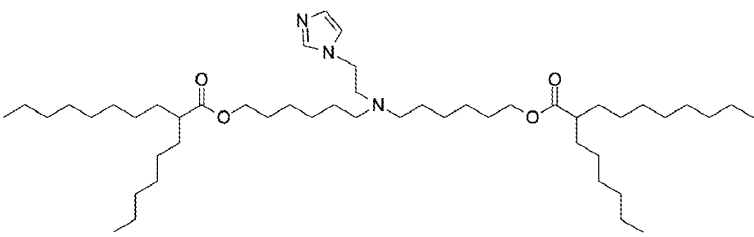
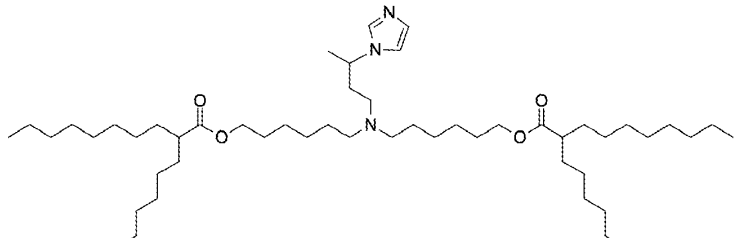
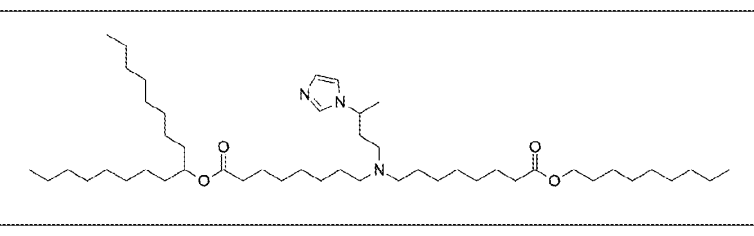
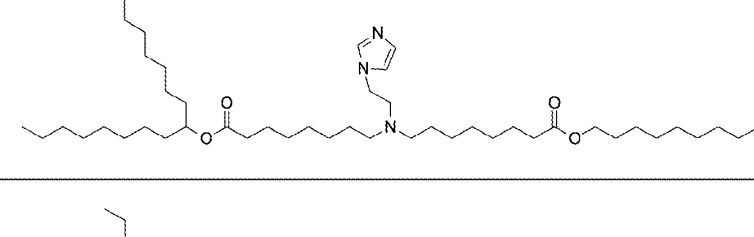
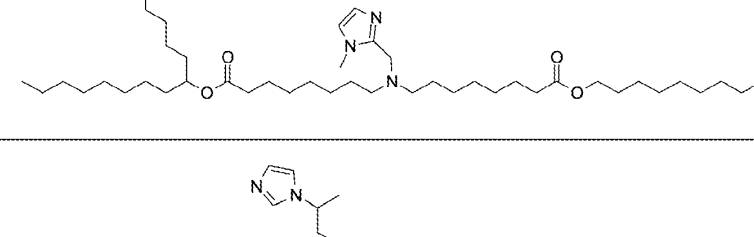
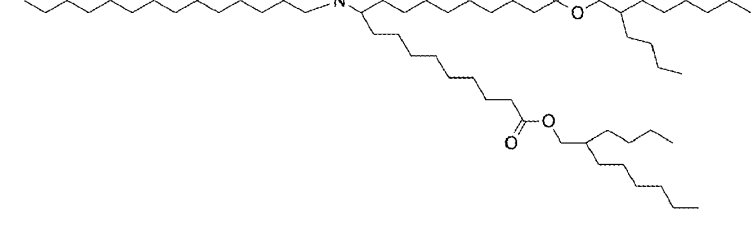
93	
94	
95	
96	
97	
98	

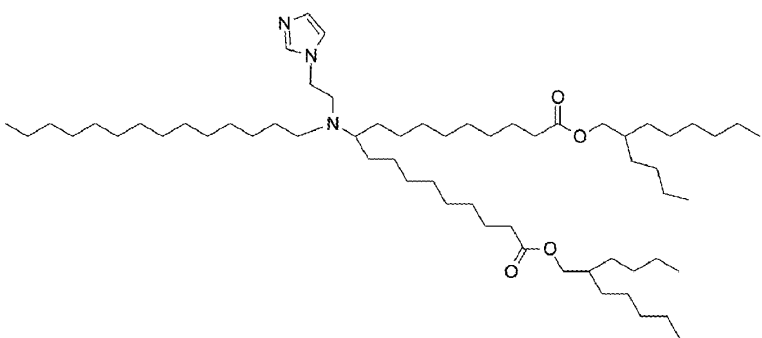
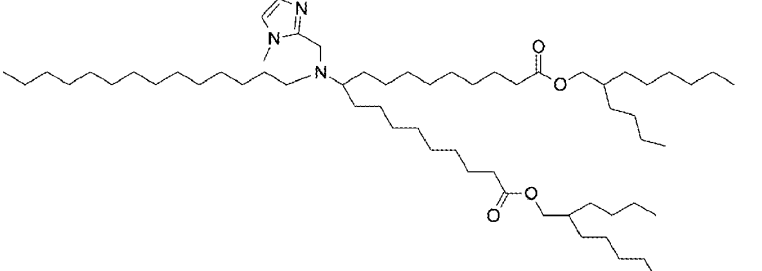
99	
100	
101	
102	
103	
104	
105	
106	

107	
108	
109	
110	
111	
112	
113	

114	
115	
116	
117	
118	
119	
120	

121	
122	
123	
124	
125	
126	
127	
128	

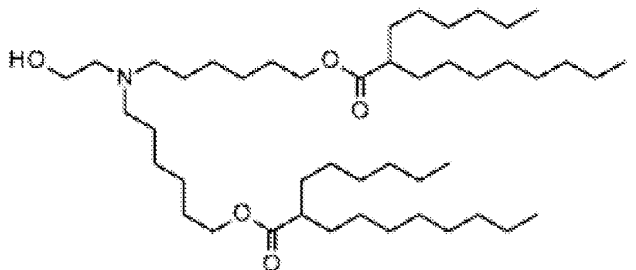
129	
130	
131	
132	
133	
134	

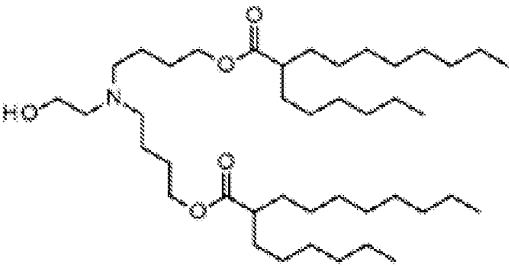
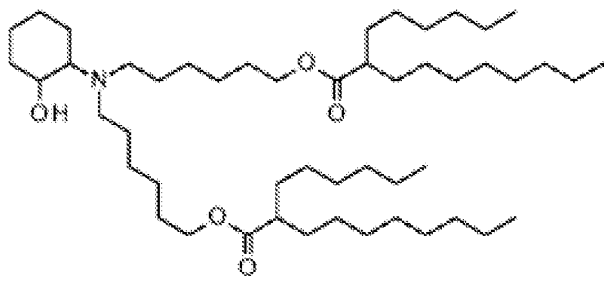
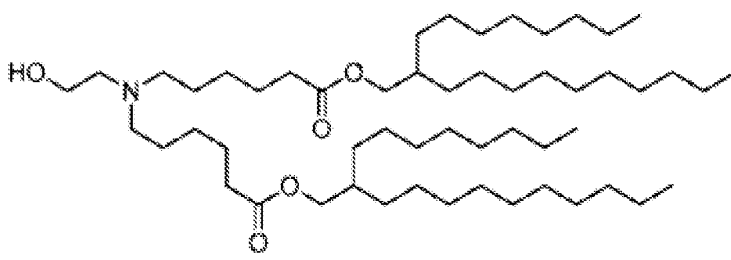
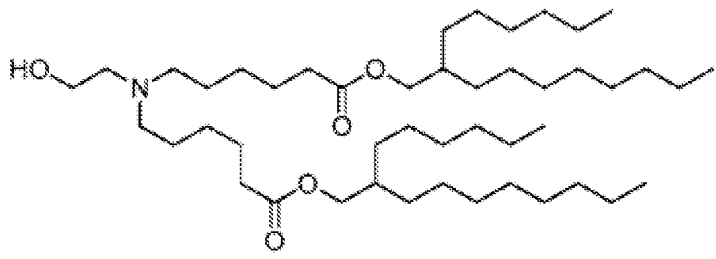
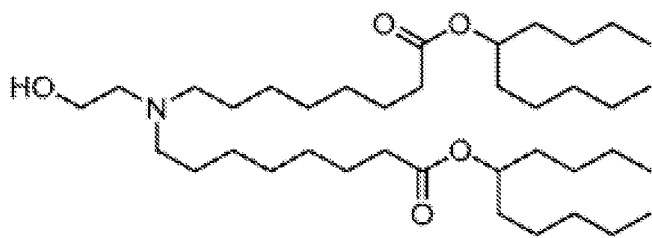
135	
136	

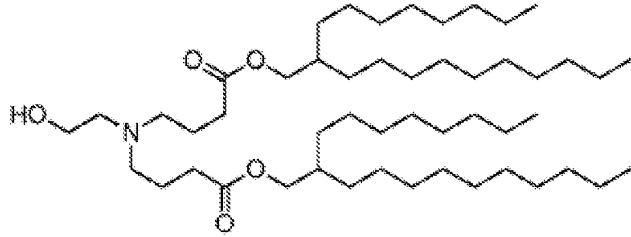
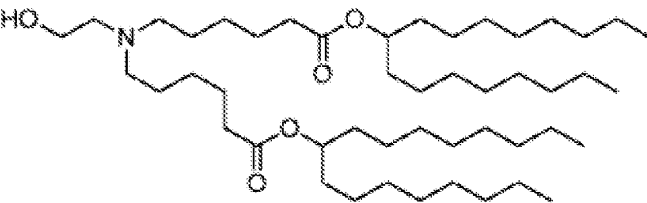
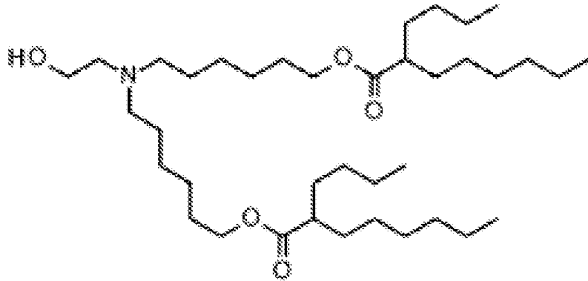
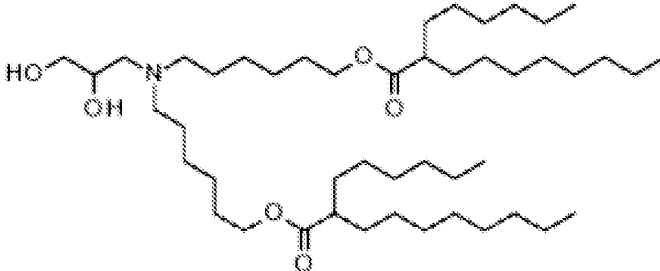
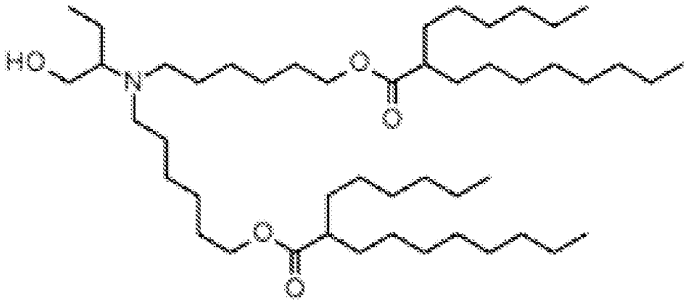
[0463] In some embodiments, the ionizable lipid has a beta-hydroxyl amine head group. In some embodiments, the ionizable lipid has a gamma-hydroxyl amine head group.

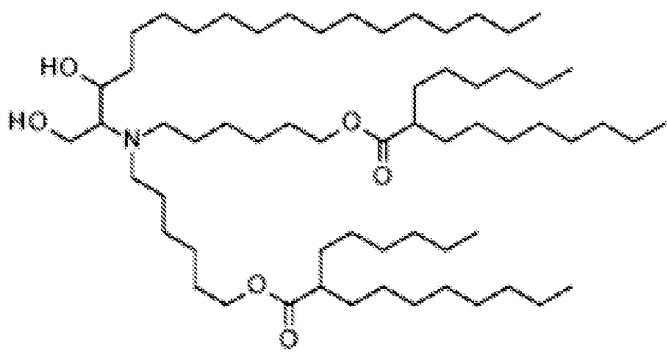
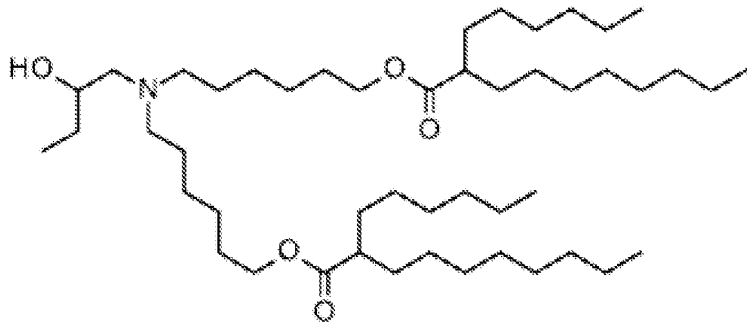
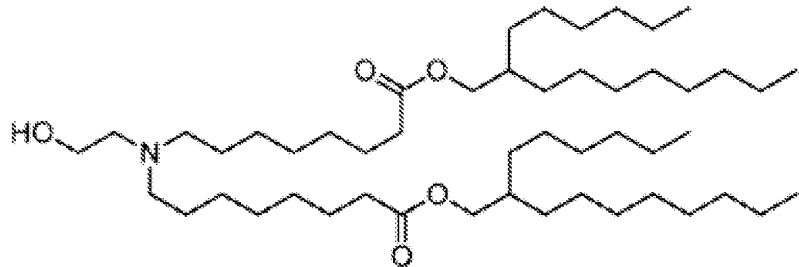
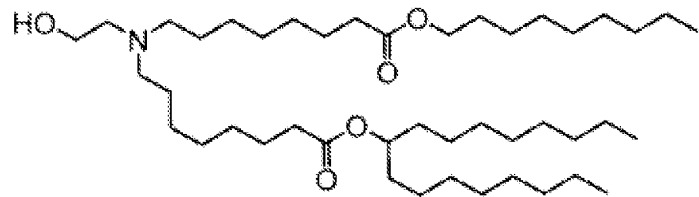
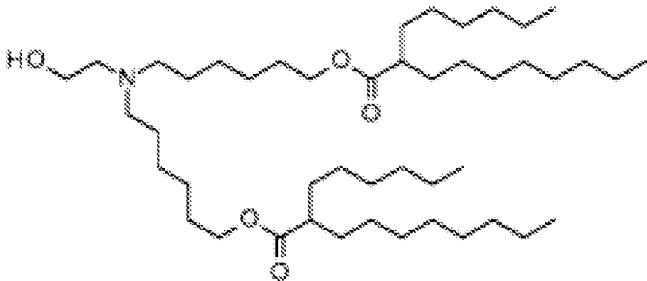
[0464] In some embodiments, an ionizable lipid of the disclosure is a lipid selected from Table 10b. In some embodiments, an ionizable lipid of the disclosure is Lipid 15 from Table 10b. In an embodiment, the ionizable lipid is described in US patent publication number US20170210697A1. In an embodiment, the ionizable lipid is described in US patent publication number US20170119904A1.

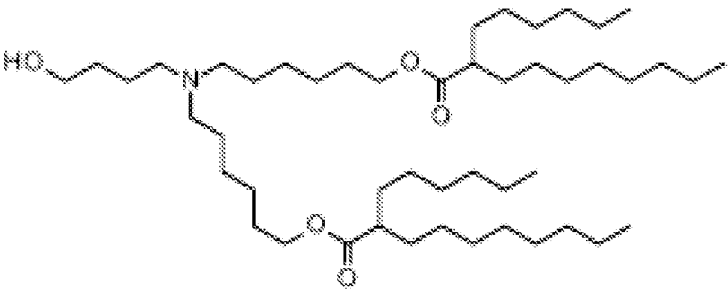
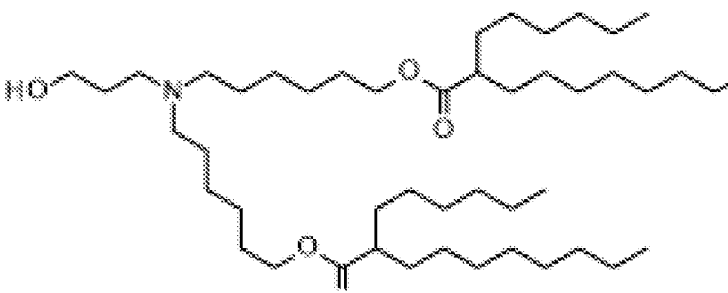
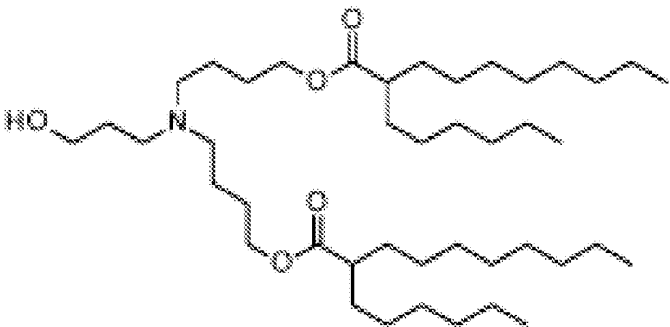
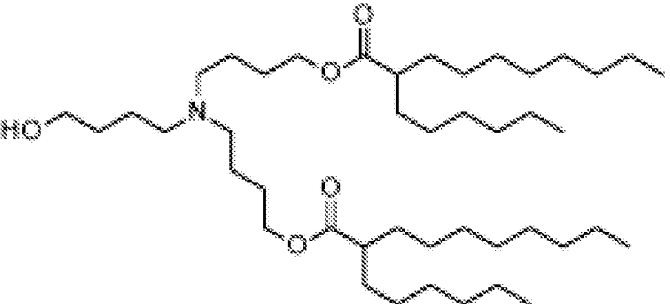
Table 10b

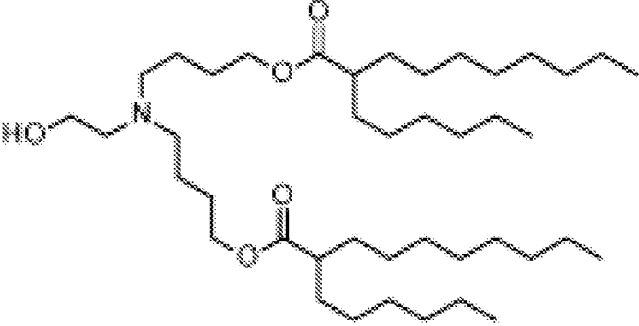
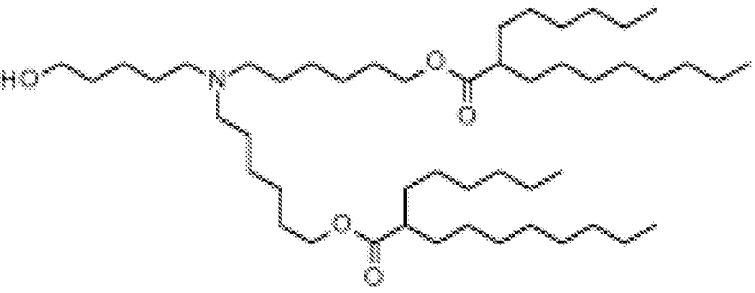
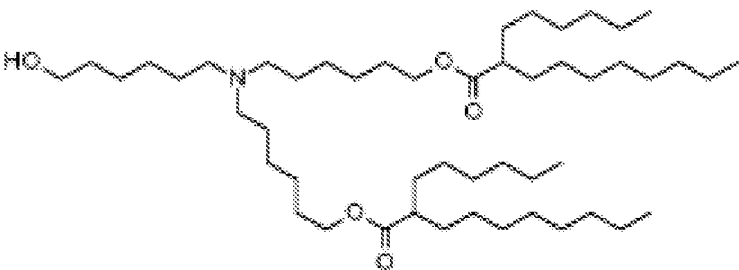
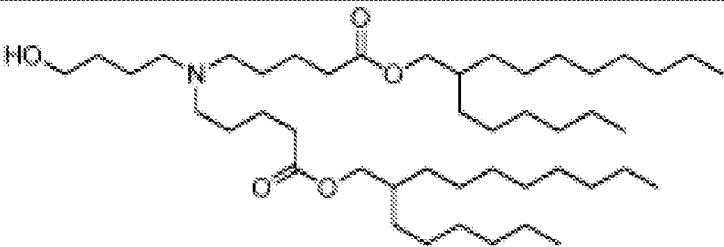
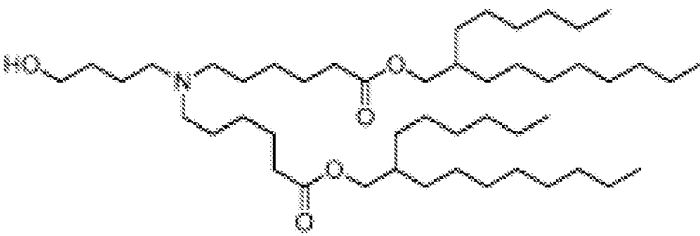
Ionizable lipid number	Structure
1	

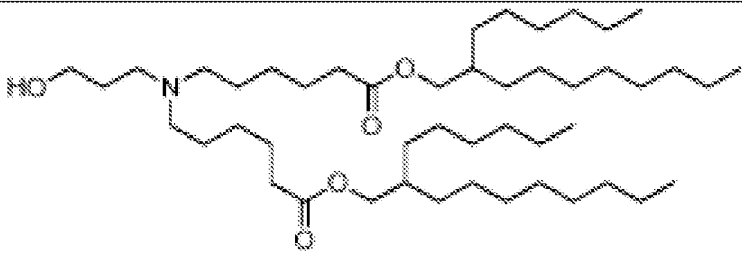
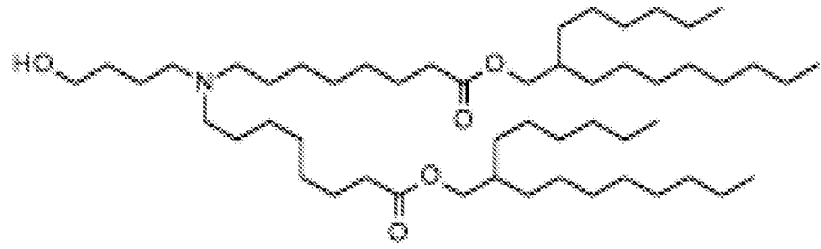
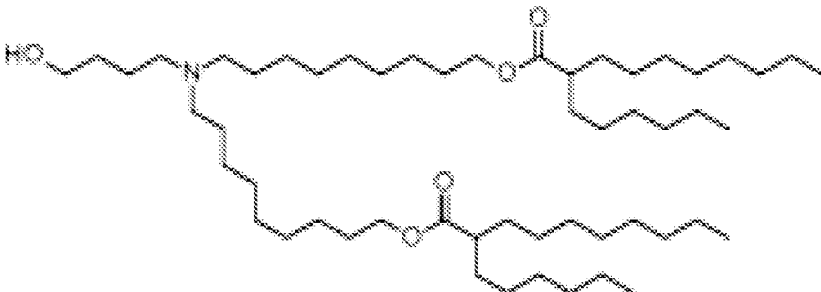
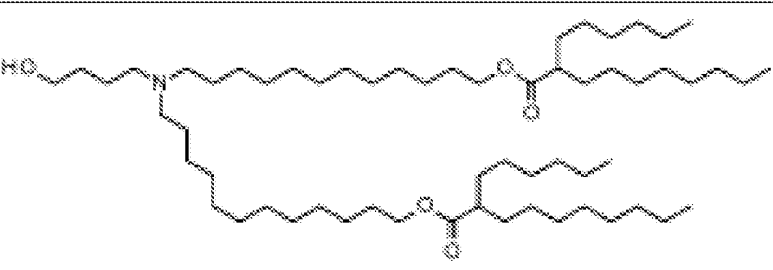
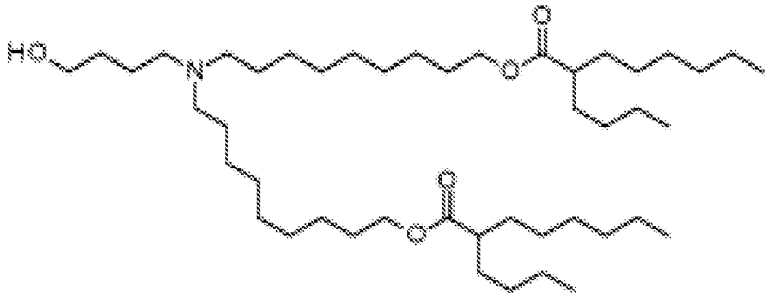
2	
3	
4	
5	
6	

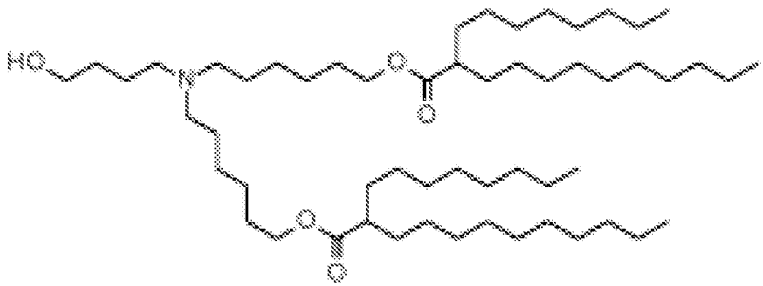
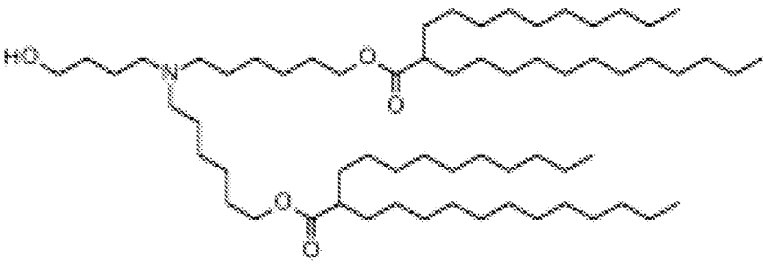
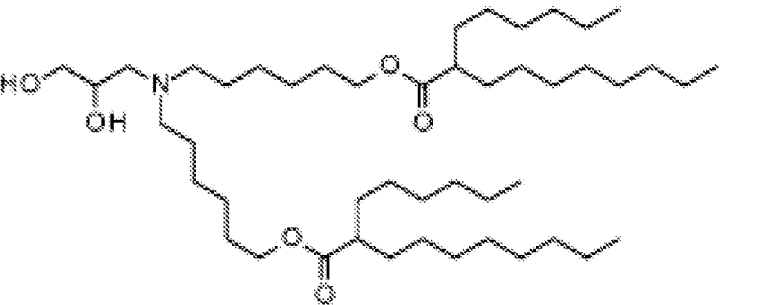
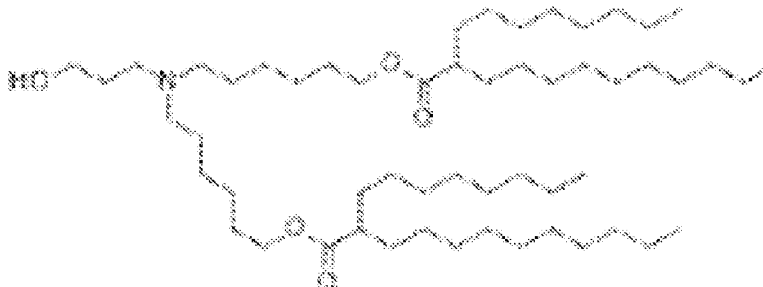
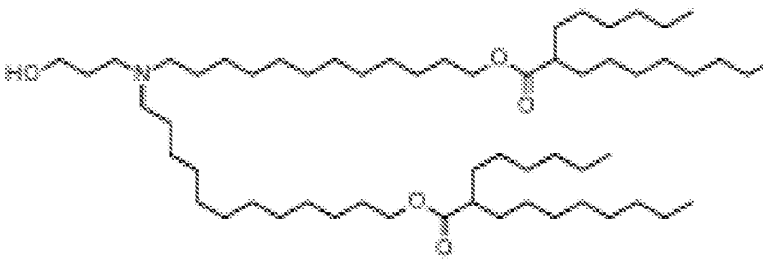
7	
8	
9	
10	
11	

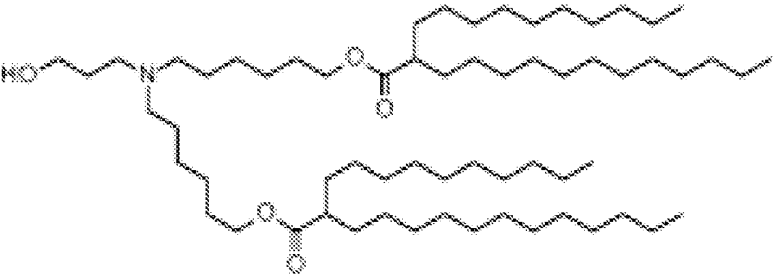
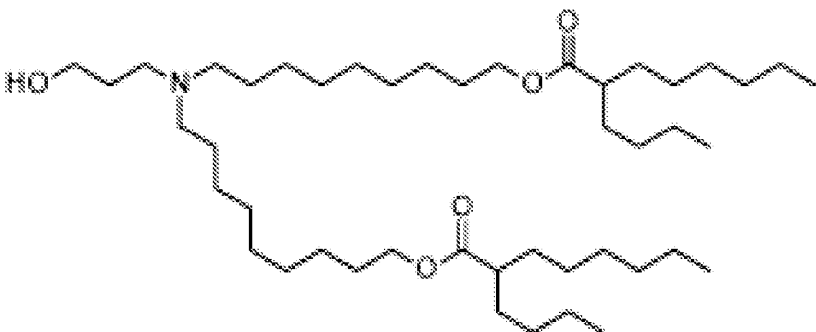
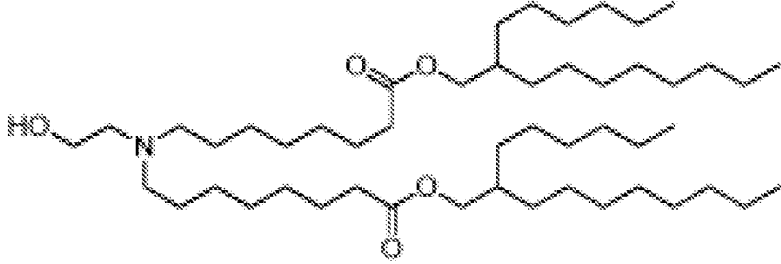
12	
13	
14	
15	
16	

17	
18	
19	
20	

21	
22	
23	
24	
25	

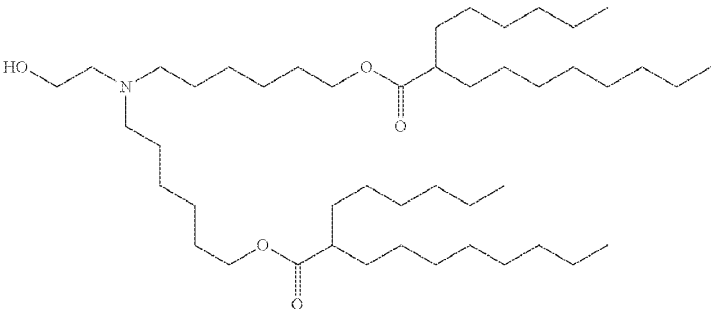
26	
27	
28	
29	
30	

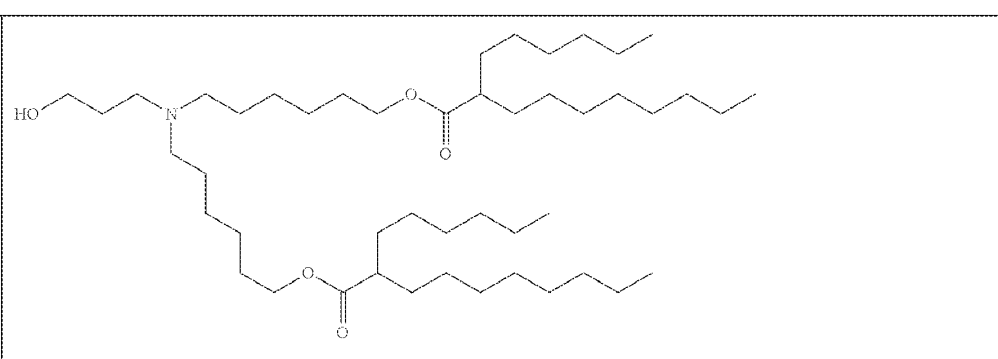
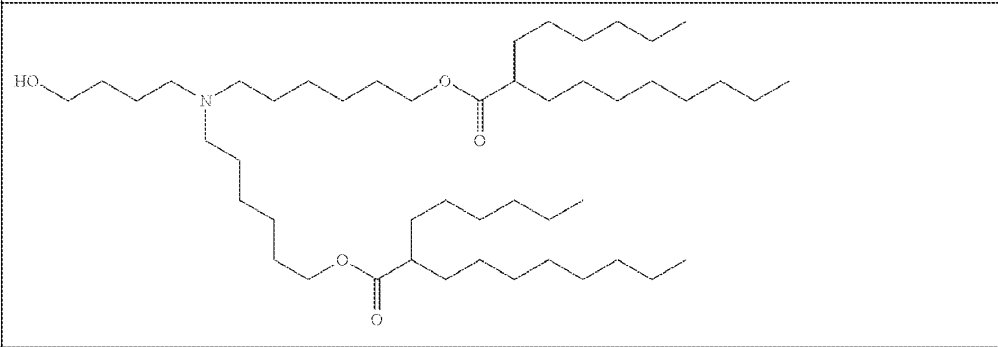
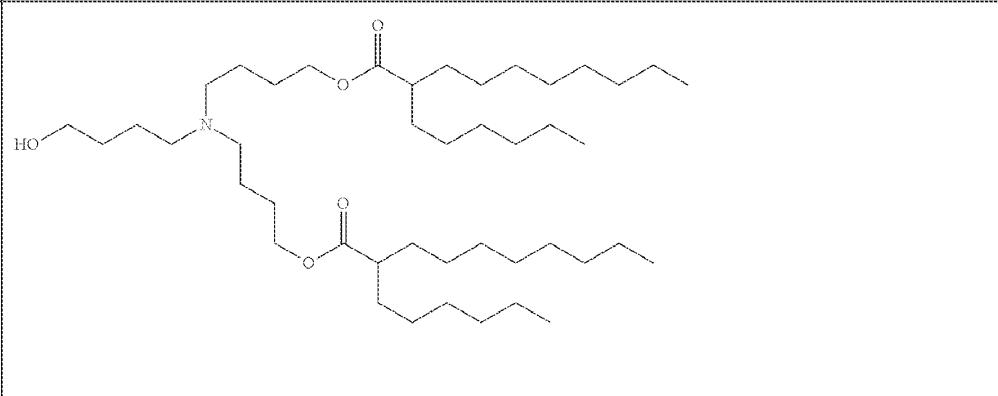
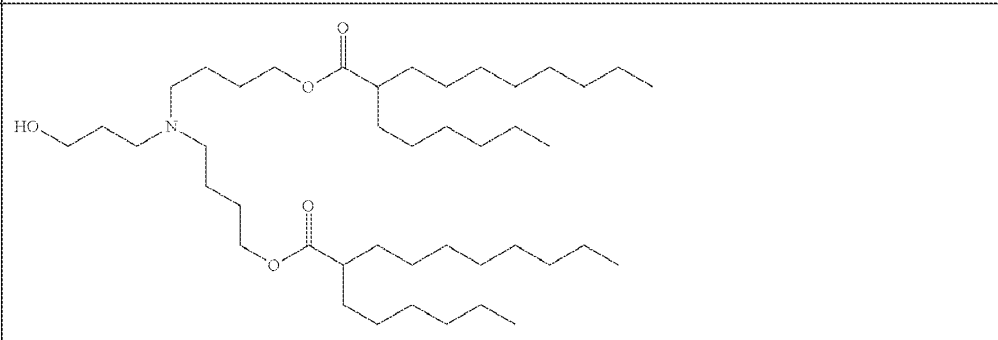
31	
32	
33	
34	
35	

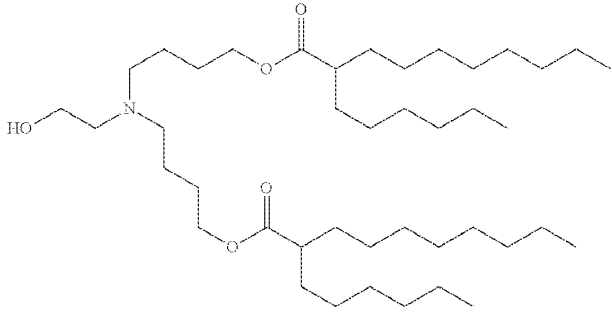
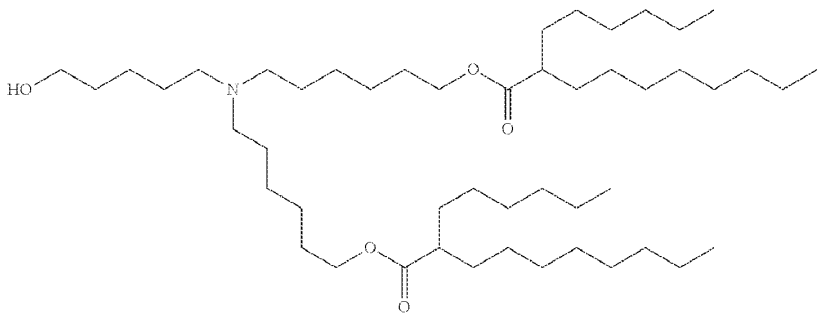
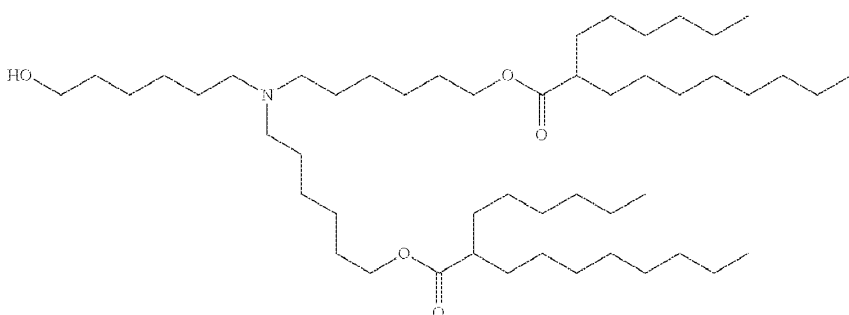
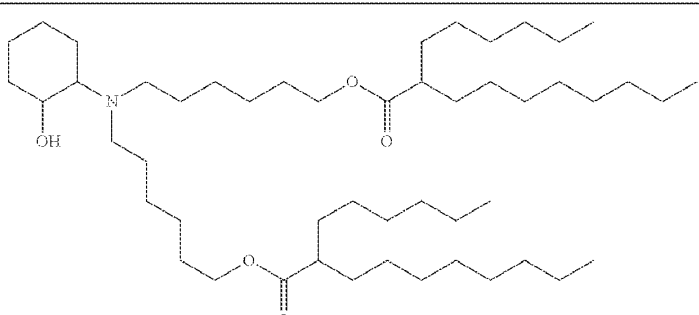
36	
37	
38	

[0465] In some embodiments, an ionizable lipid has one of the structures set forth in Table 10 below.

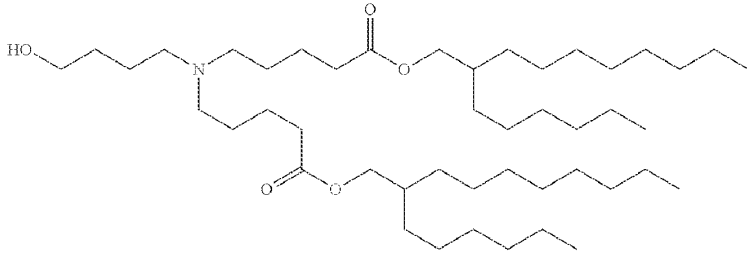
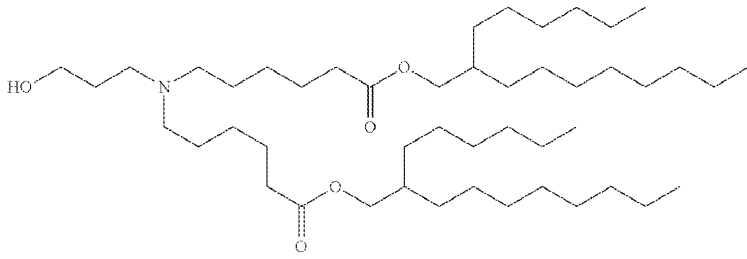
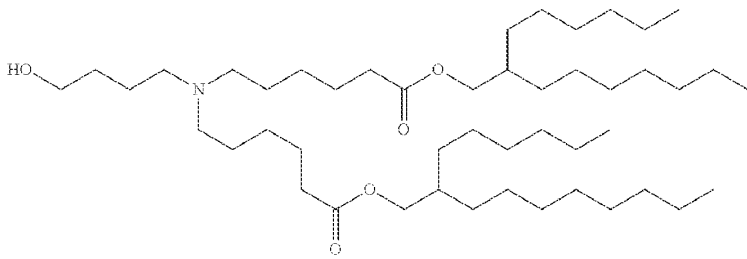
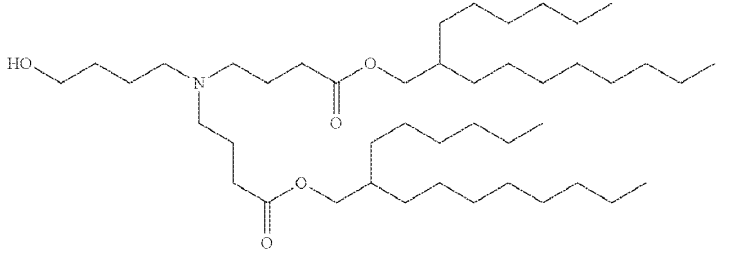
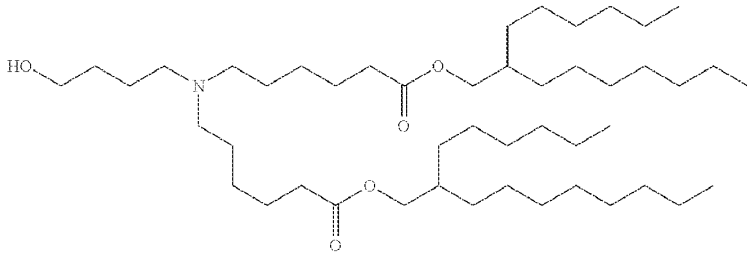
Table 10

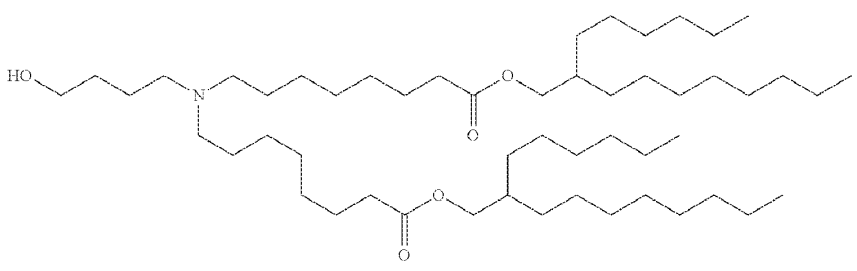
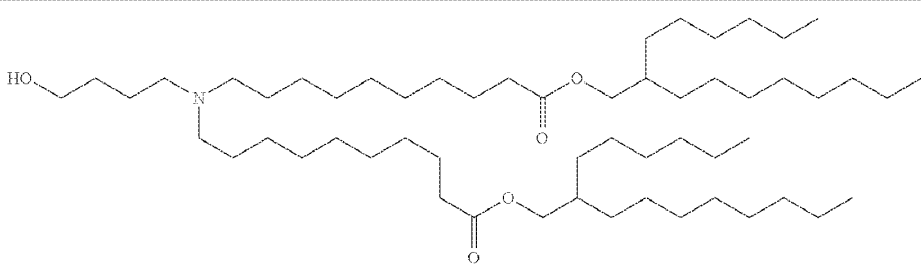
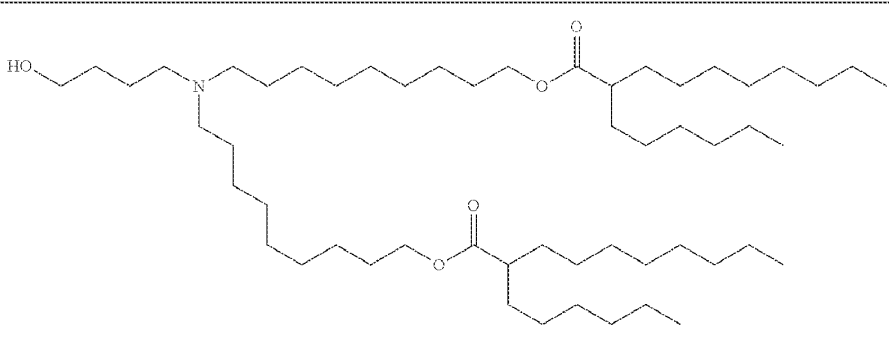
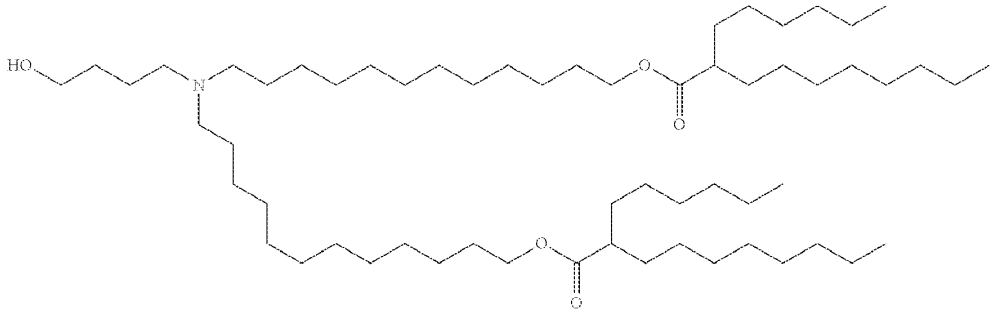
Number	Structure
1	

2	 <chem>CCCC(O)CN(CCCCCC(=O)OCC(C)CCCC)CCCCC(=O)OCC(C)CCCC</chem>
3	 <chem>CCCC(O)CN(CCCCCC(=O)OCC(C)CCCC)CCCCC(=O)OCC(C)CCCC</chem>
4	 <chem>CCCC(O)CN(CCCCCC(=O)OCC(C)CCCC)CCCCC(=O)OCC(C)CCCC</chem>
5	 <chem>CCCC(O)CN(CCCCCC(=O)OCC(C)CCCC)CCCCC(=O)OCC(C)CCCC</chem>

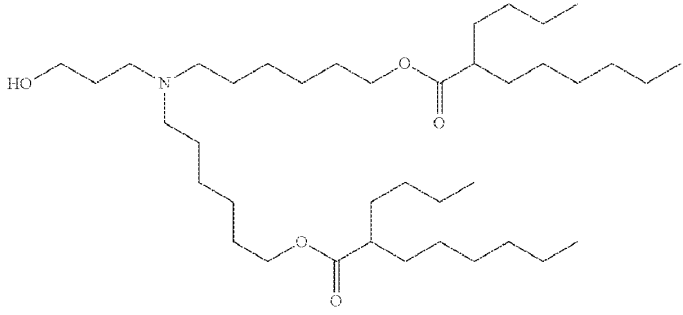
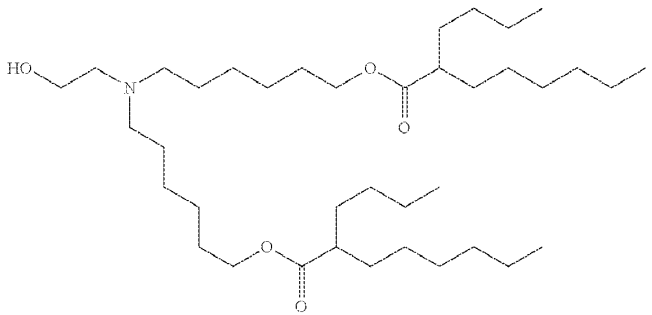
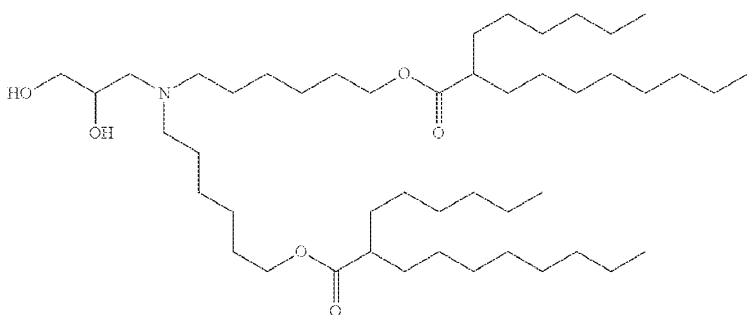
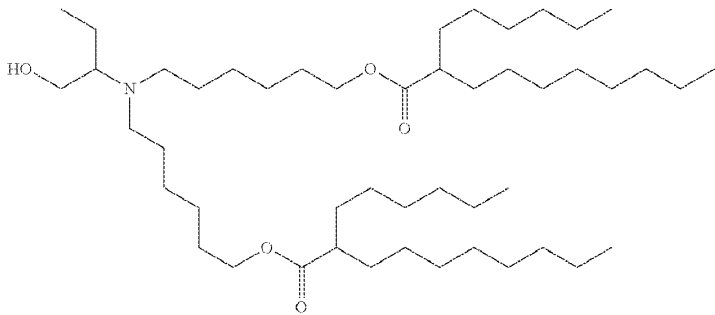
6	
7	
8	
9	

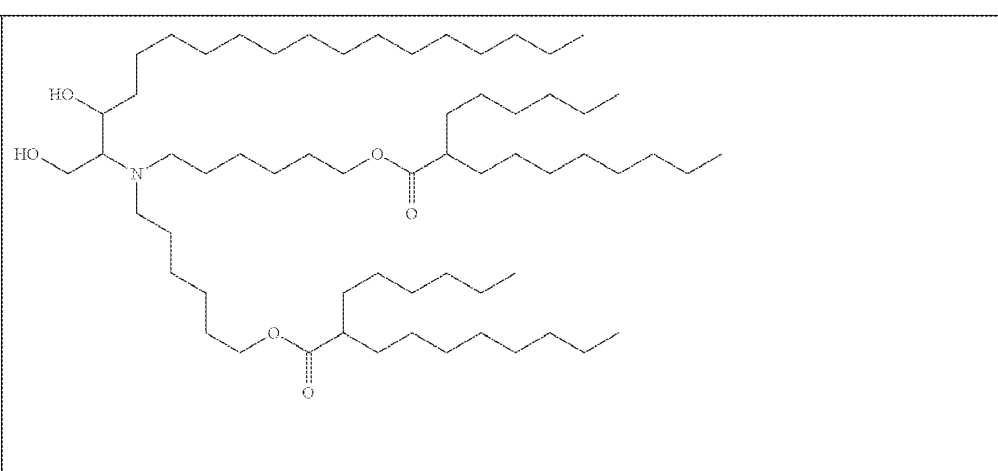
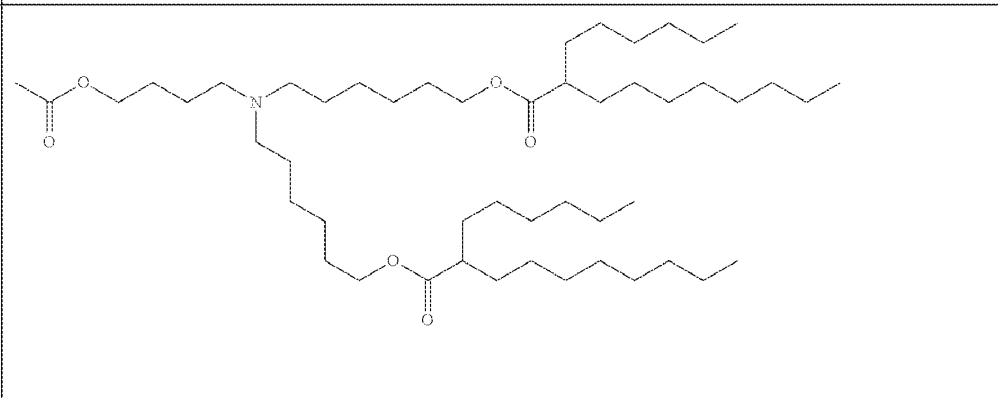
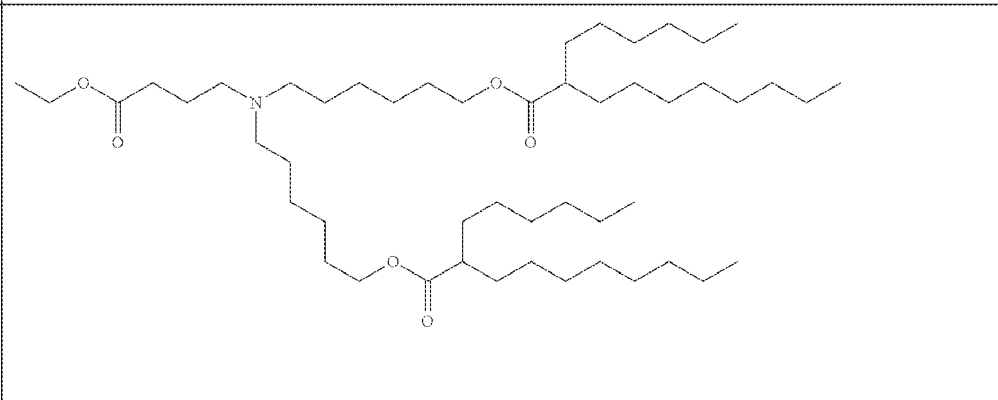
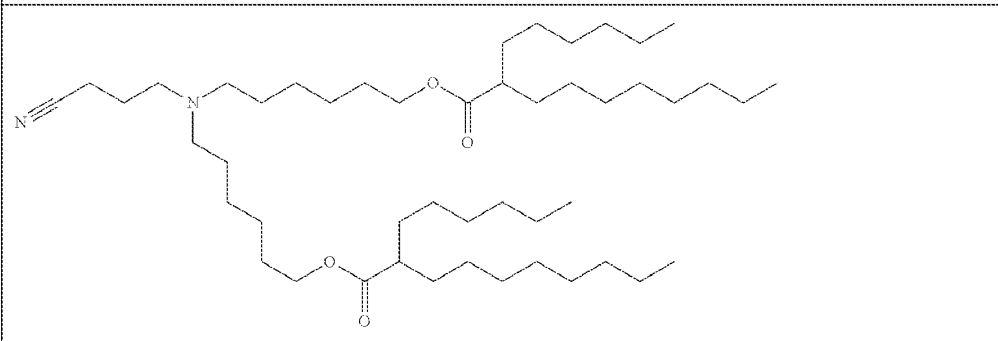
10	<chem>CCCCCCCC(=O)OCC(CCCCCC)N1CC(CCCOCC(=O)OCCCCC)CC1</chem>
11	<chem>CCCCCCCC(=O)OCC(CCCCCC)N1CC(CCCOCC(=O)OCCCCC)CC1</chem>
12	<chem>CCCCCCCC(=O)OCC(CCCCCC)N1CC(CCCOCC(=O)OCCCCC)CC1</chem>
13	<chem>CCCCCCCC(=O)OCC(CCCCCC)N1CC(CCCOCC(=O)OCCCCC)CC1</chem>
14	<chem>CCCCCCCC(=O)OCC(CCCCCC)N1CC(CCCOCC(=O)OCCCCC)CC1</chem>

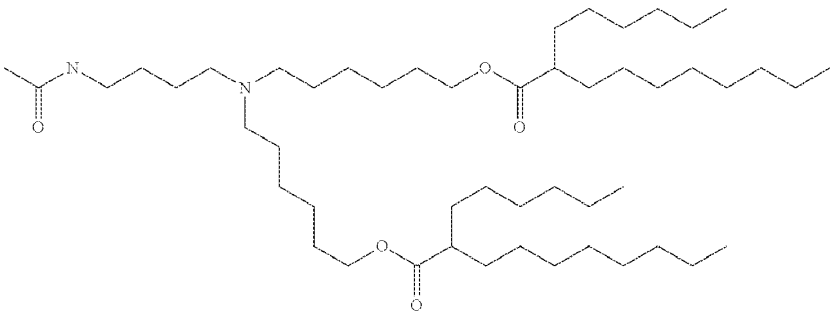
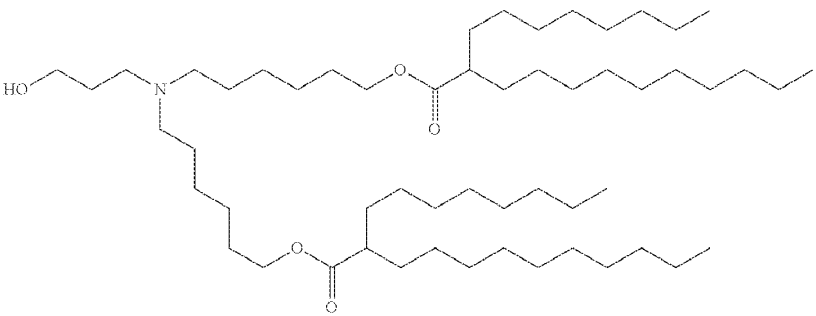
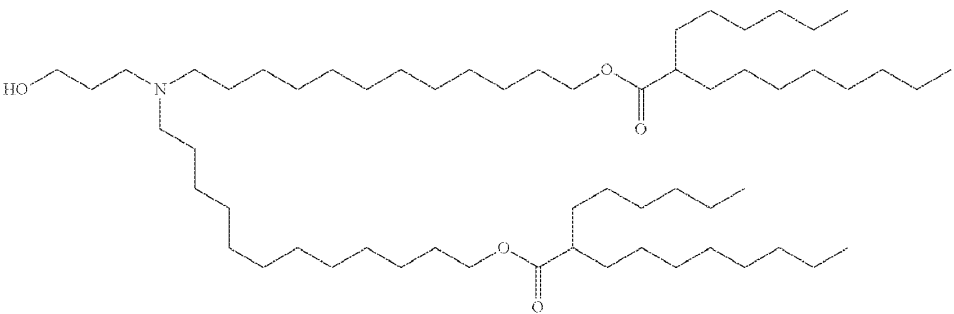
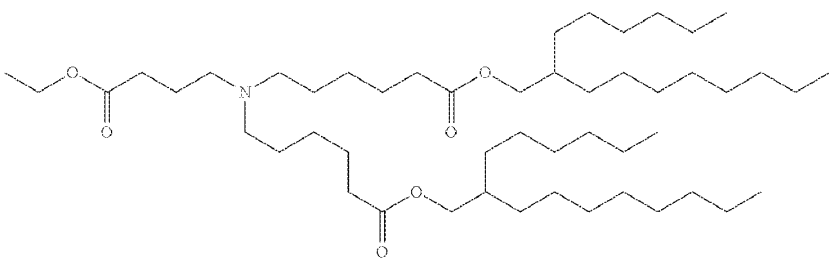
15	
16	
17	
18	
19	

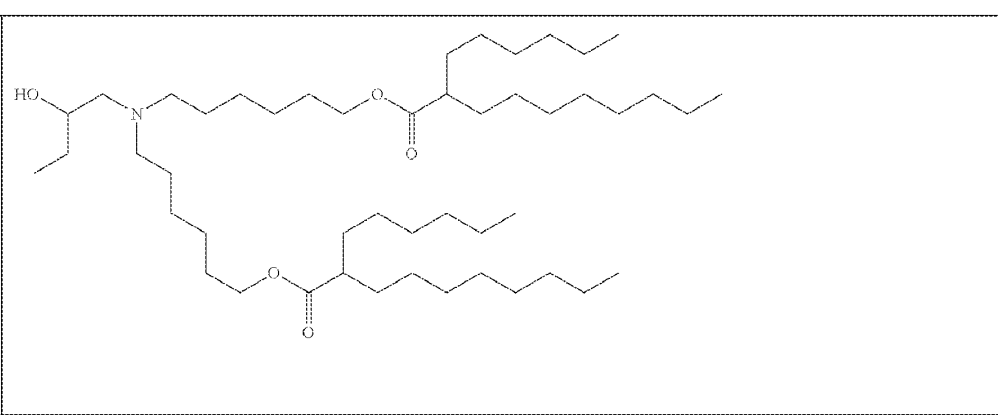
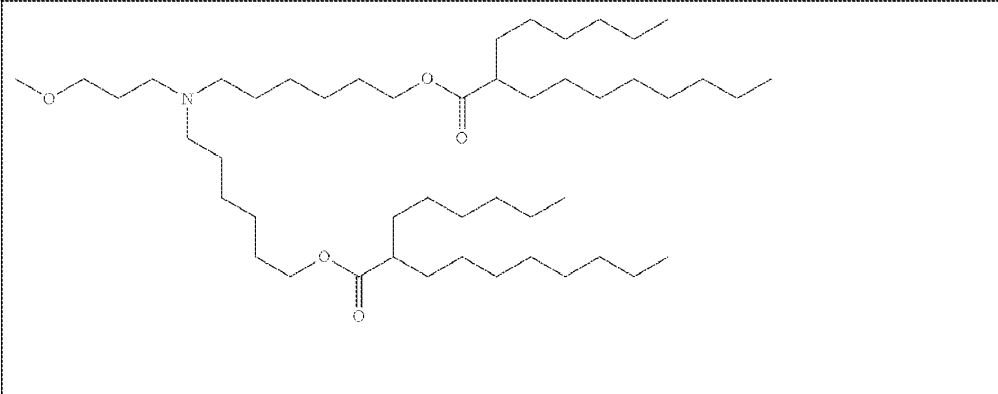
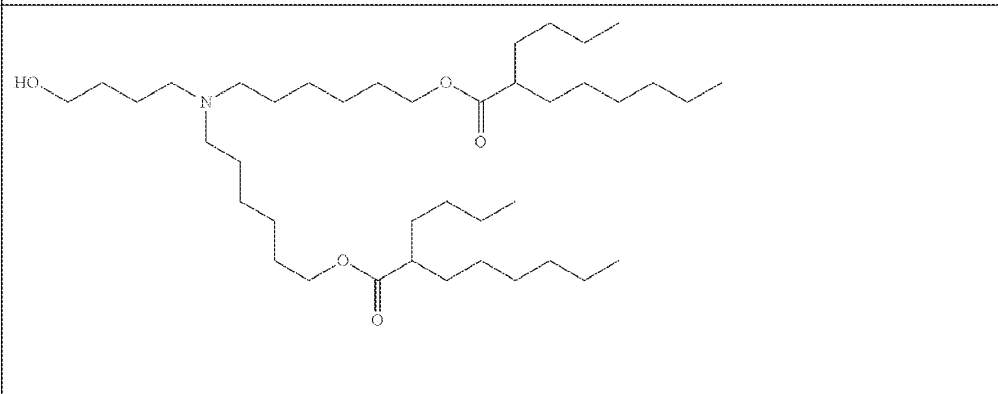
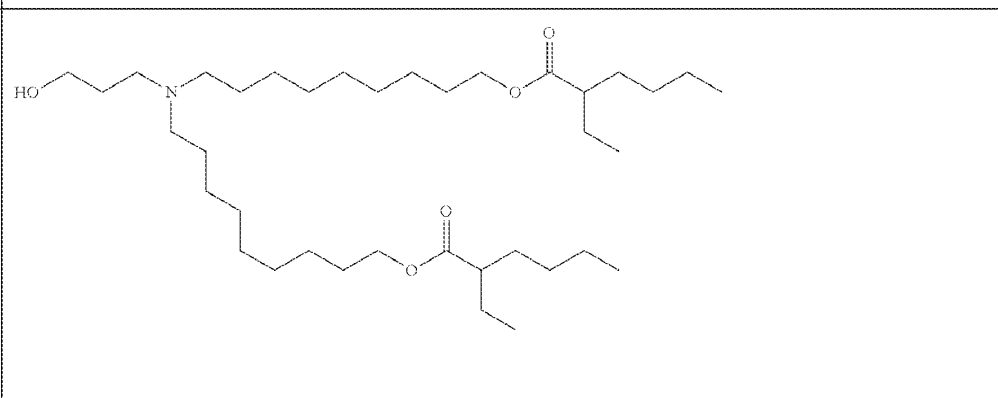
20	
21	
22	
23	

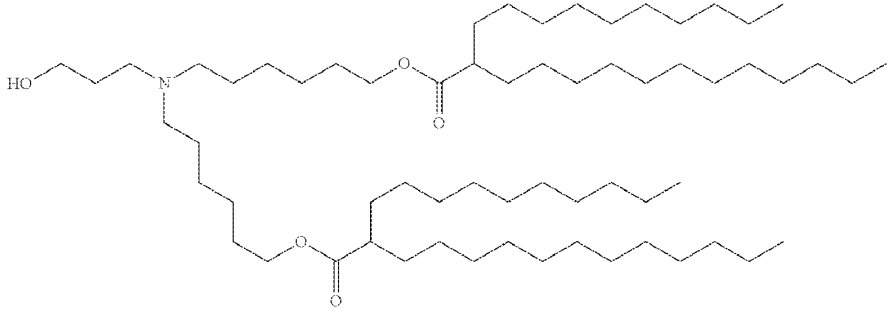
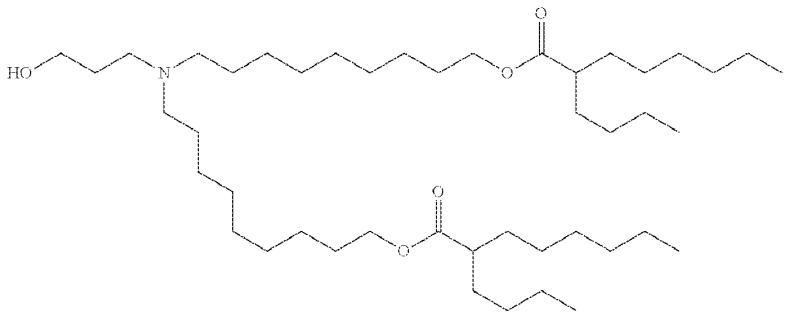
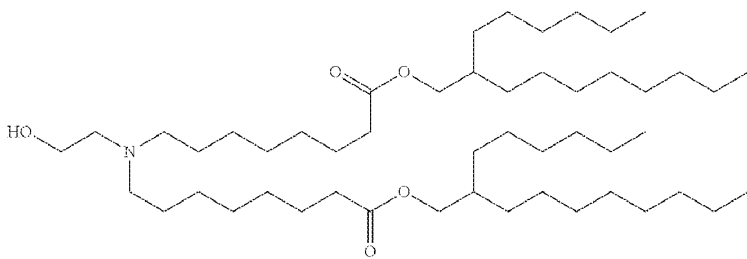
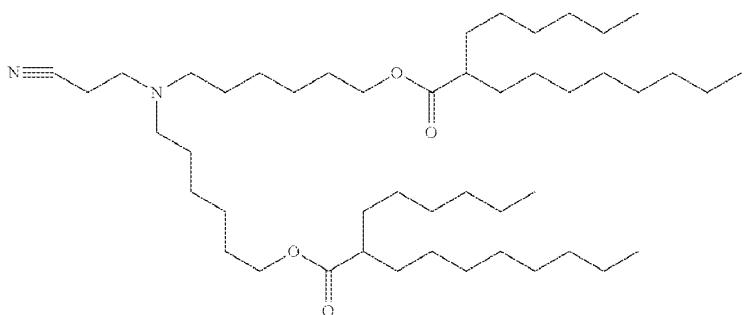
24	
25	
26	
27	

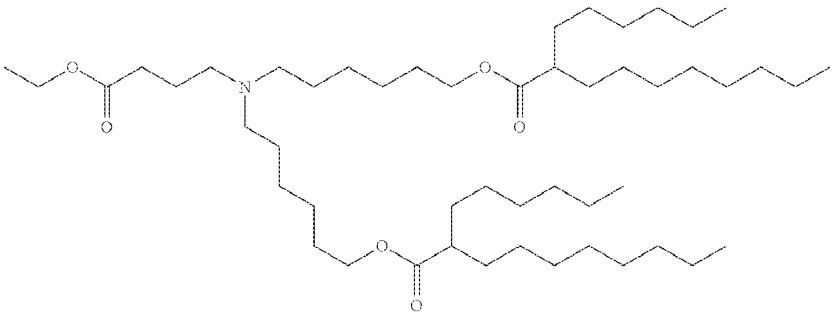
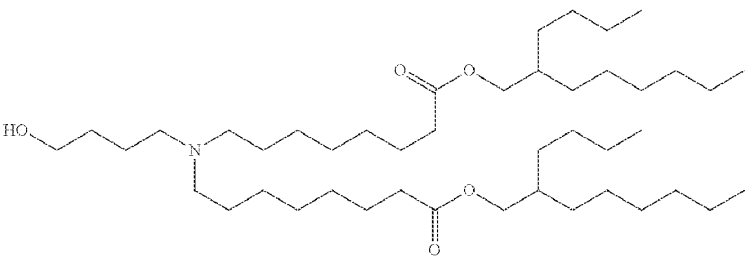
28	
29	
30	
31	

32	 <p>Chemical structure 32: A tertiary amine with a 1-hydroxy-2-(10-hydroxydecyl)ethyl group, a 10-oxodecyl group, and a 10-oxodecyl group.</p>
33	 <p>Chemical structure 33: A tertiary amine with a 10-oxodecyl group, a 10-oxodecyl group, and a 10-oxodecyl group.</p>
34	 <p>Chemical structure 34: A tertiary amine with a 10-oxodecyl group, a 10-oxodecyl group, and a 10-oxodecyl group.</p>
35	 <p>Chemical structure 35: A tertiary amine with a 10-oxodecyl group, a 10-oxodecyl group, and a 10-oxodecyl group.</p>

36	
37	
38	
39	

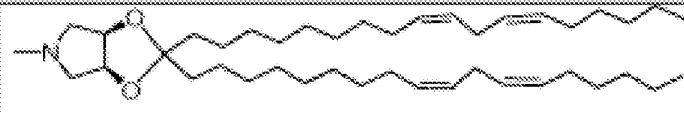
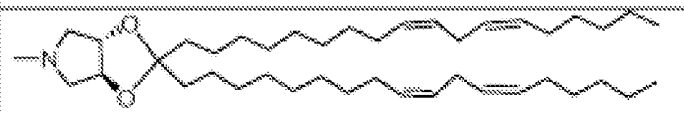
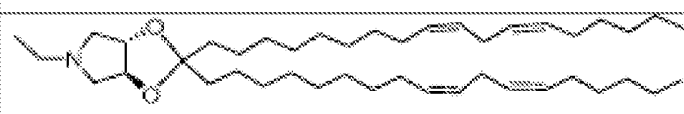
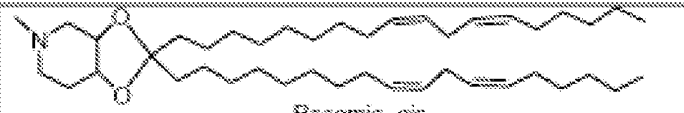
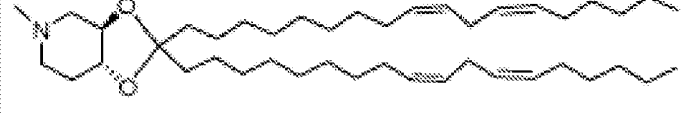
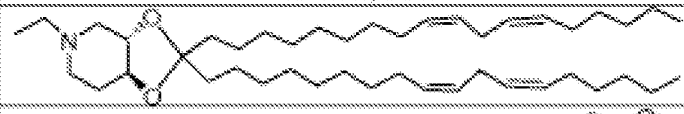
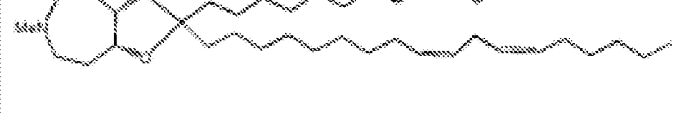
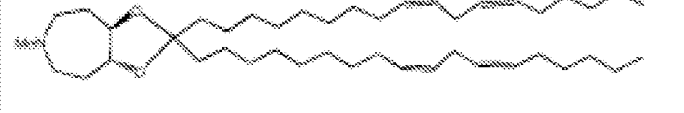
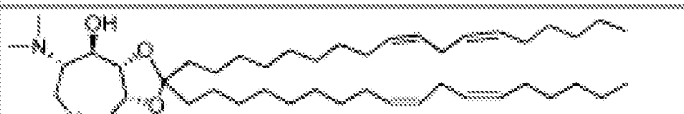
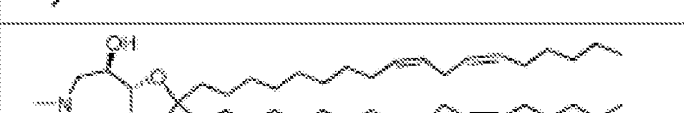
40	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCCCC(=O)OCC)C(C)CCCC</chem>
41	 <chem>CCCCCCCCCCCC(=O)OCCN(COCC)CCCCCCCCCCCC(=O)OCC</chem>
42	 <chem>CCCCCCCCCCCC(=O)OCCN(COCCC)CCCCCCCCCCCC(=O)OCC</chem>
43	 <chem>CCCCCCCCCCCC(=O)OCCN(COCCC)CCCCCCCCCCCC(=O)OCC</chem>

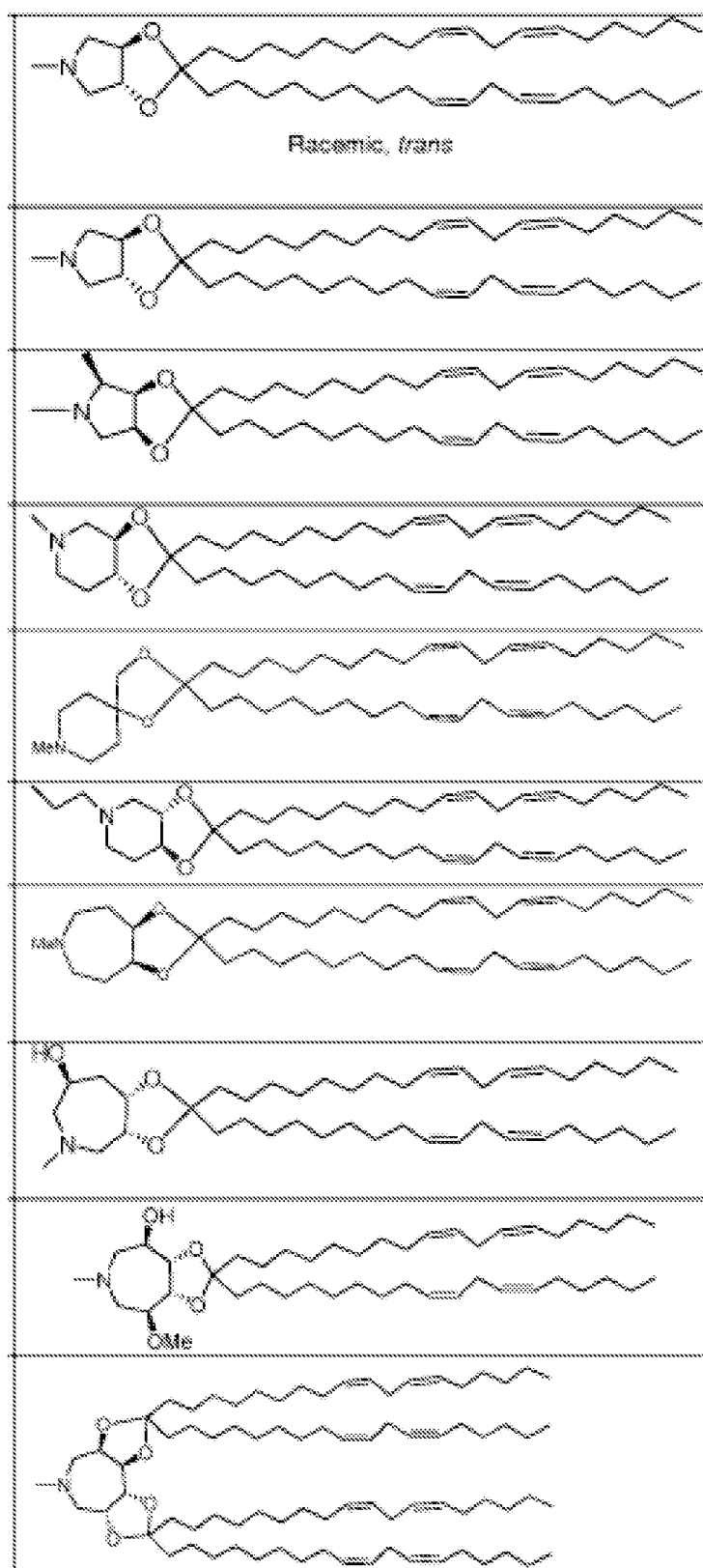
44	
45	
46	
47	

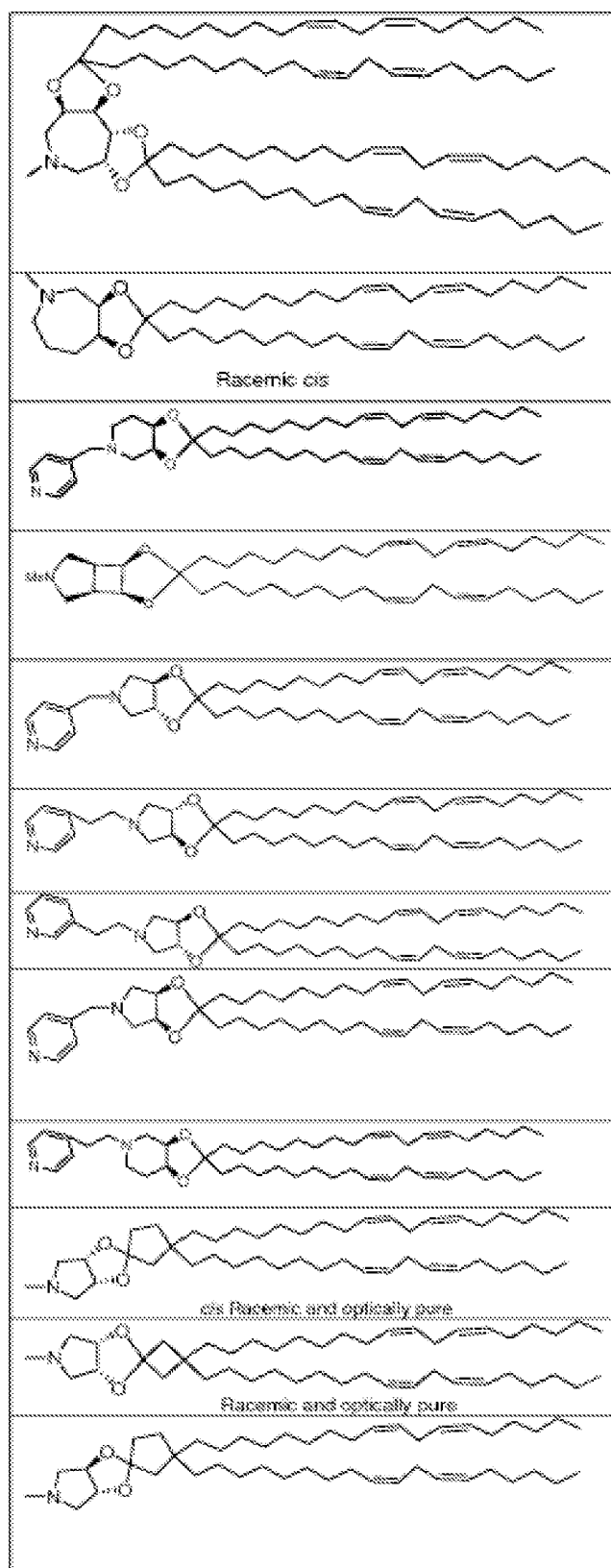
48	
49	

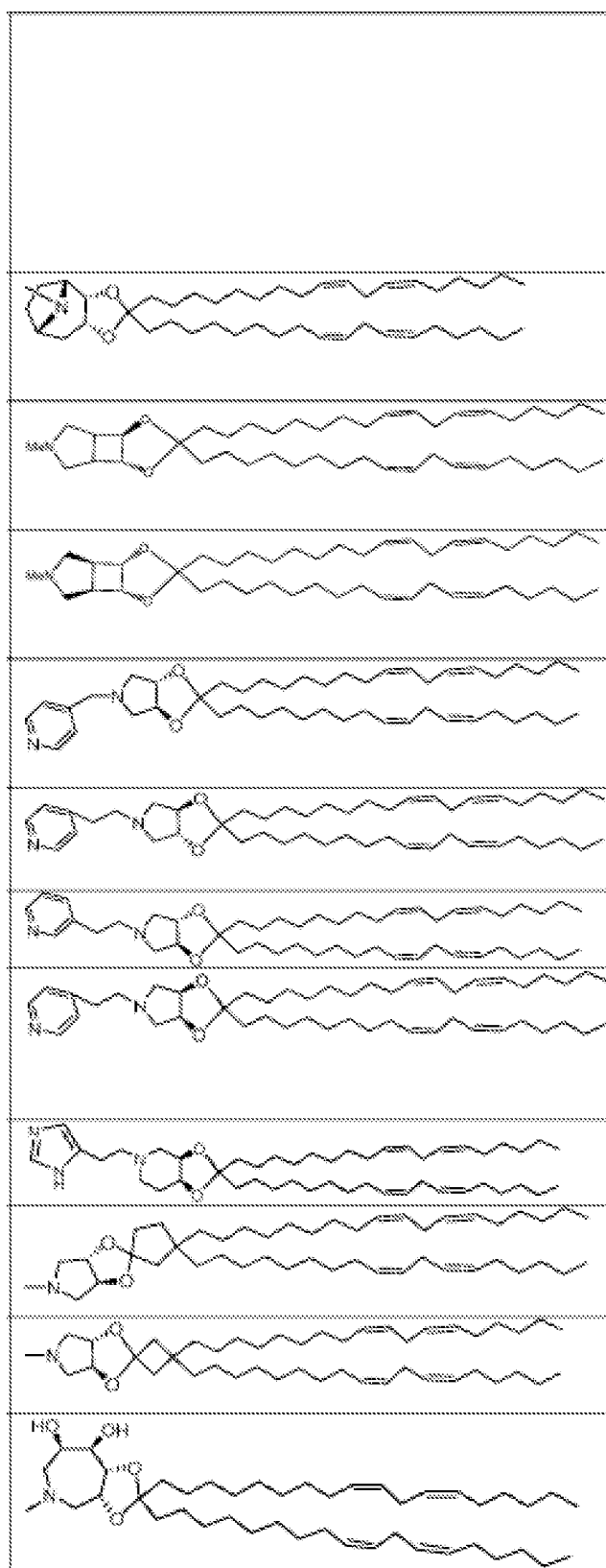
[0466] In some embodiments, the ionizable lipid has one of the structures set forth in Table 11 below. In some embodiments, the ionizable lipid as set forth in Table 11 is as described in international patent application PCT/US2010/061058.

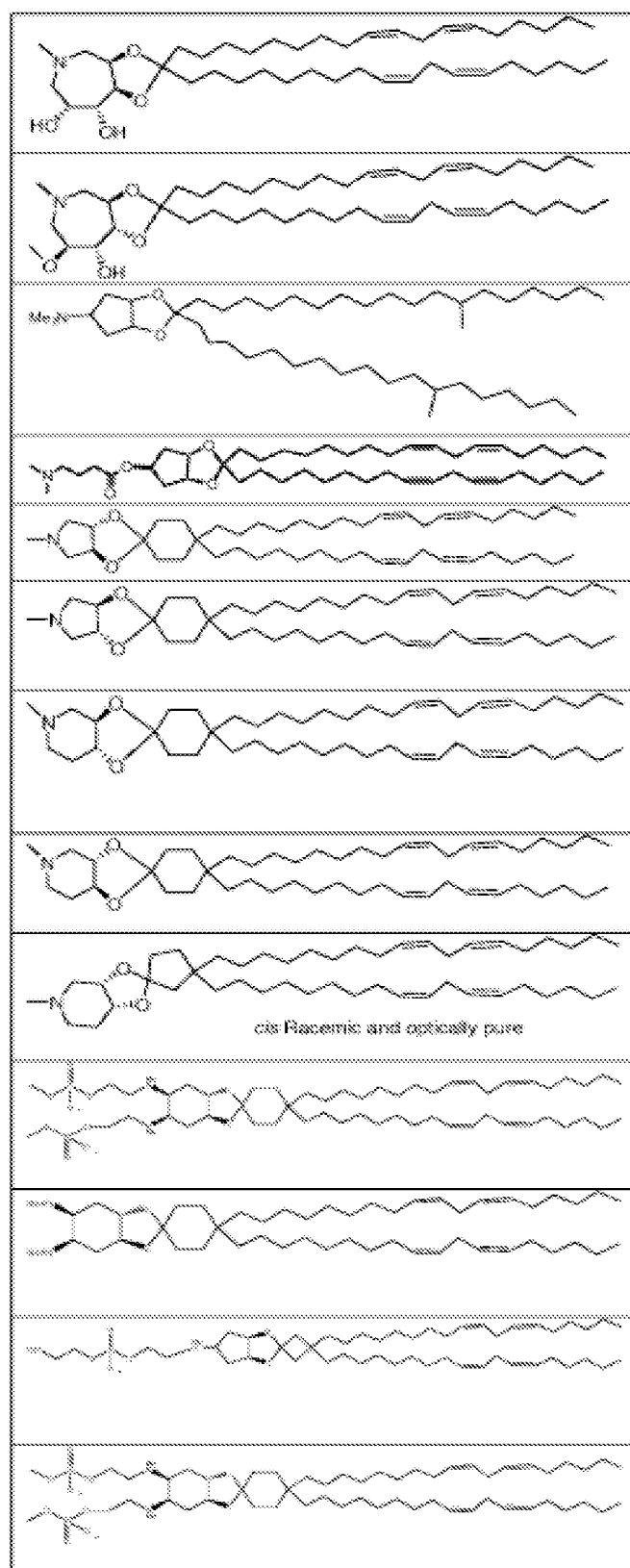
Table 11

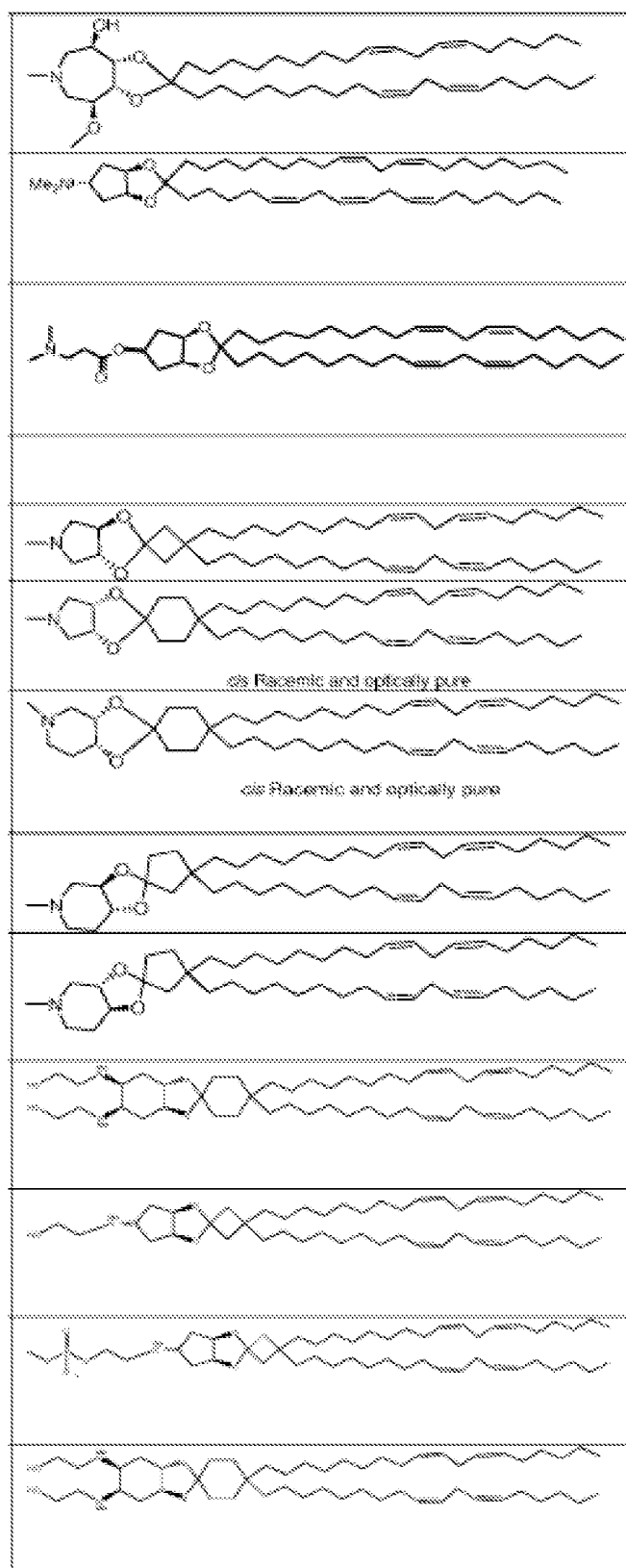
 <i>meso, cis</i>


 <i>Racemic, cis</i>
 <i>Racemic, trans</i>






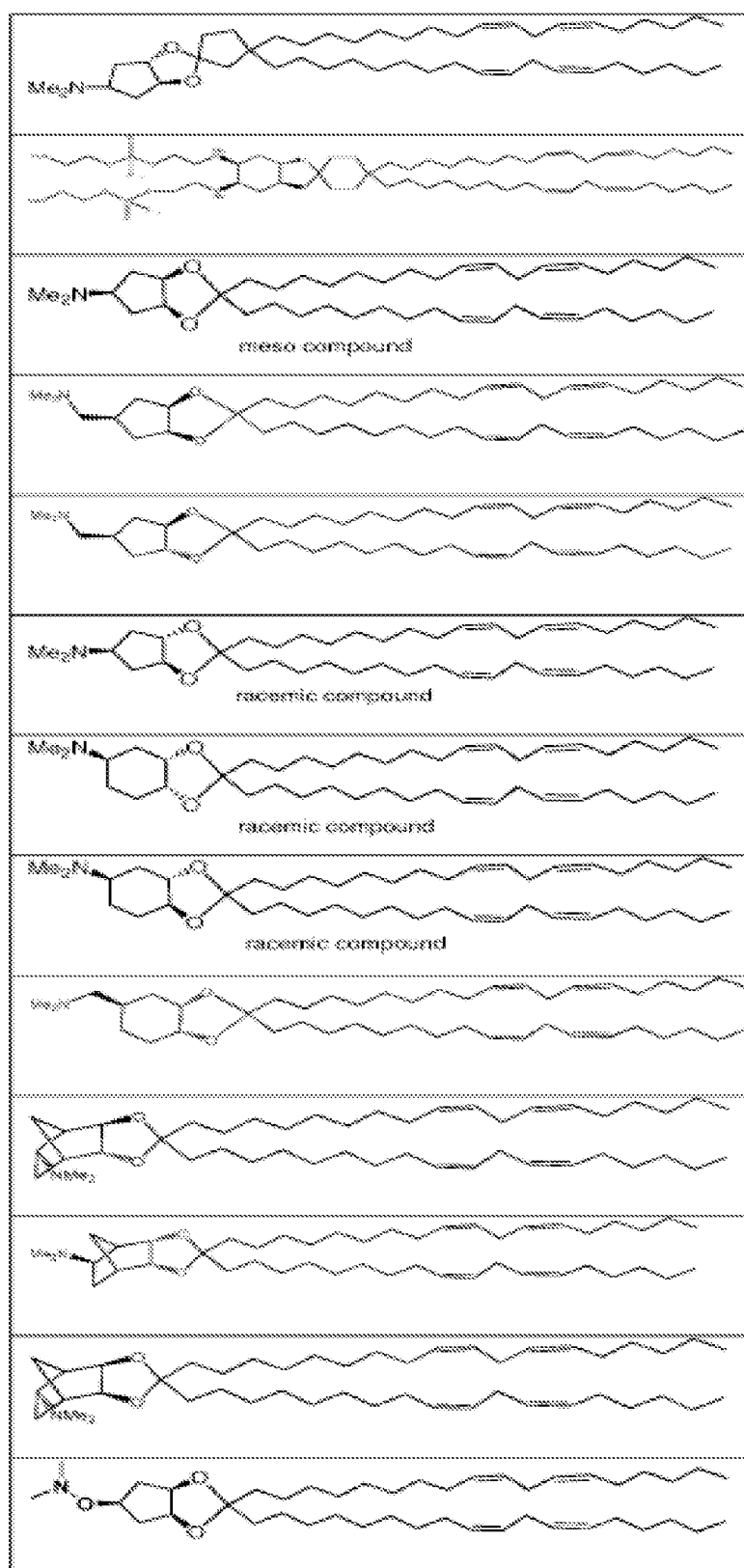


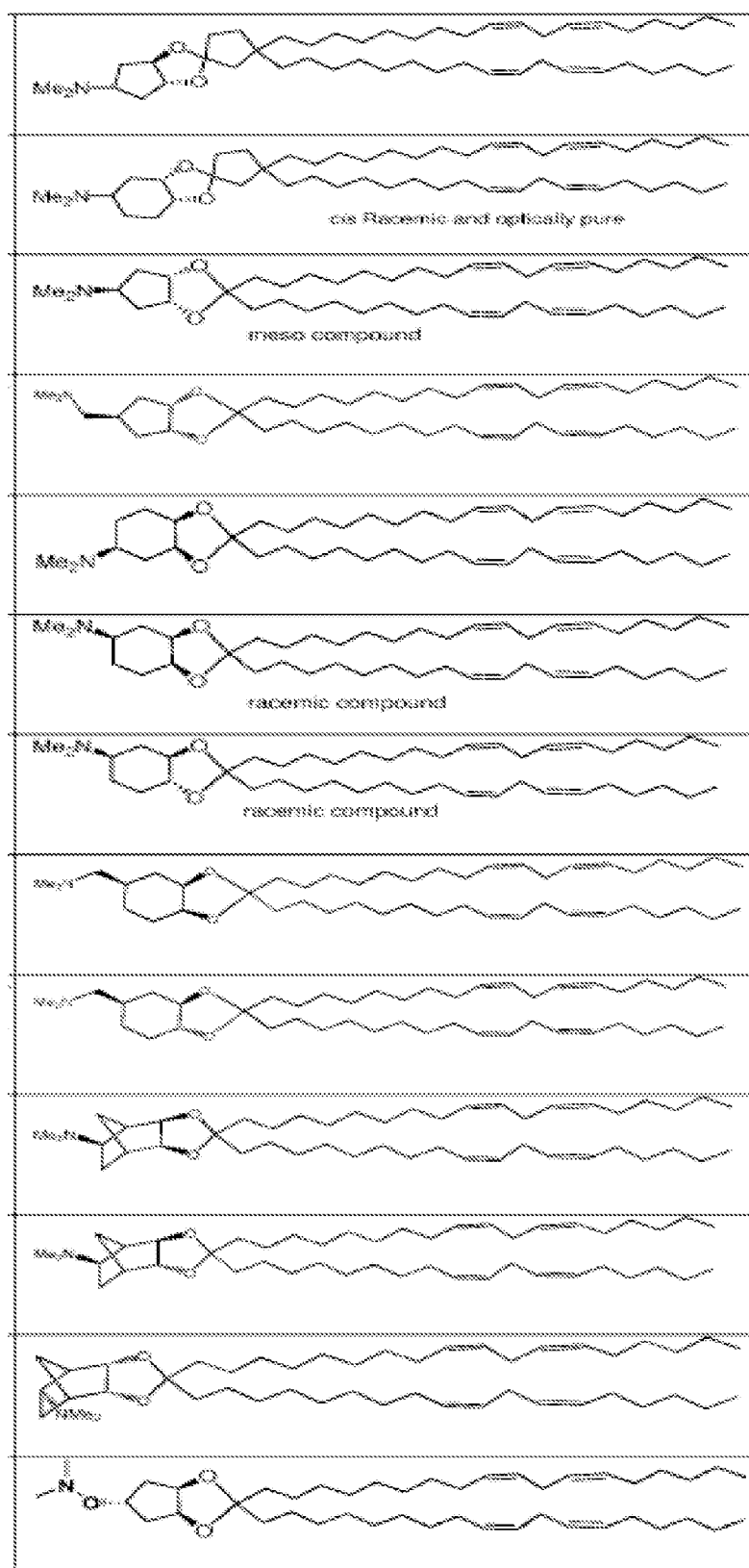


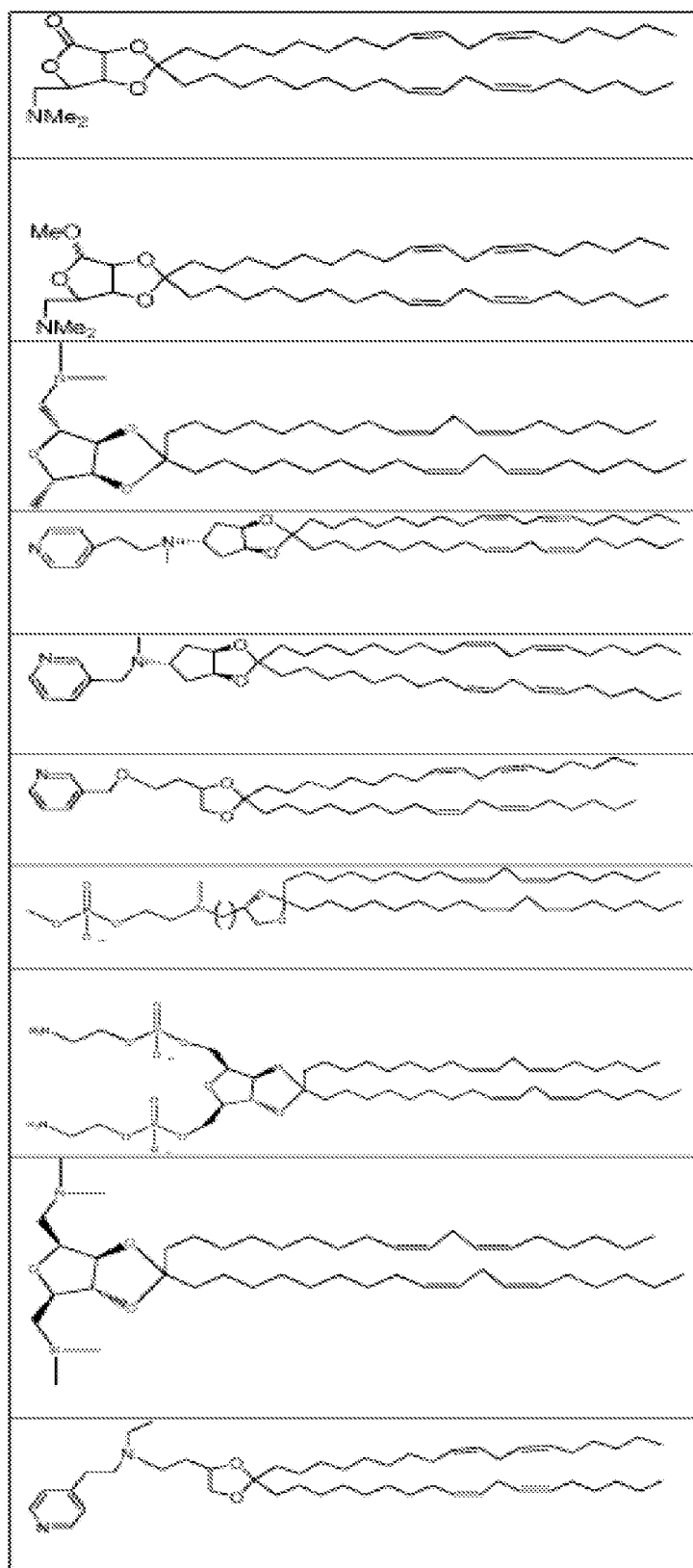


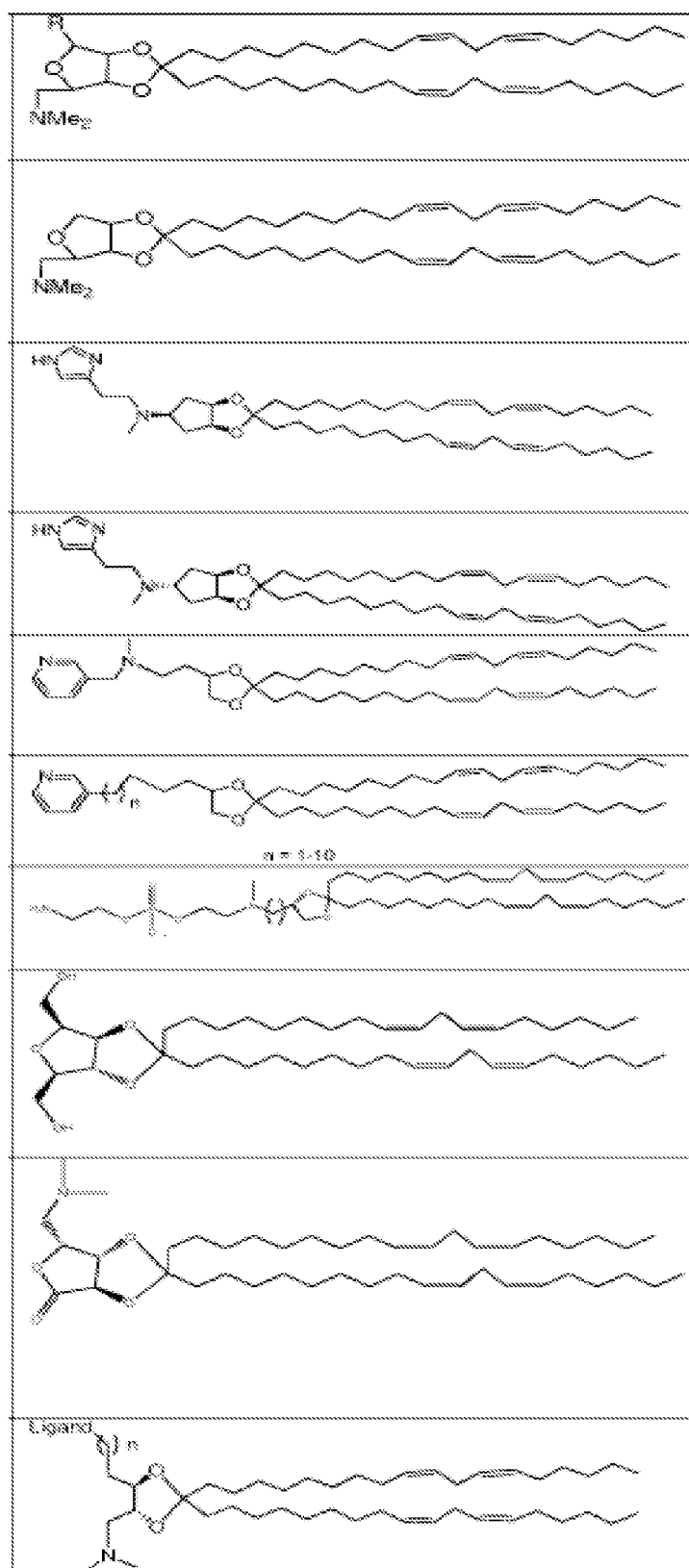


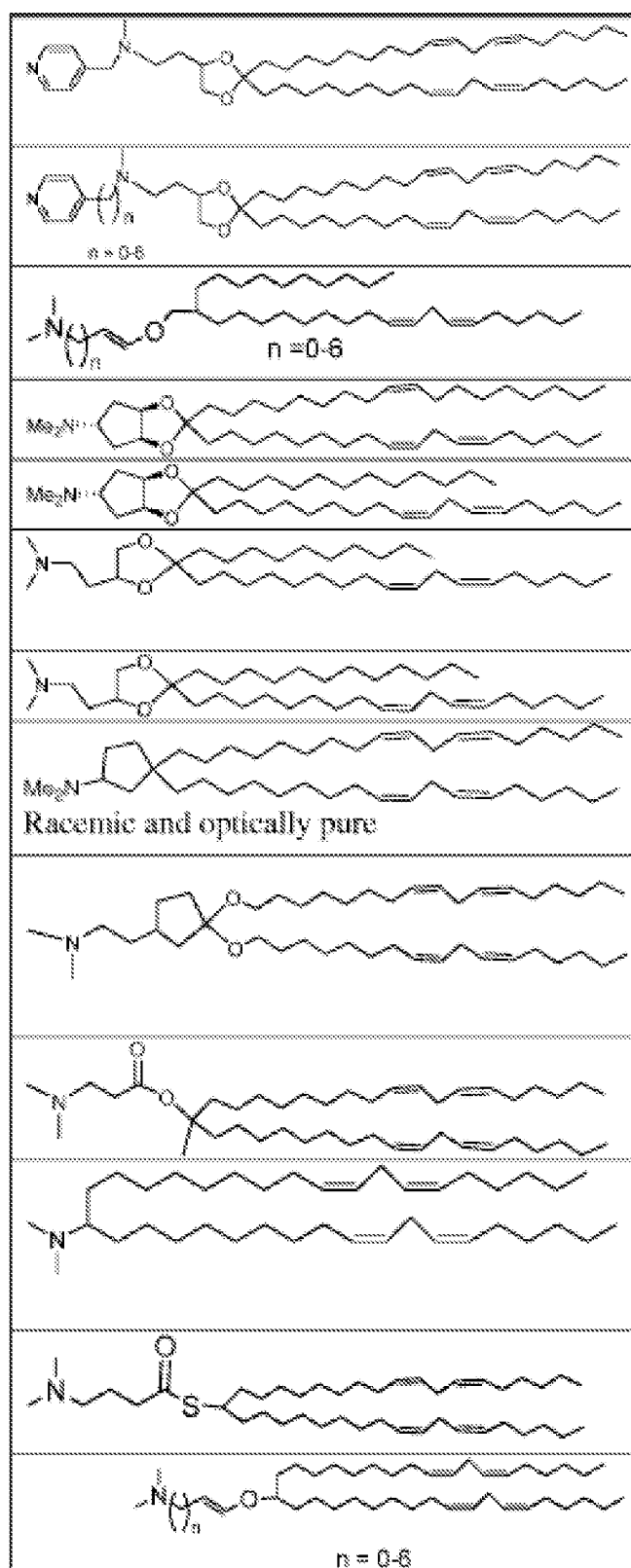


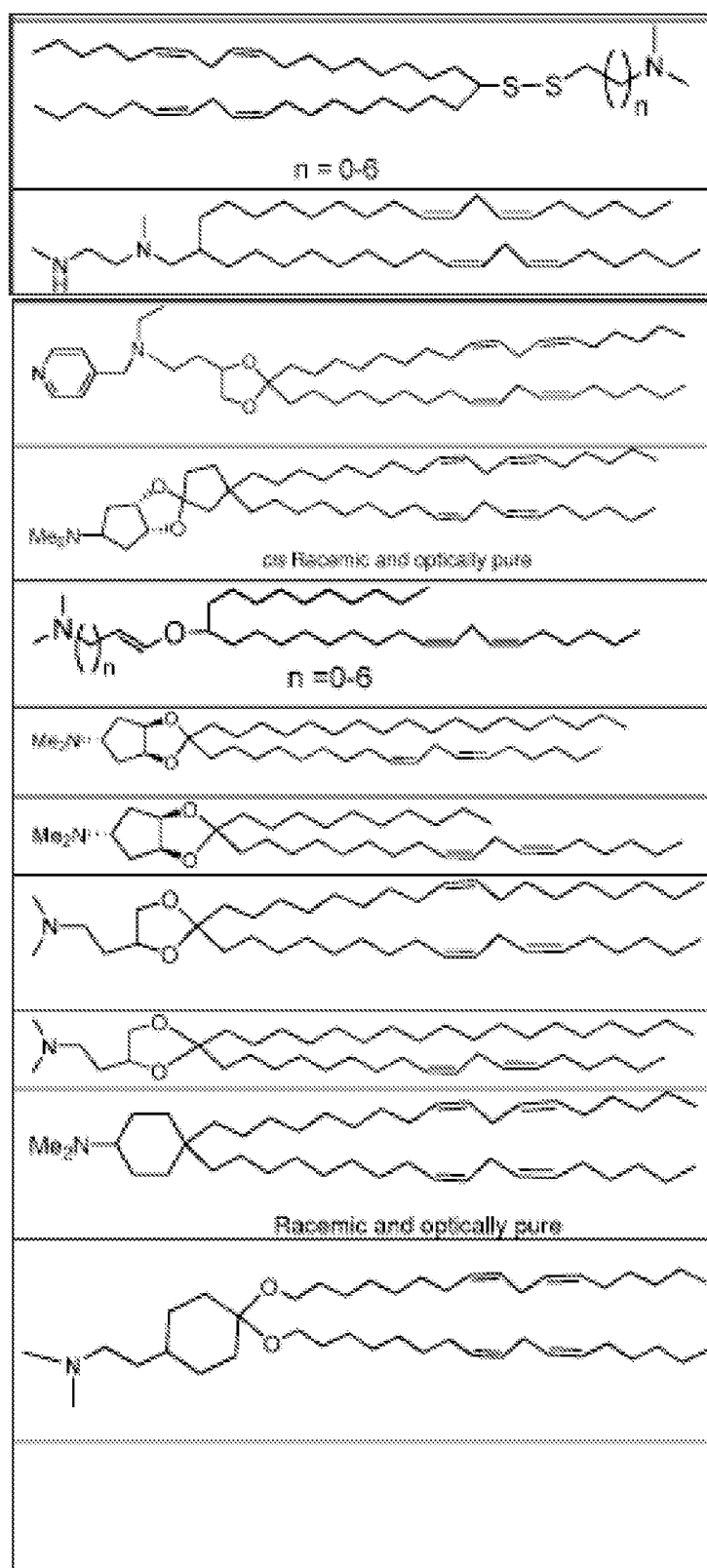


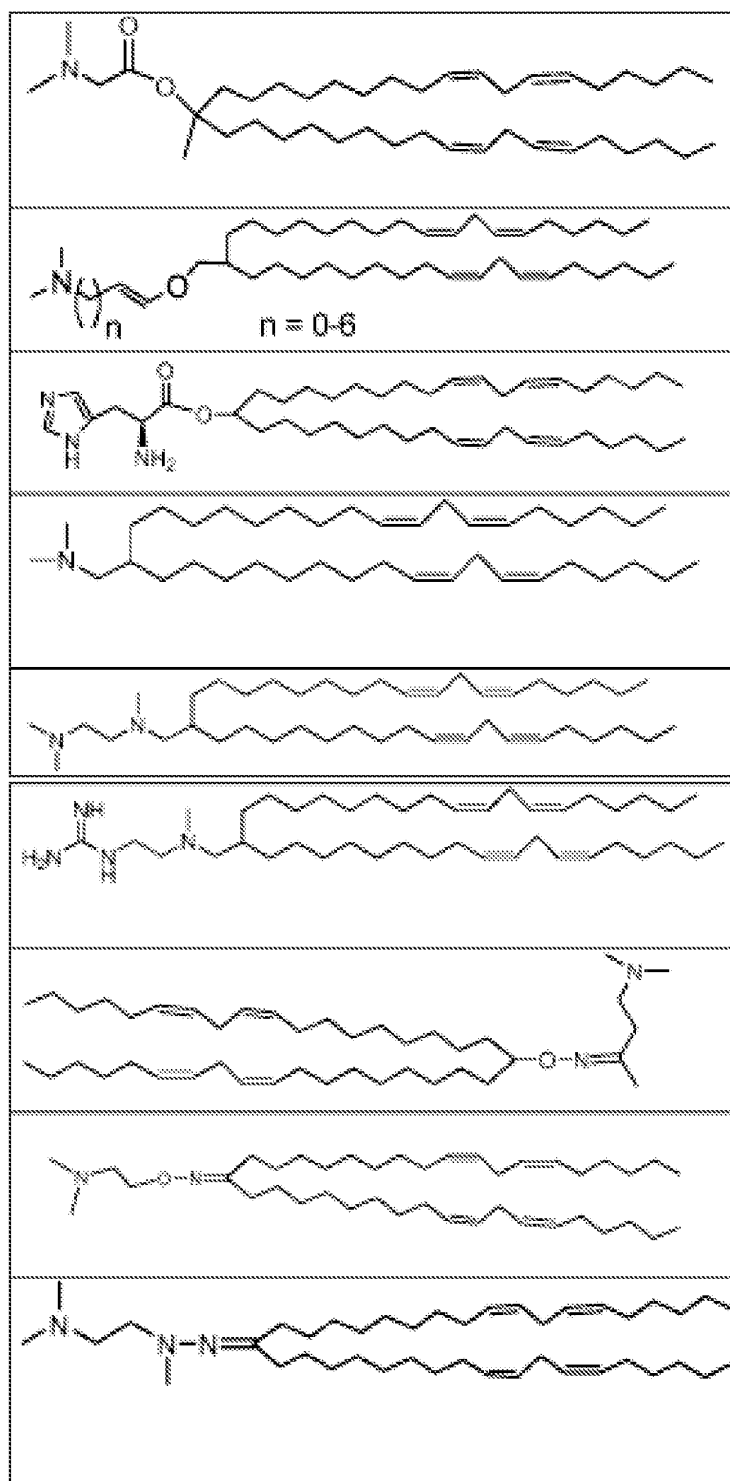


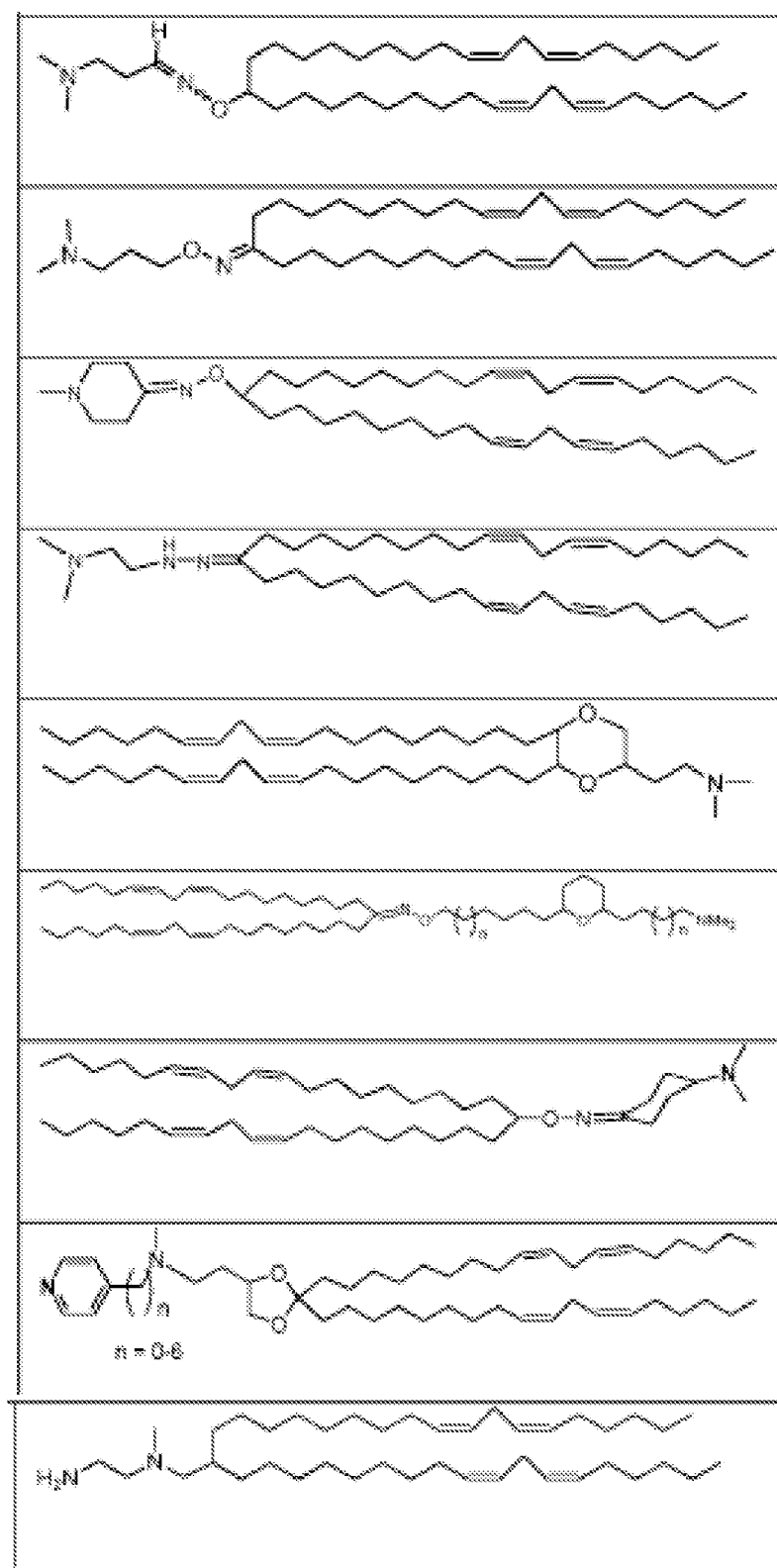


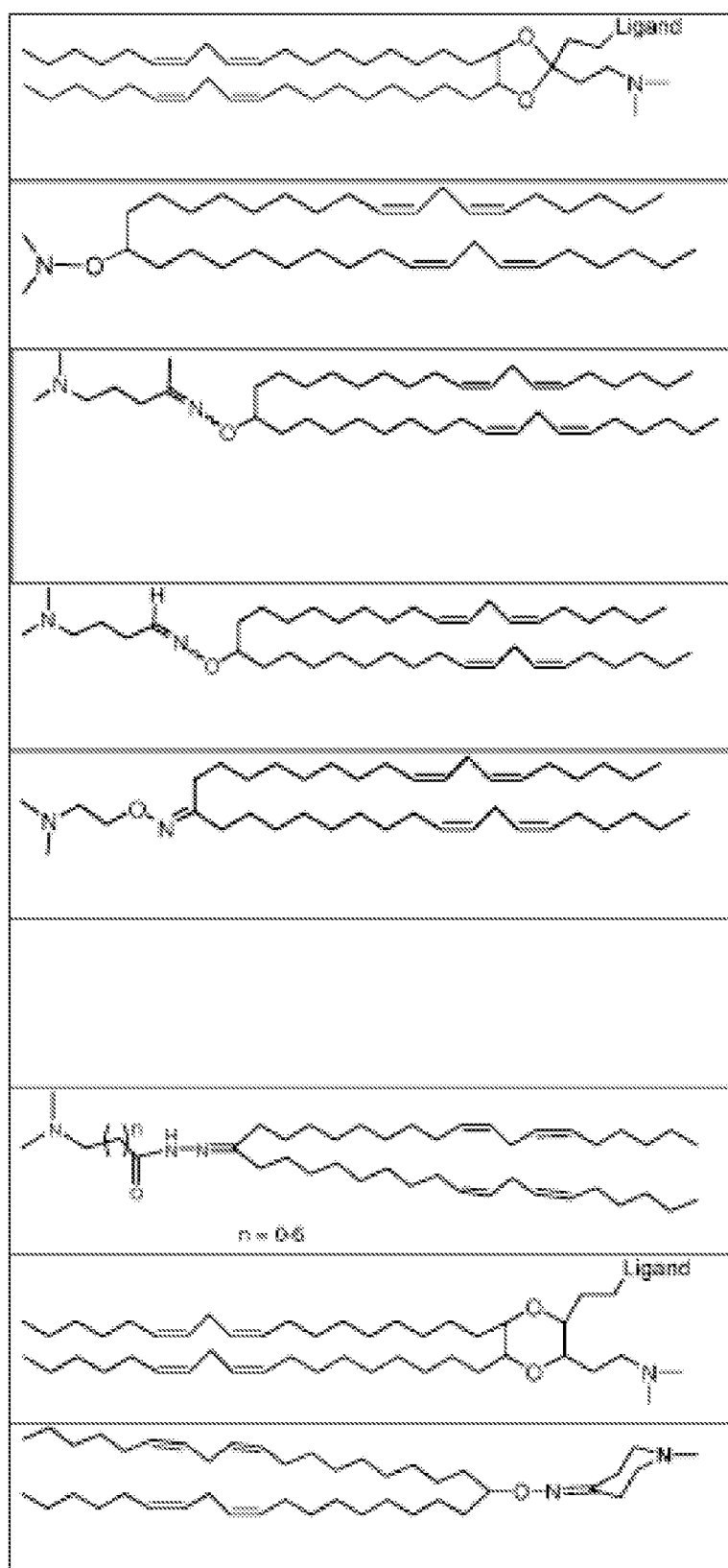


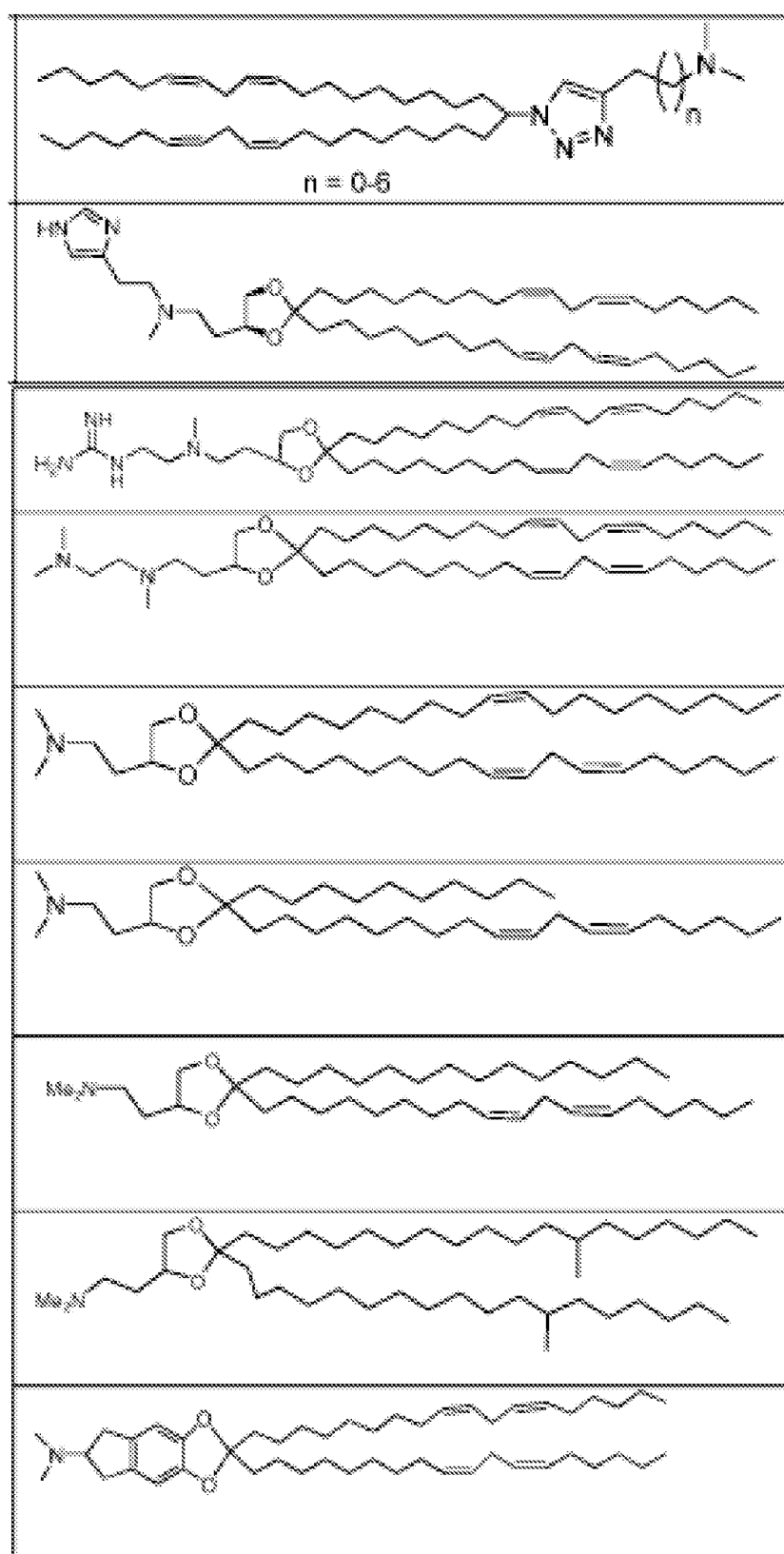


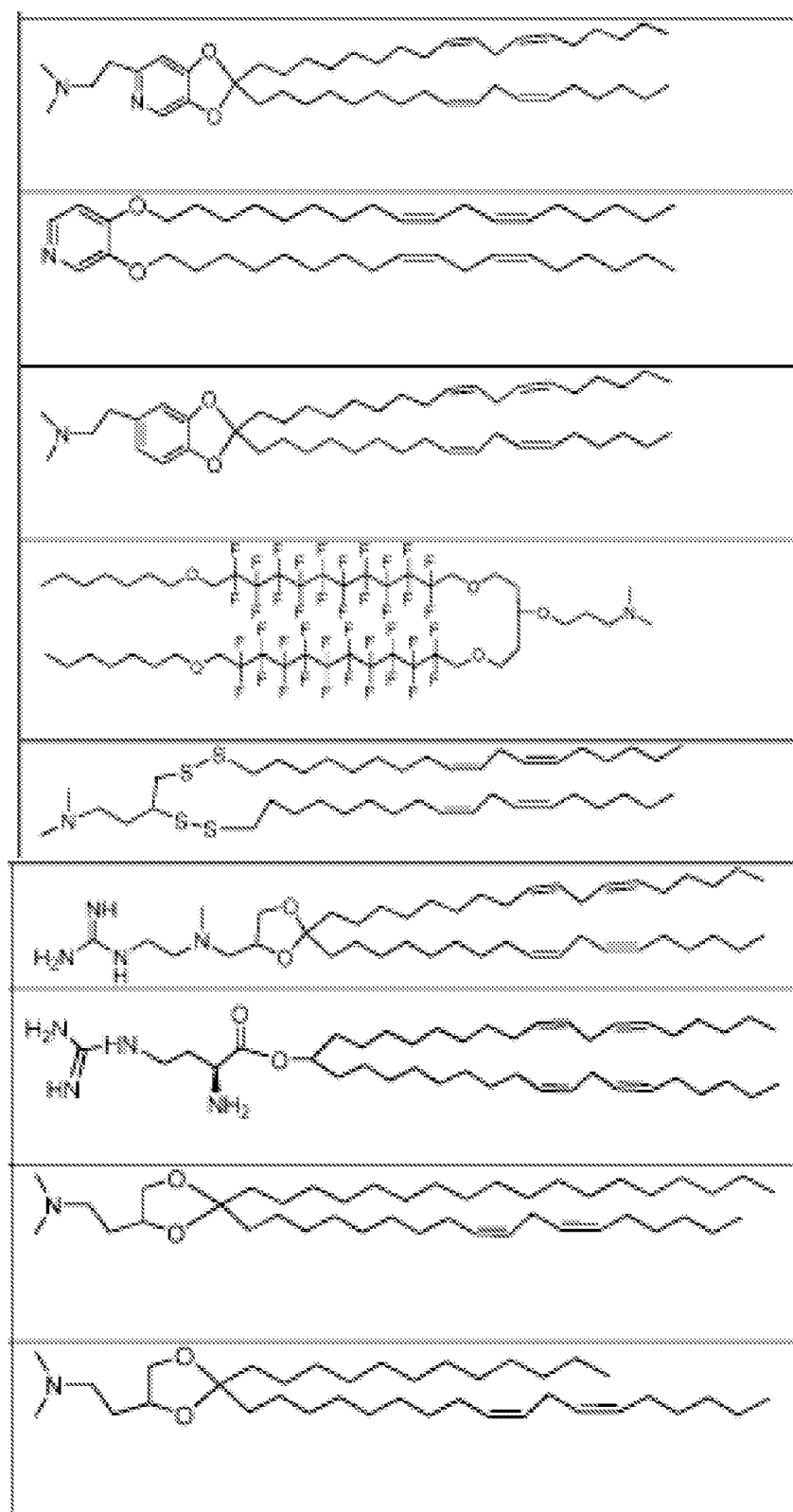


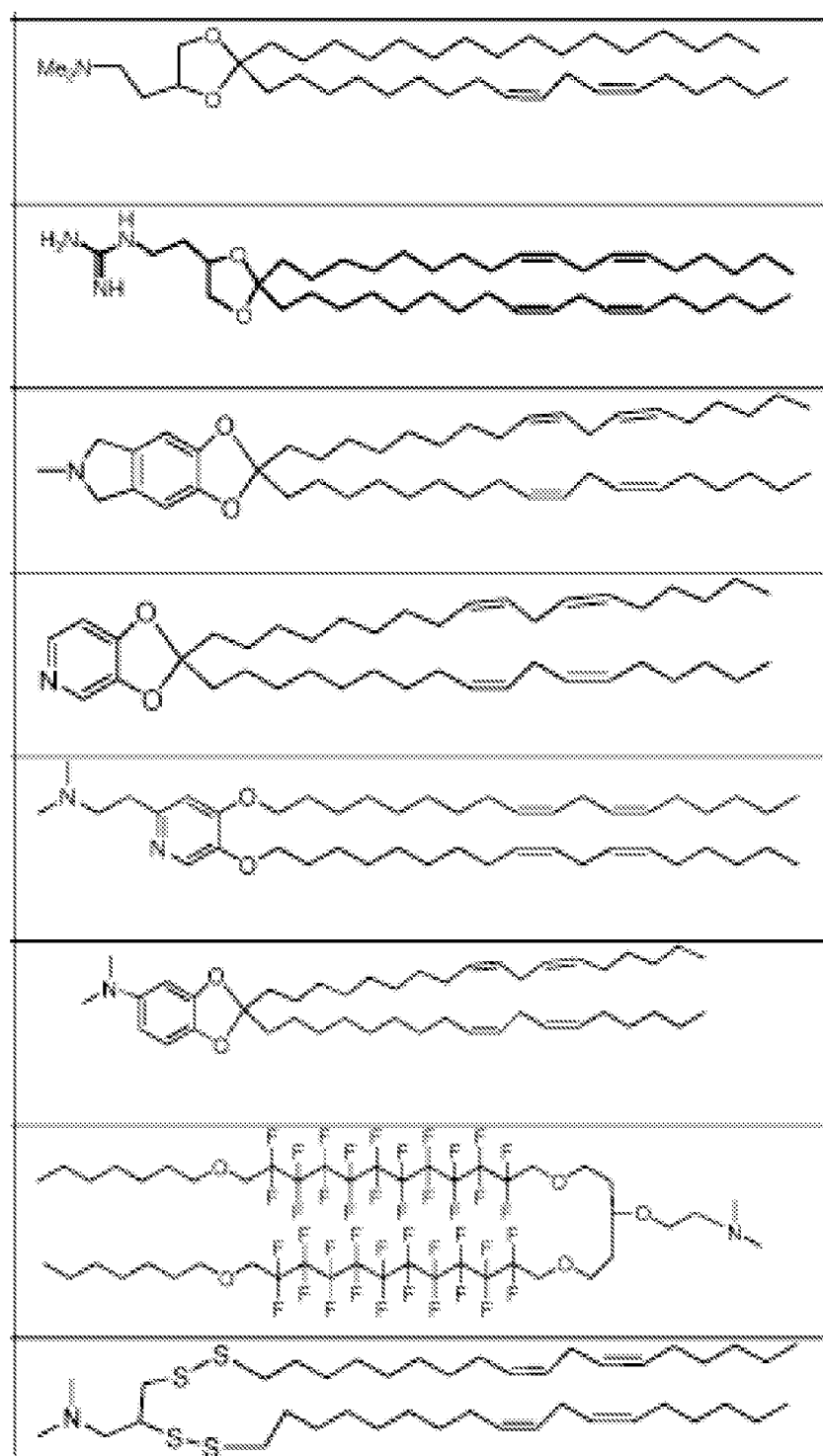


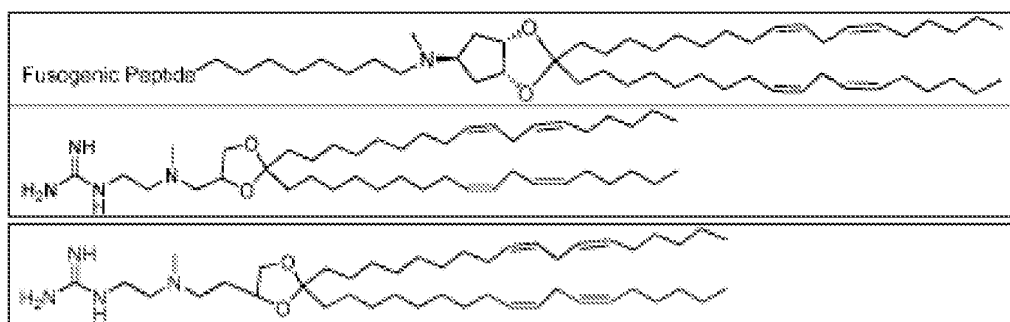
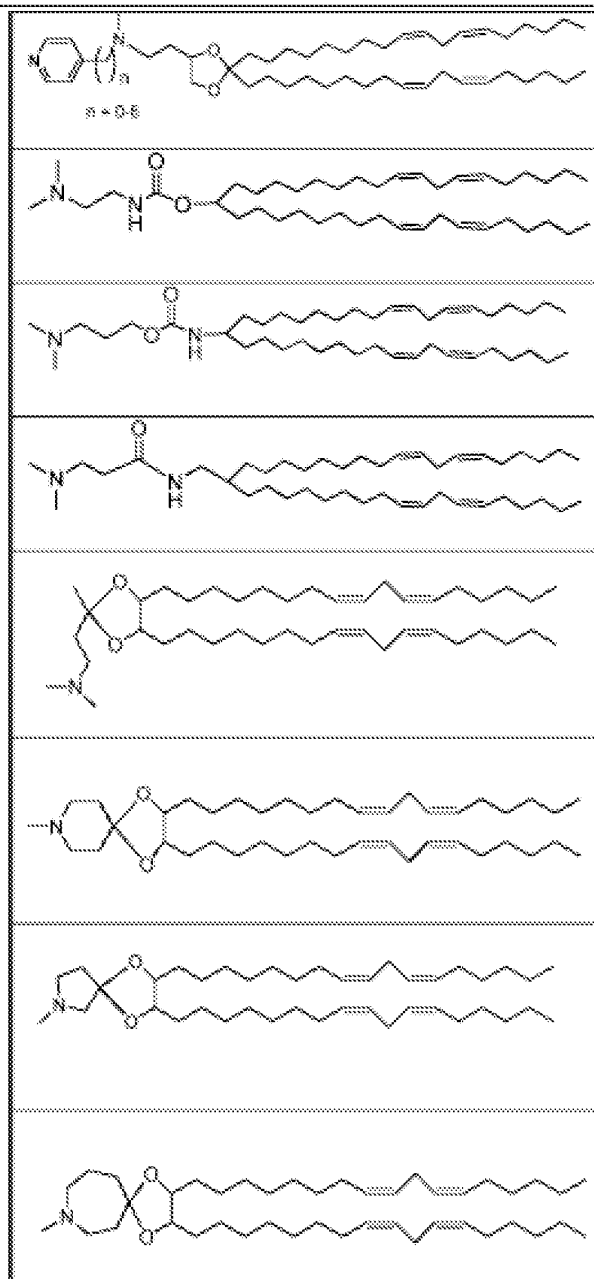


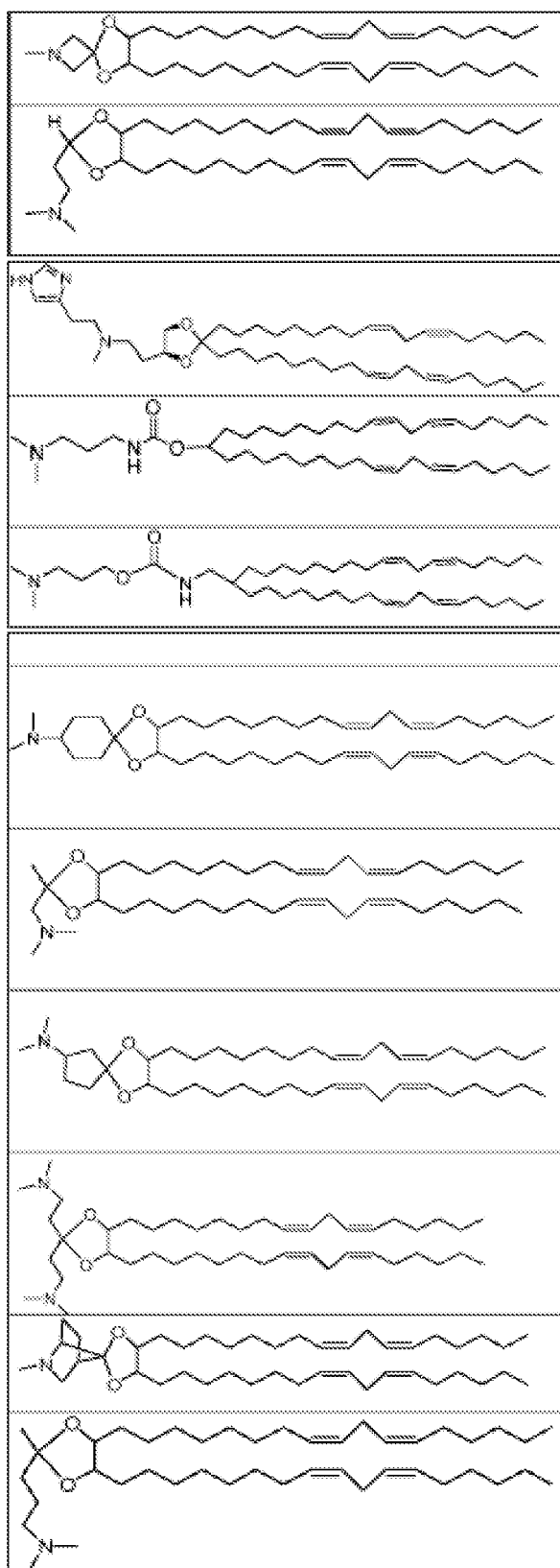


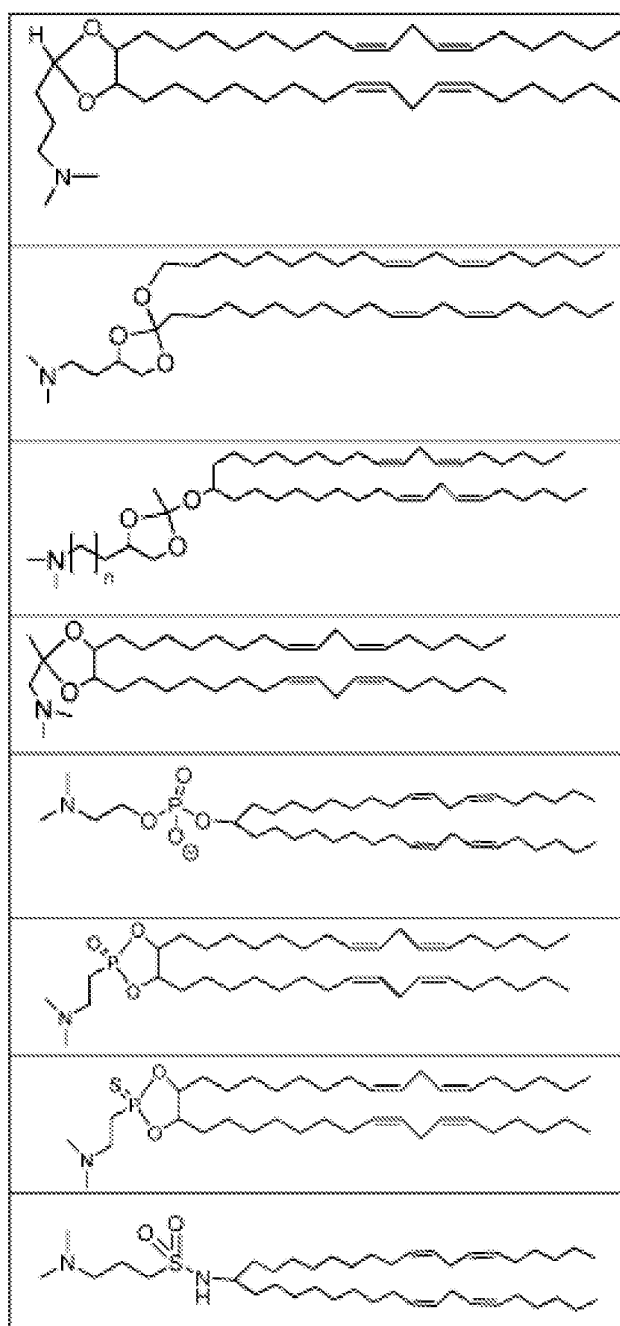


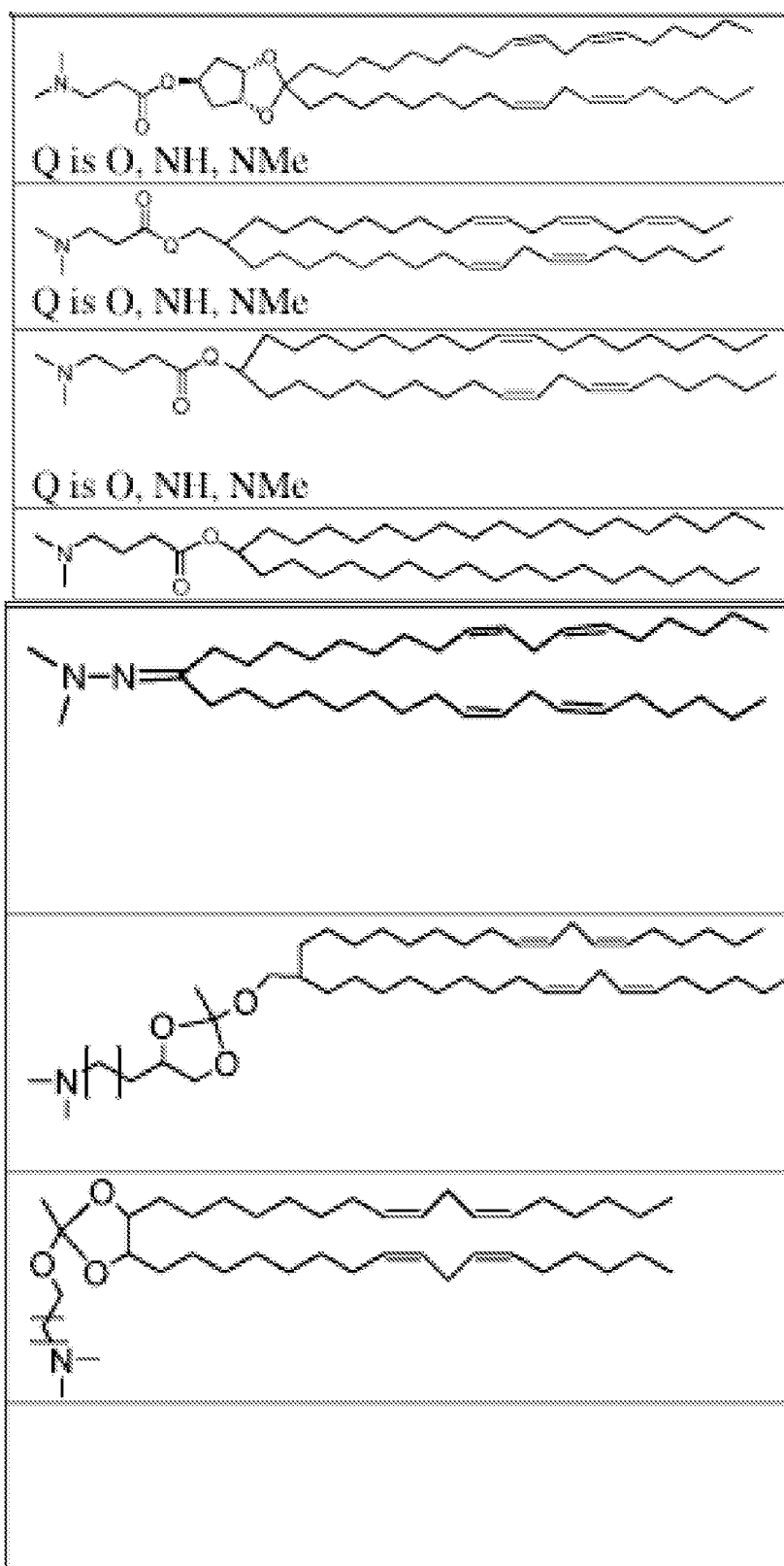


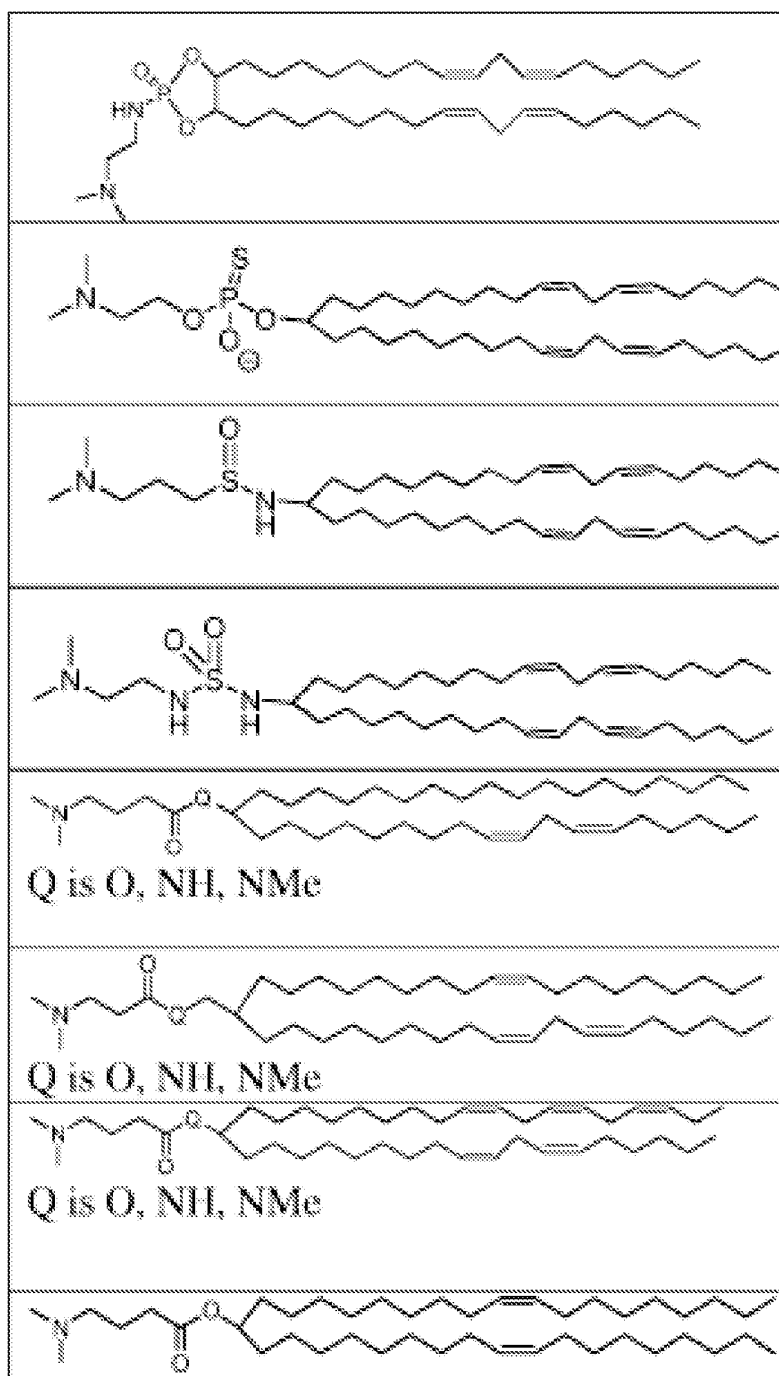


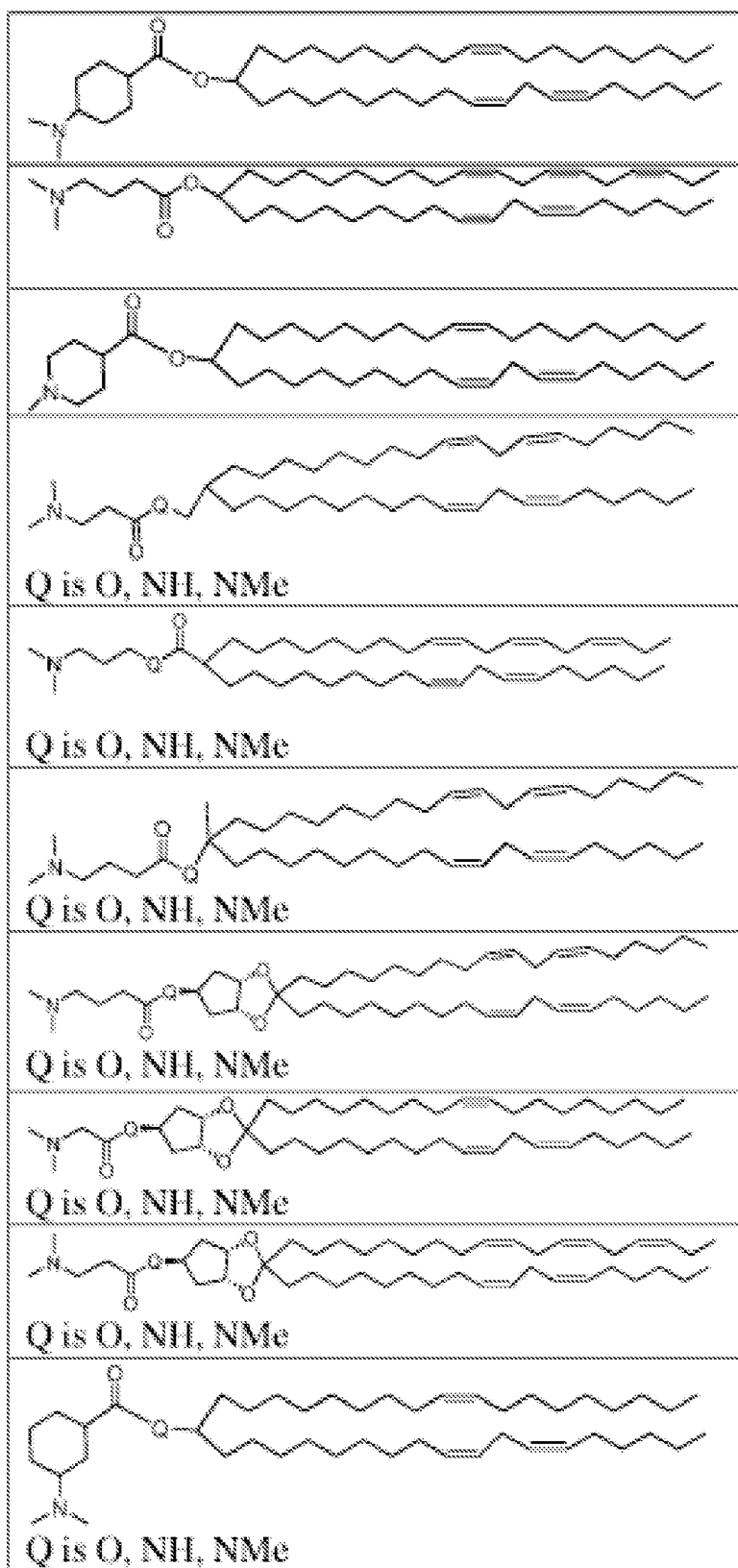


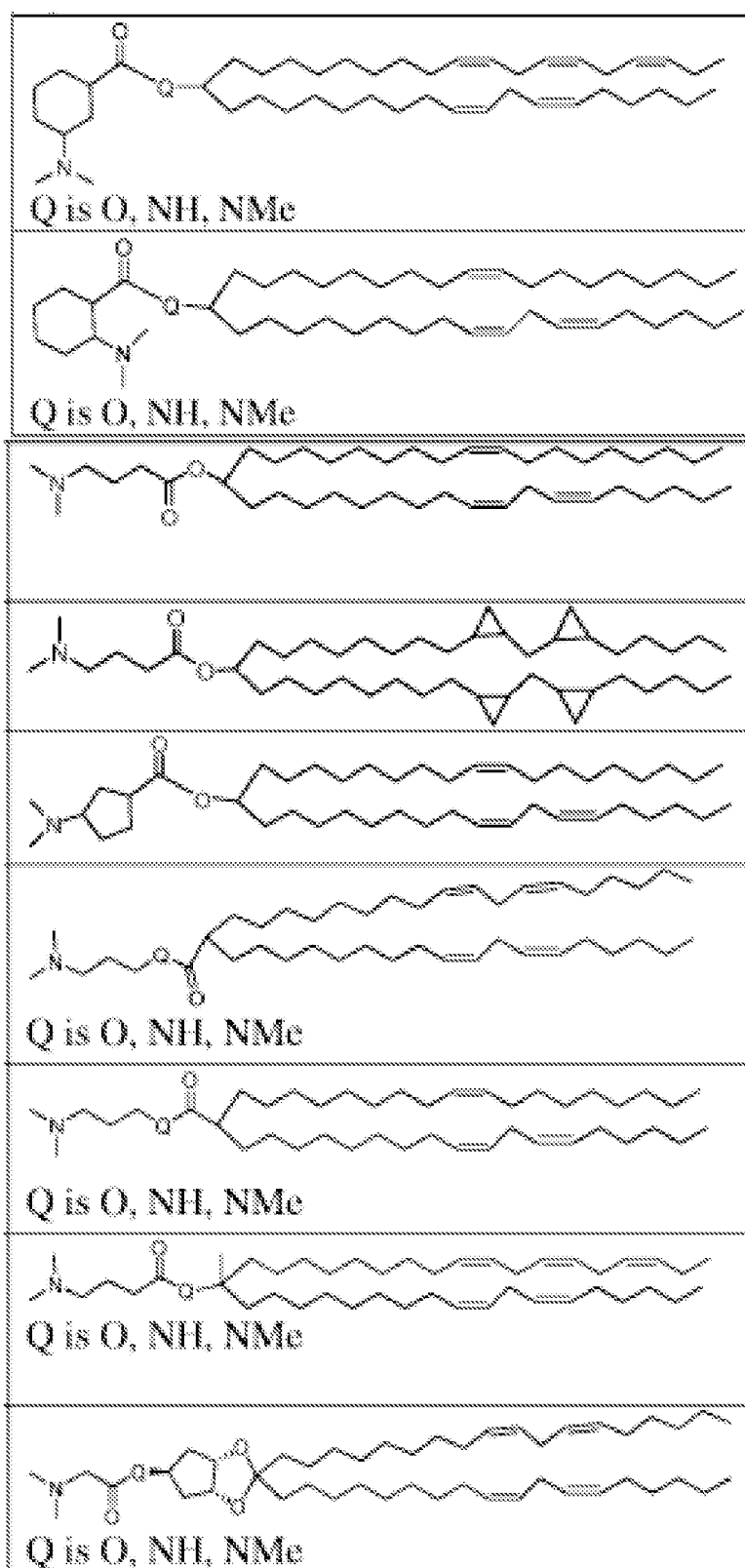


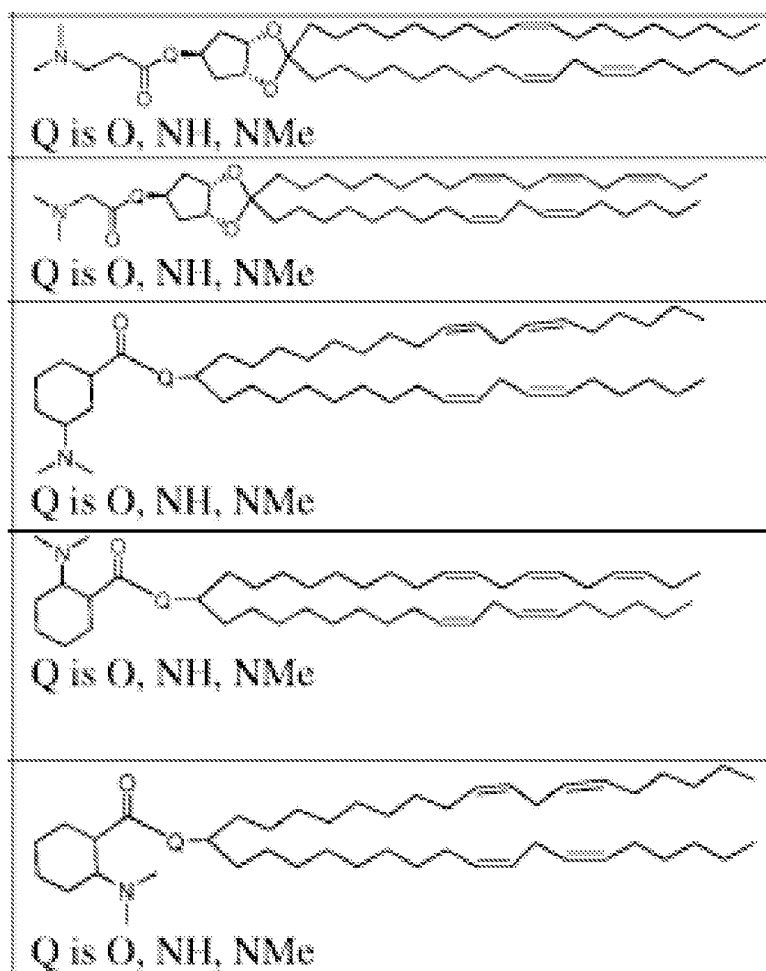


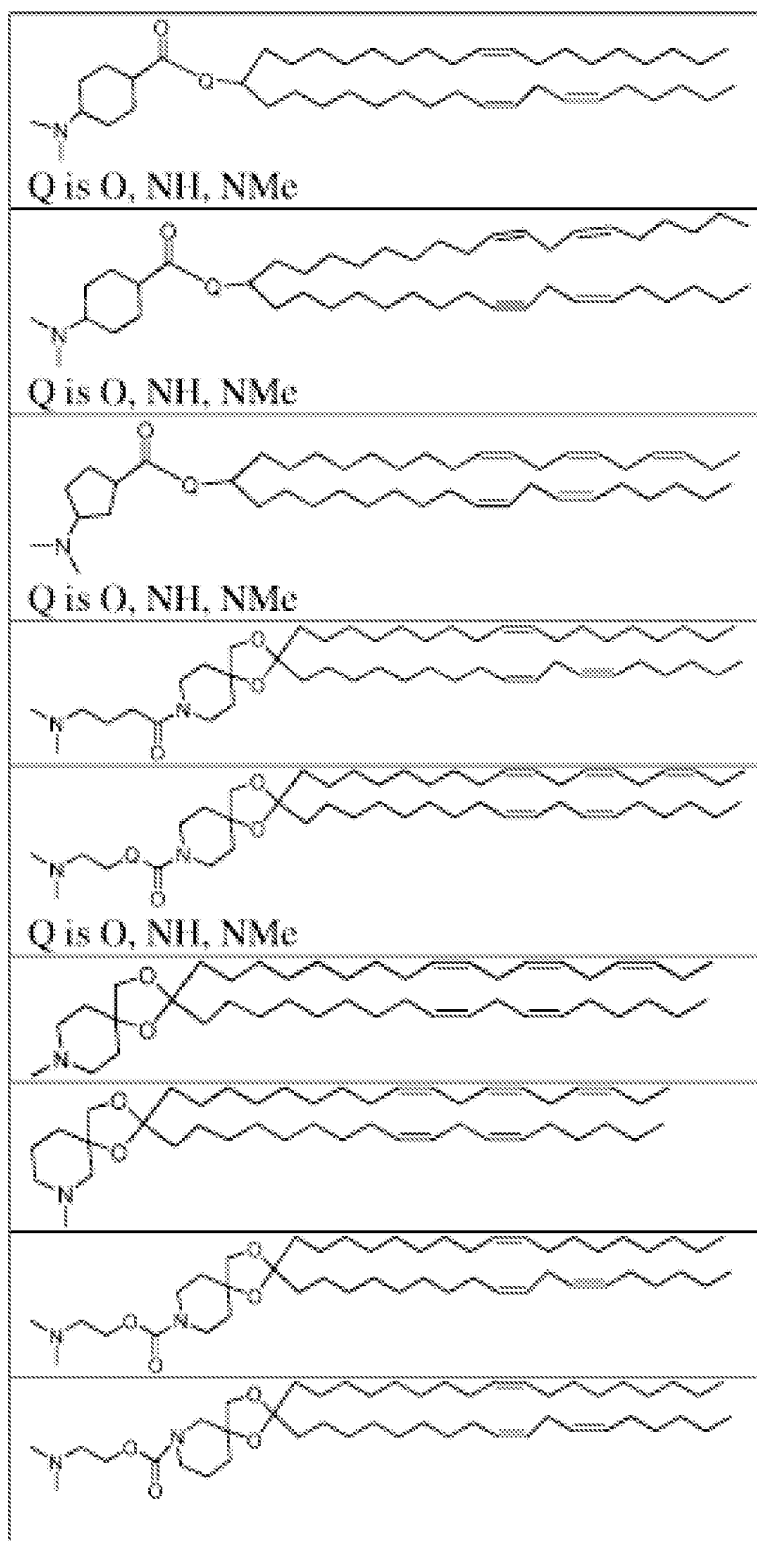


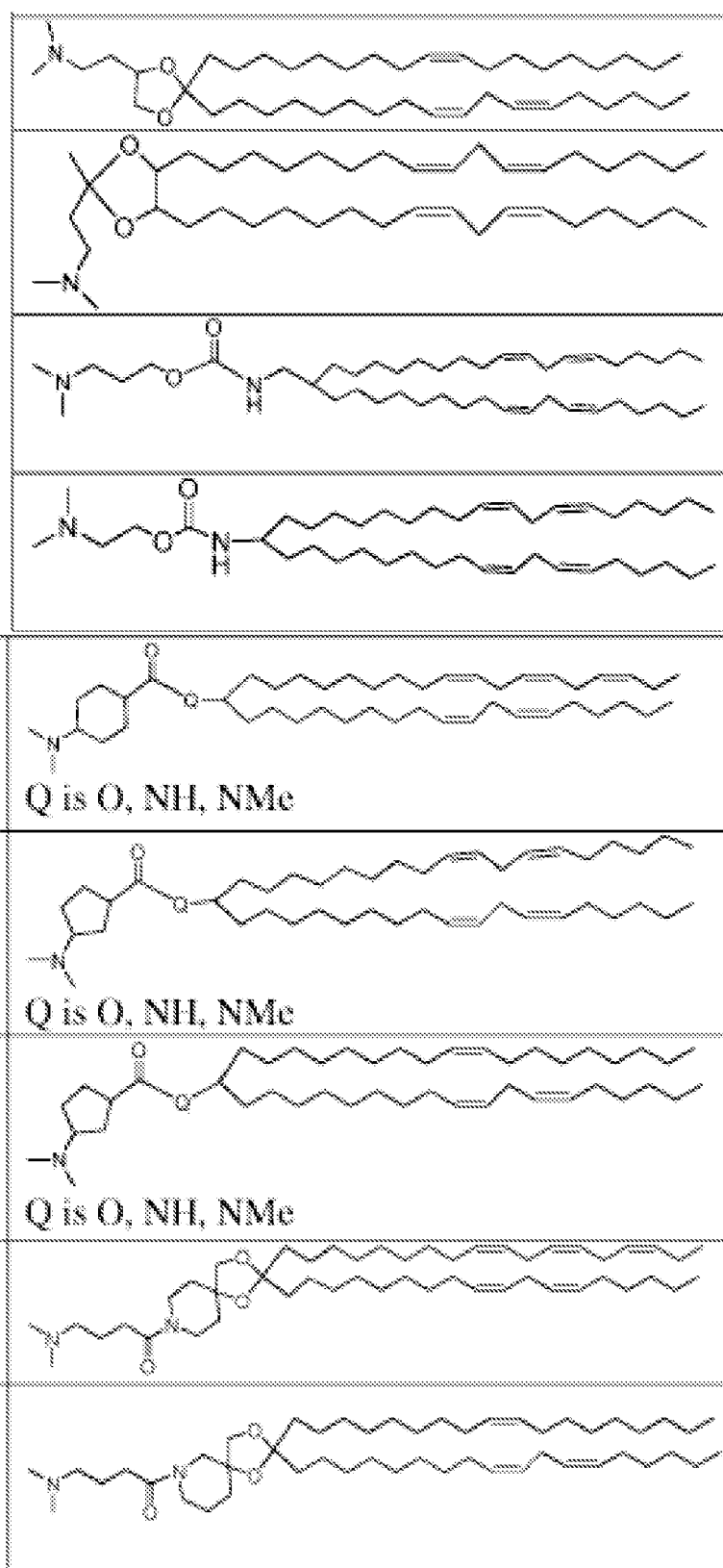


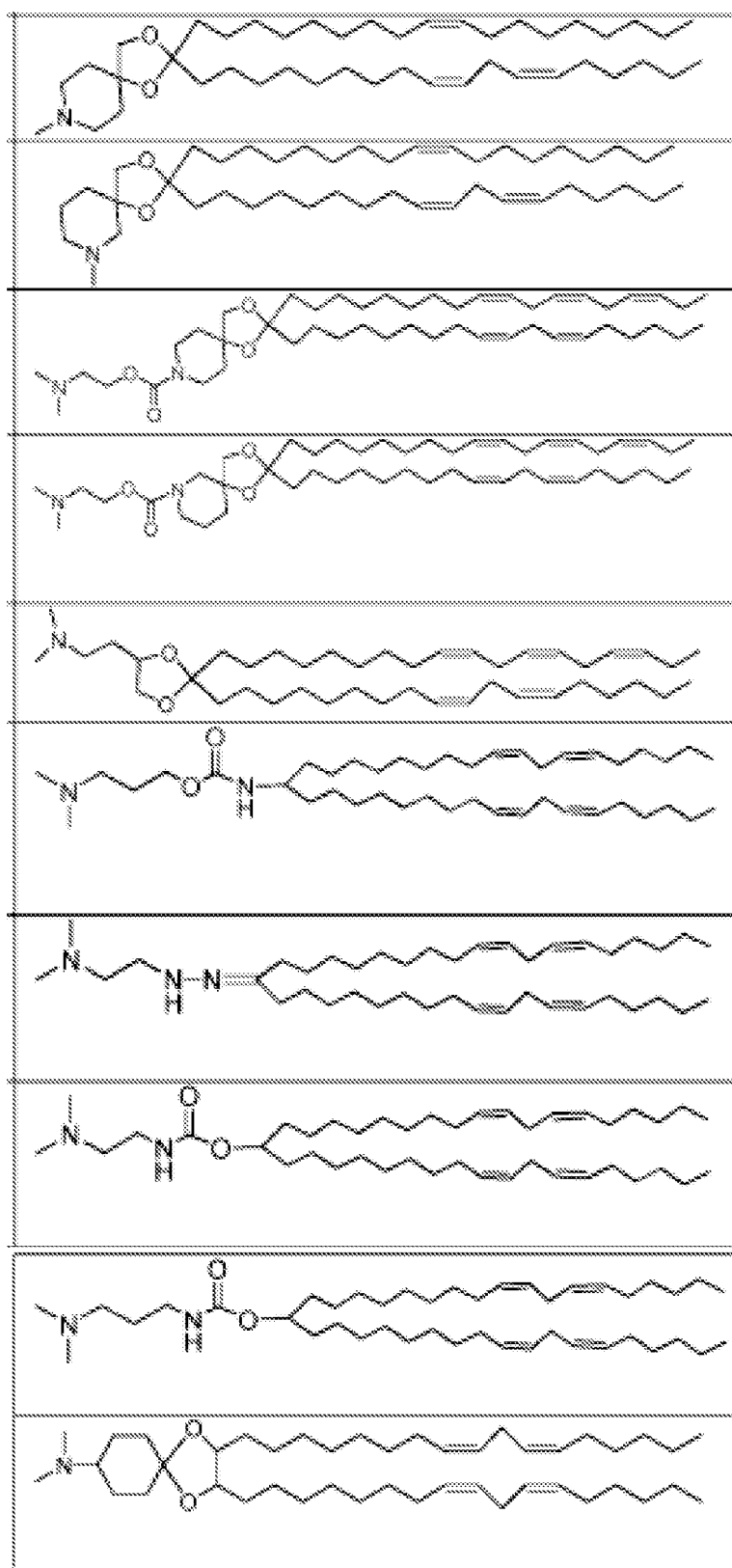


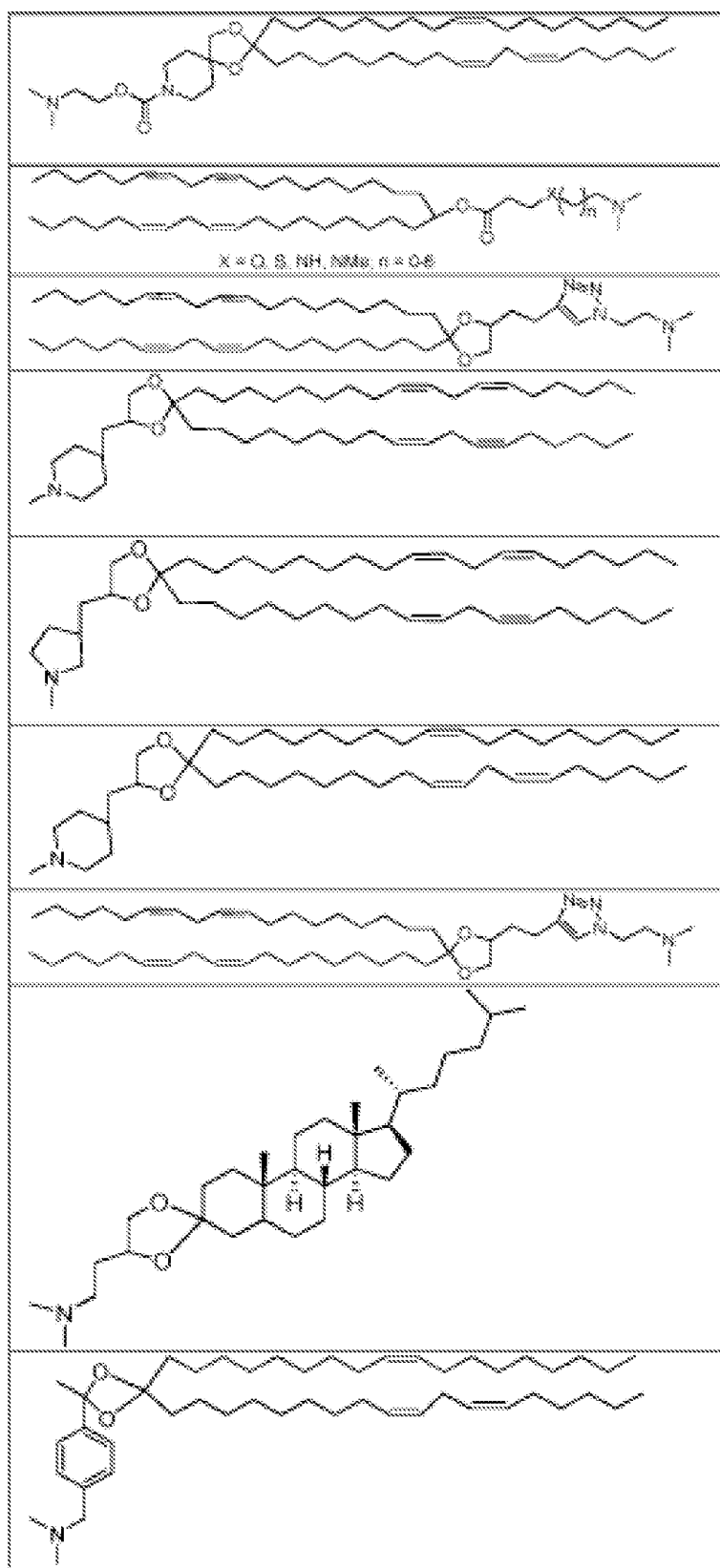


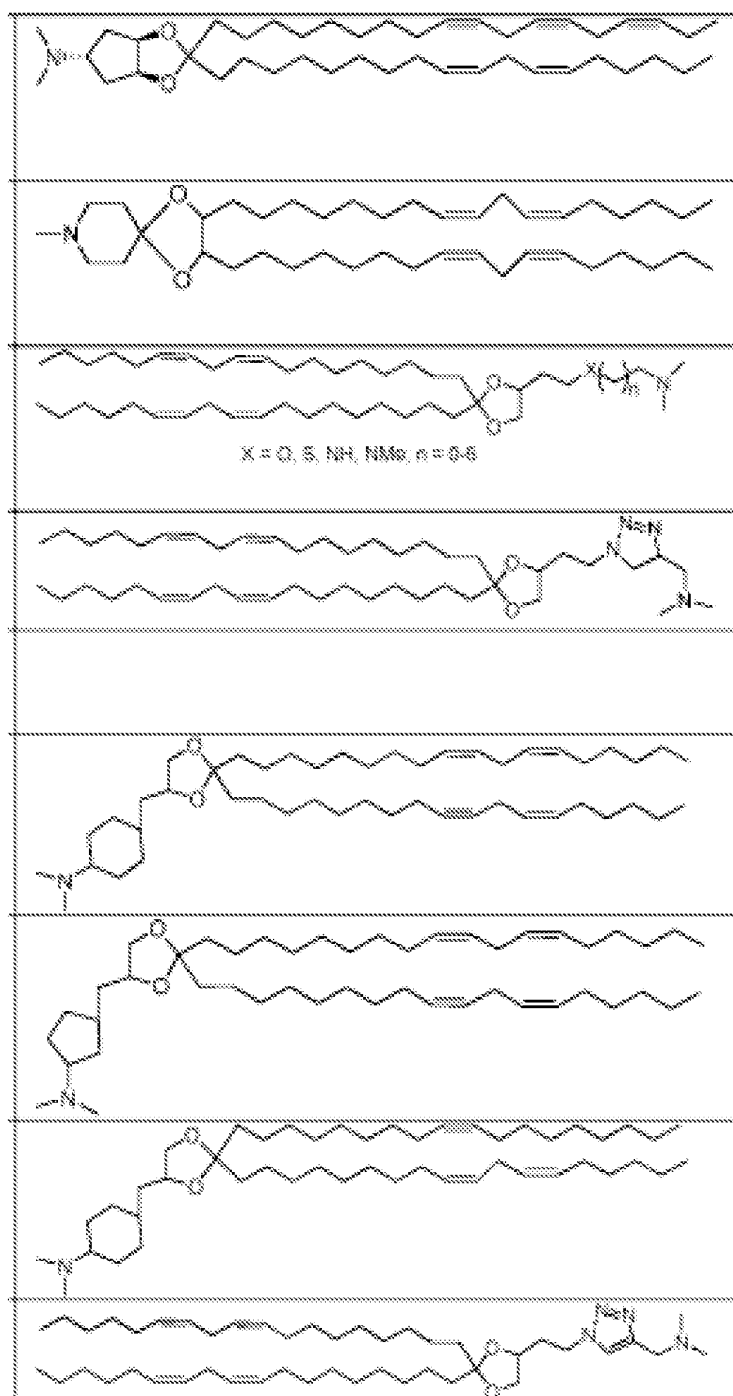


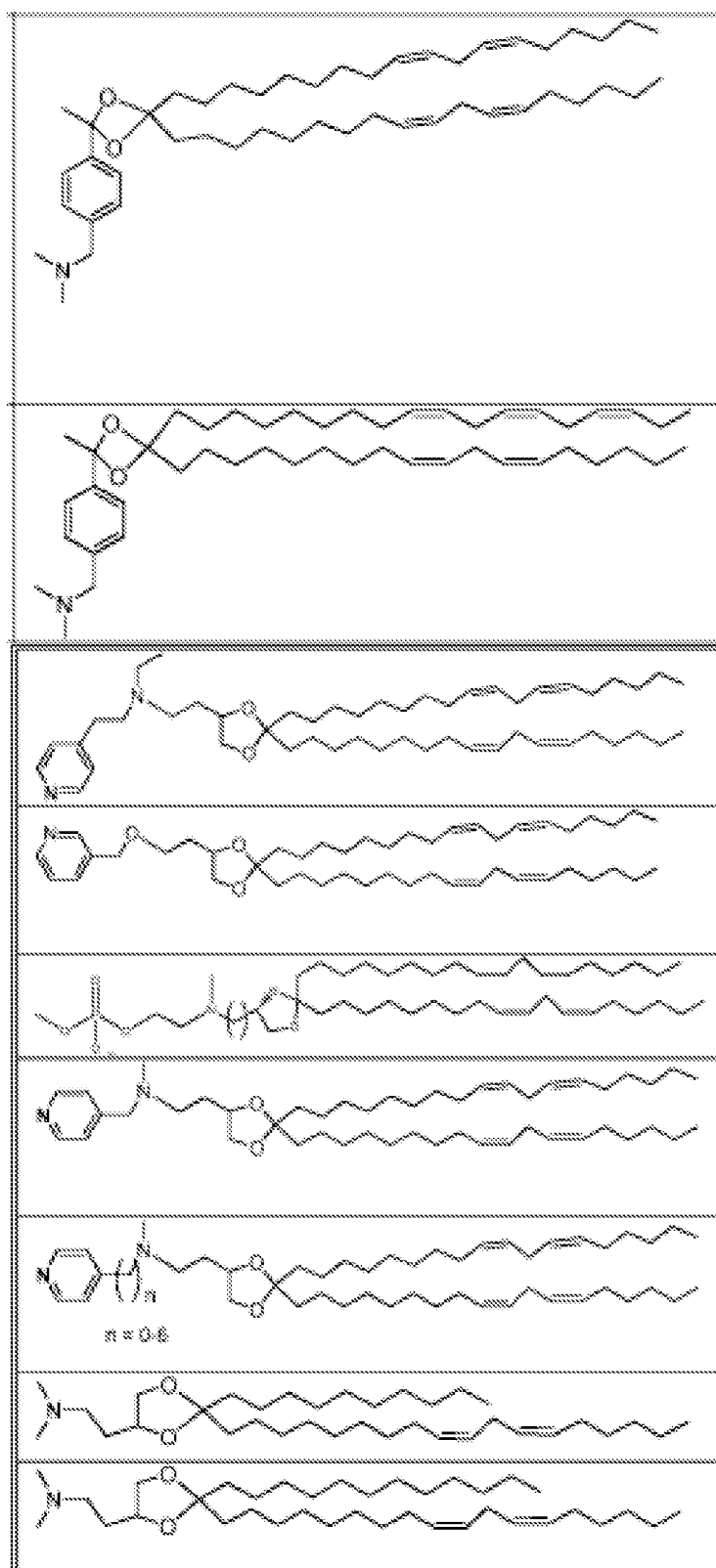


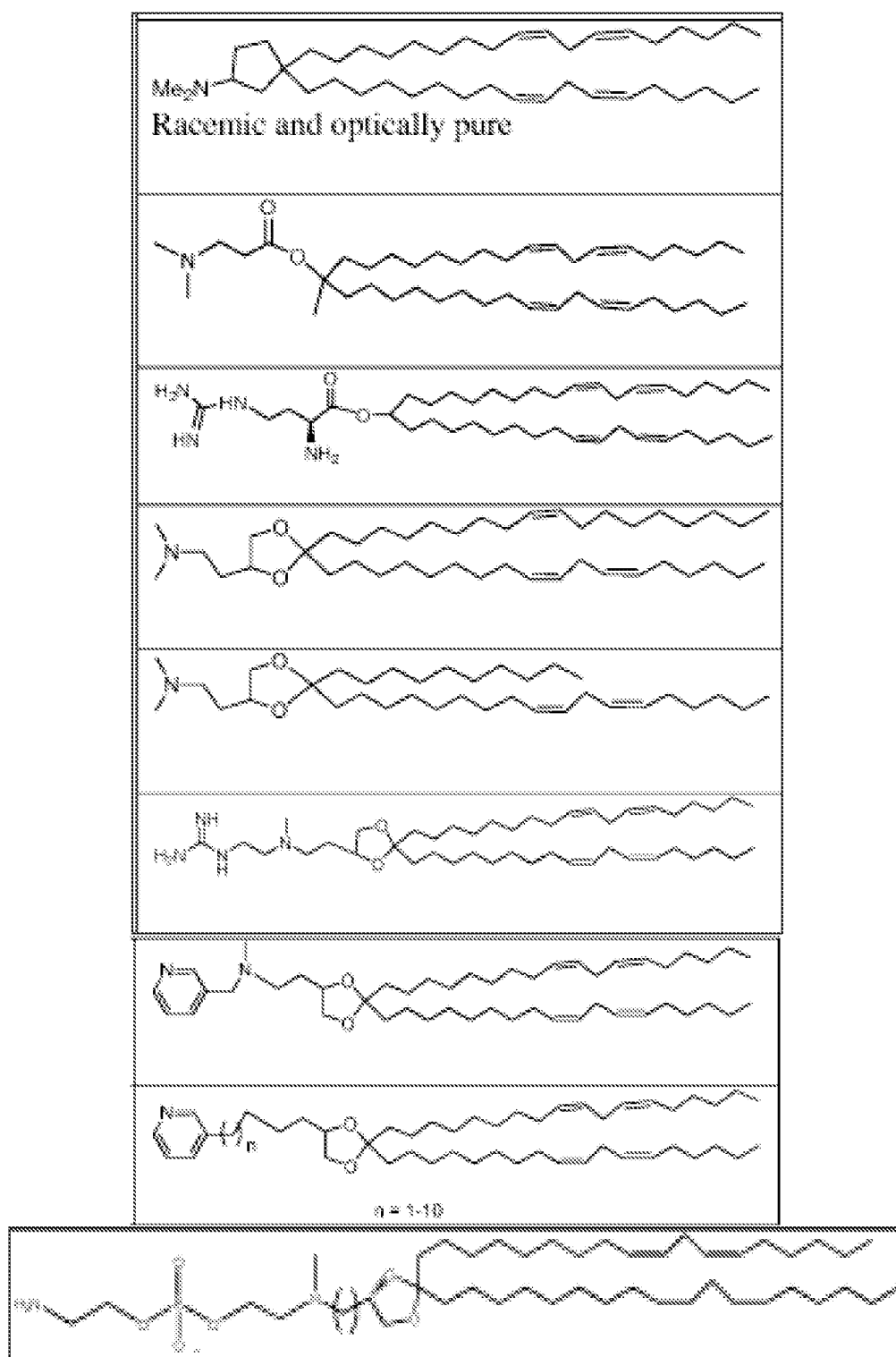


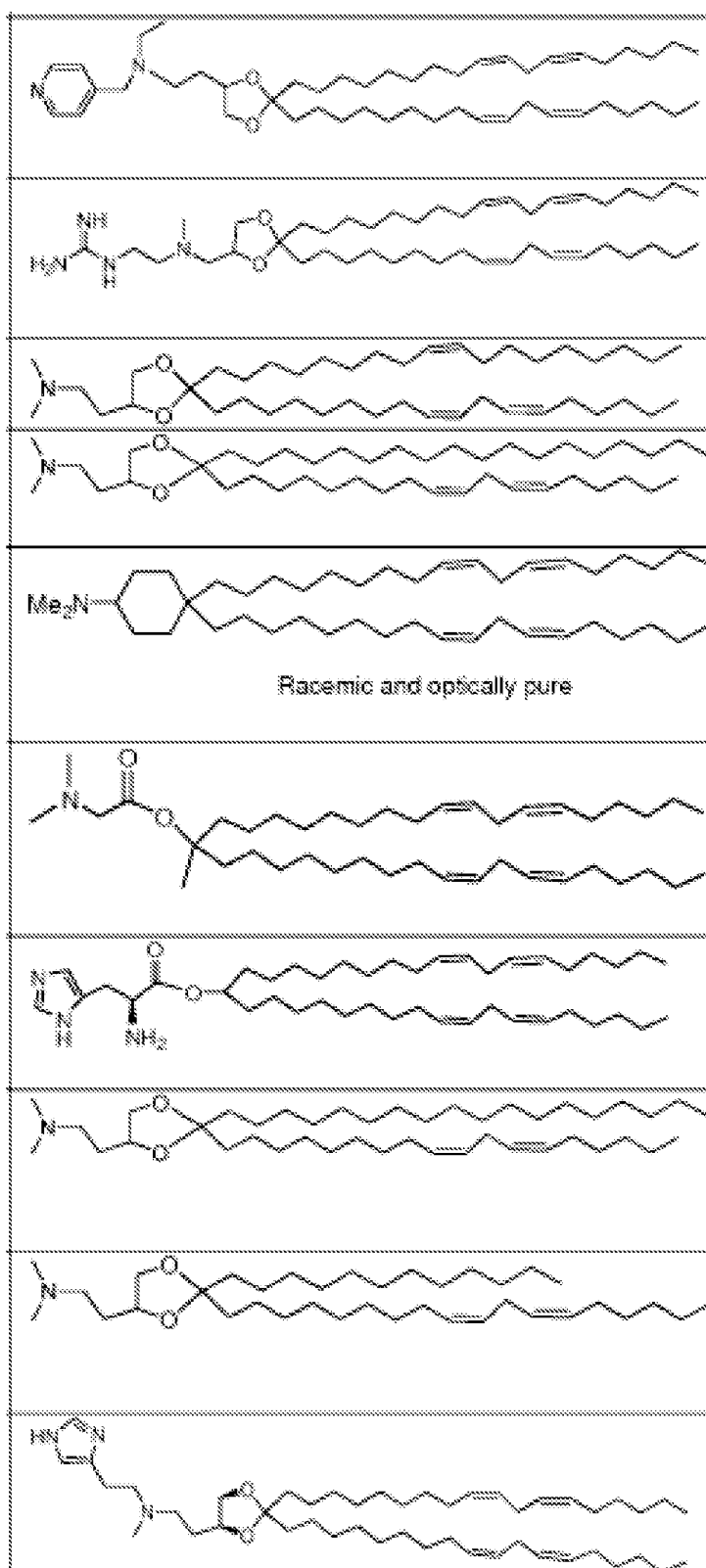


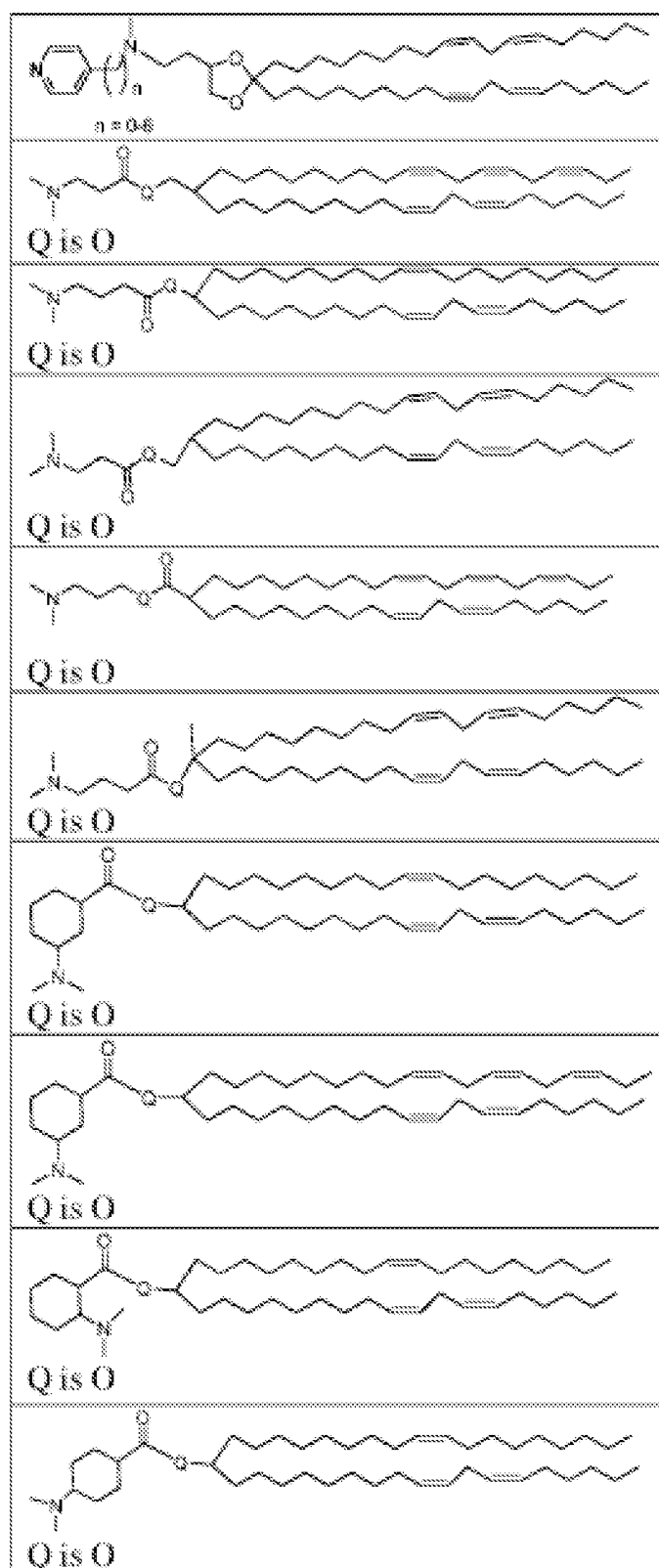


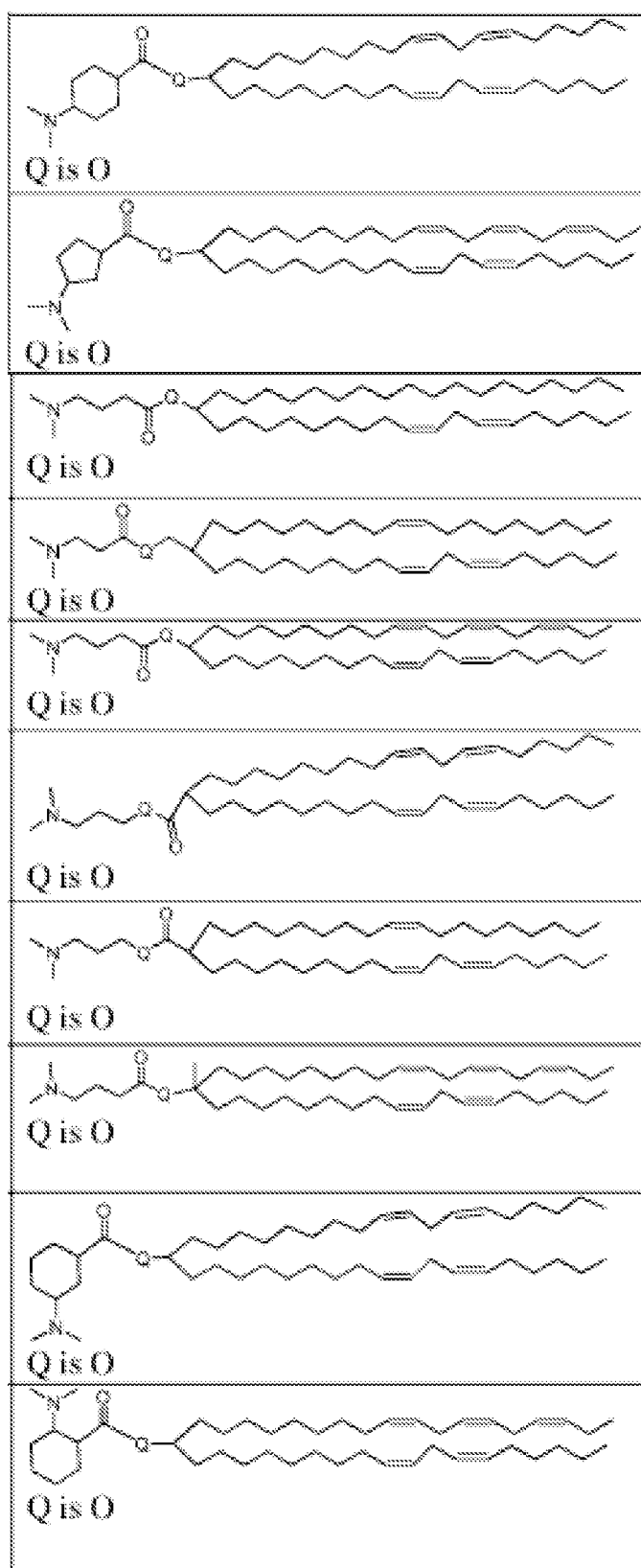


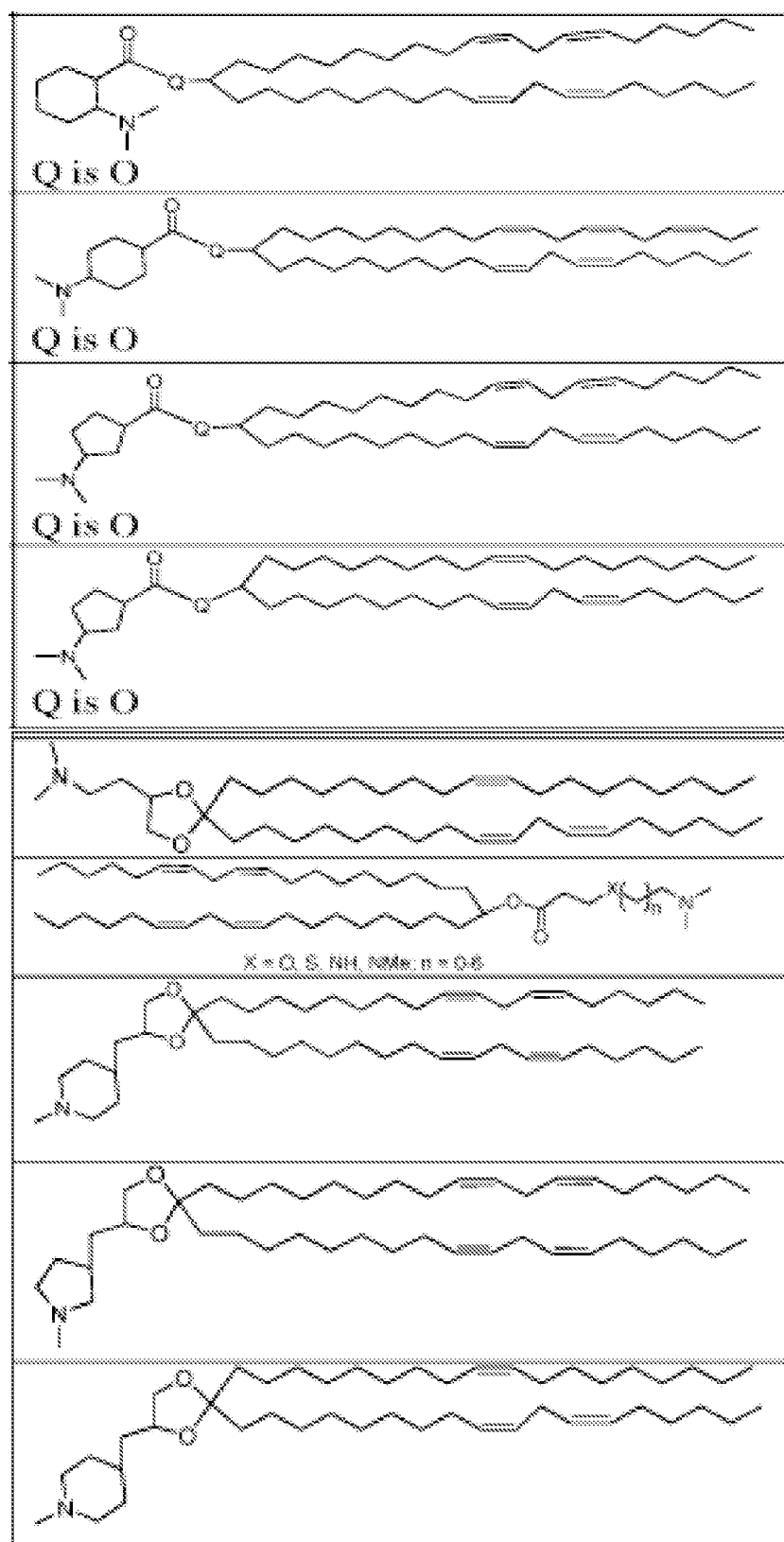


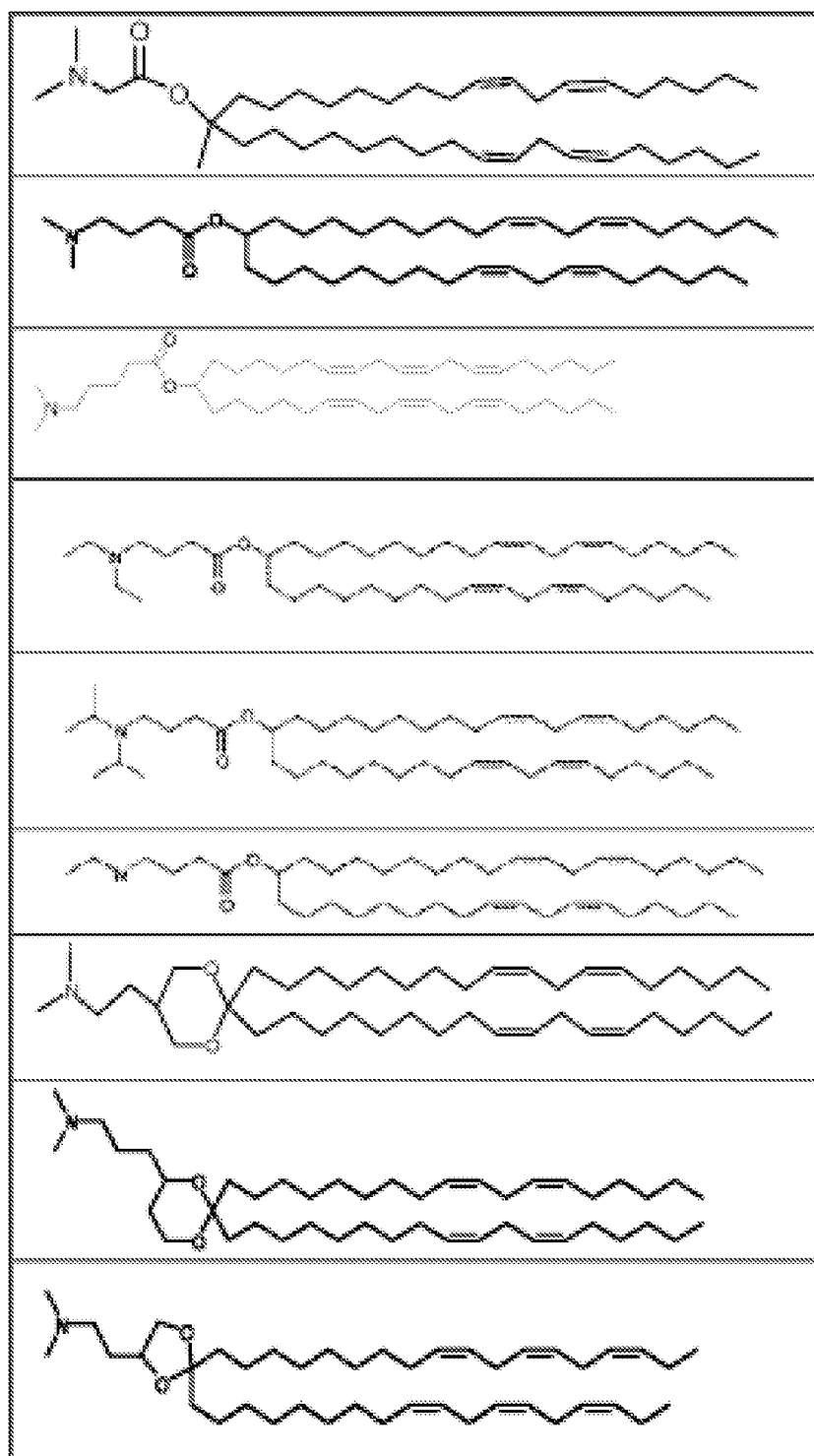


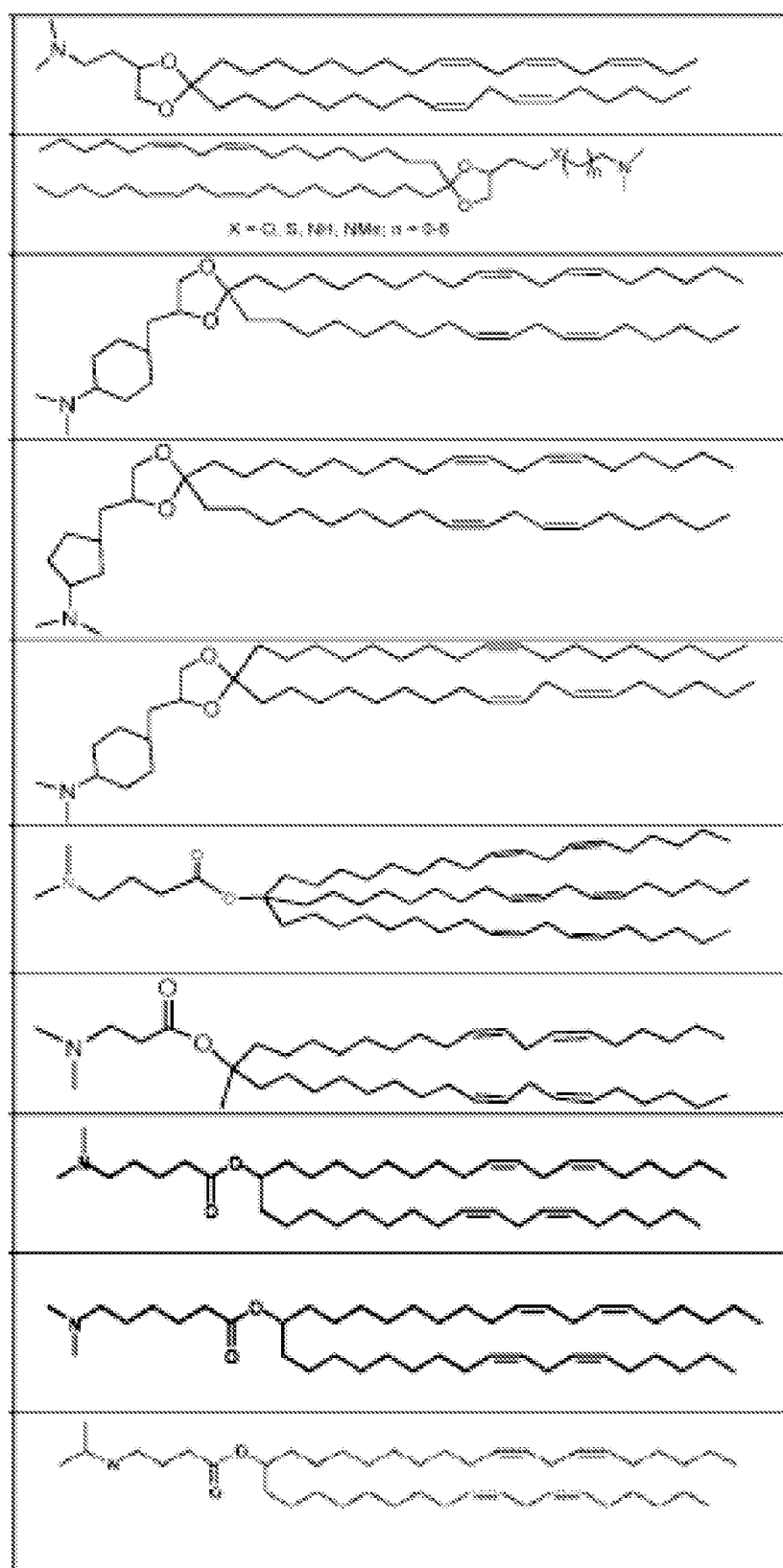


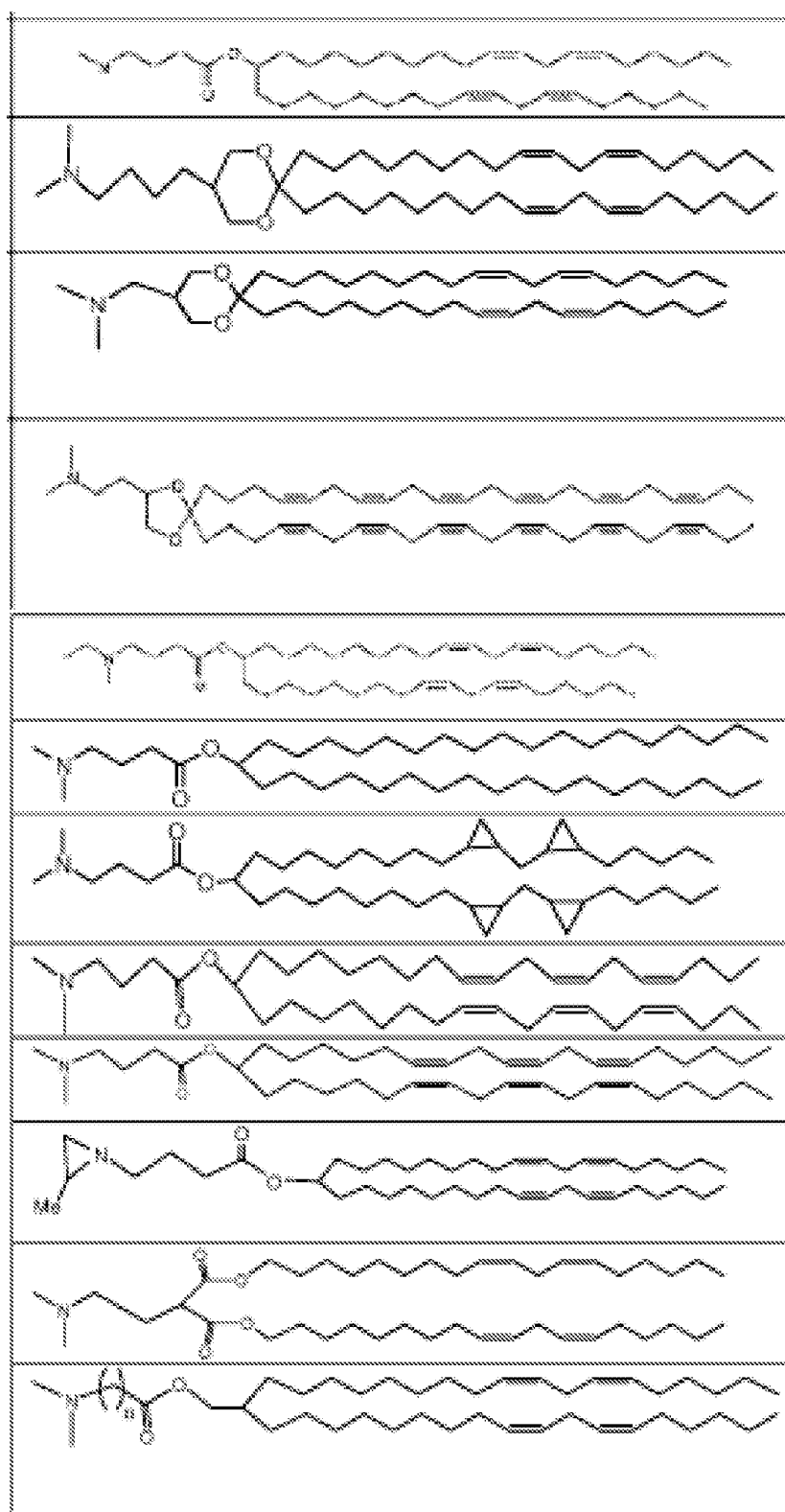


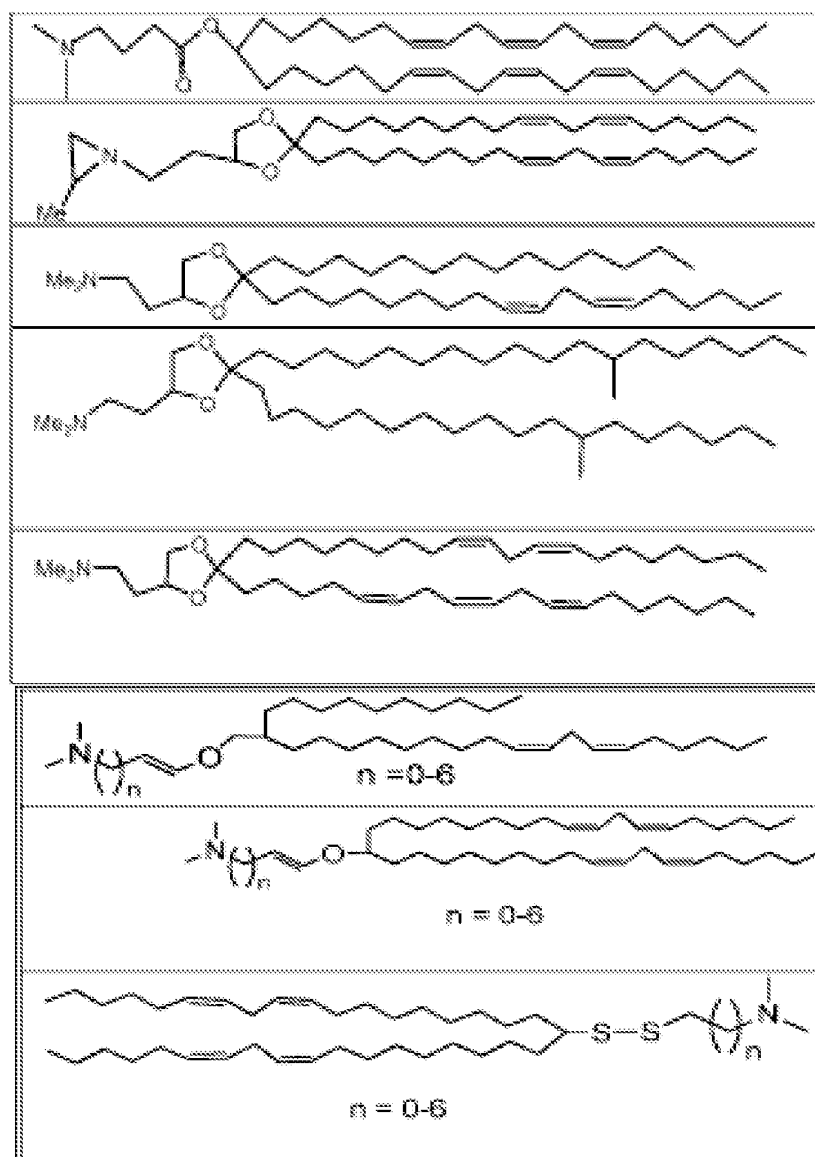


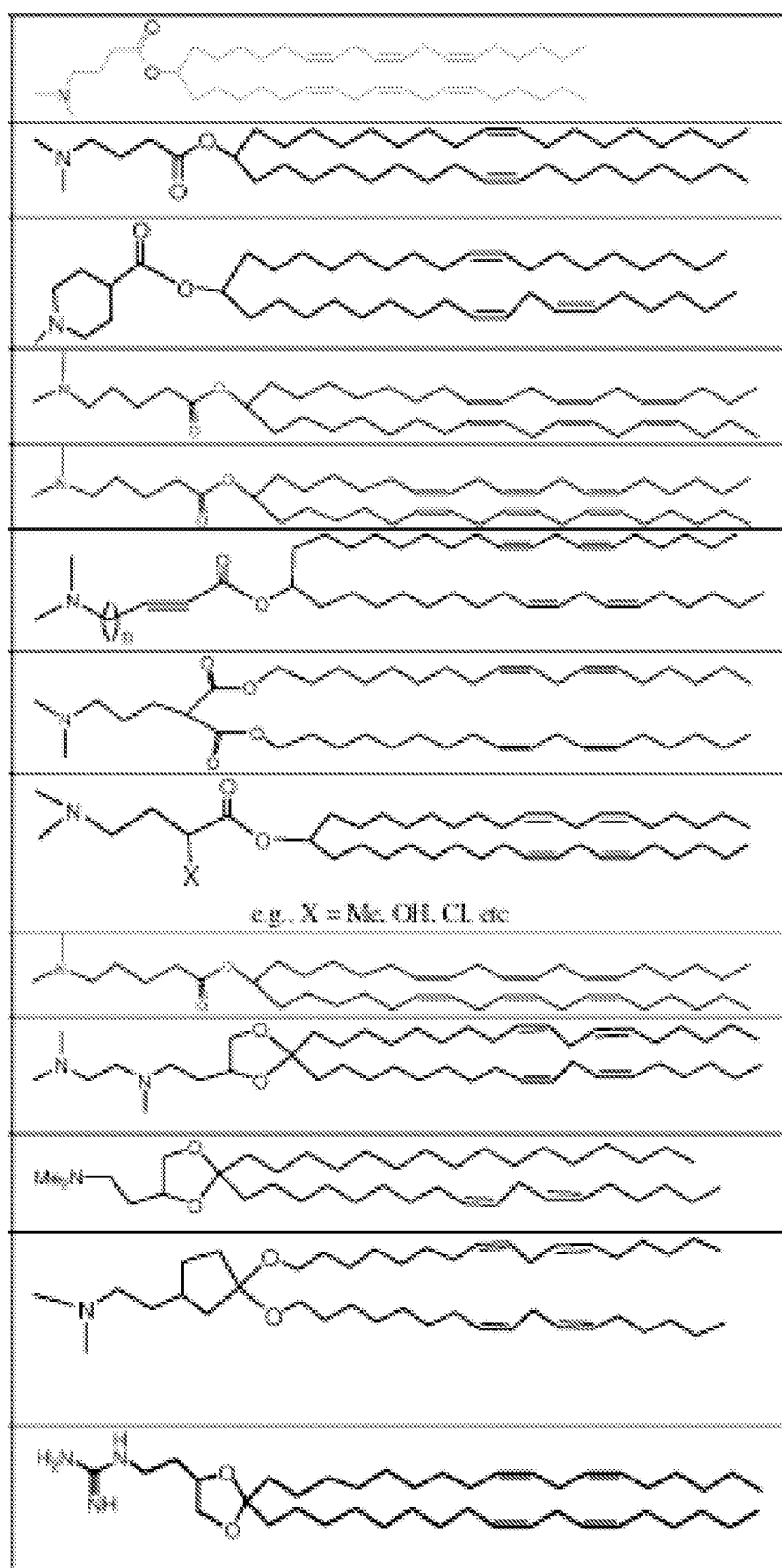


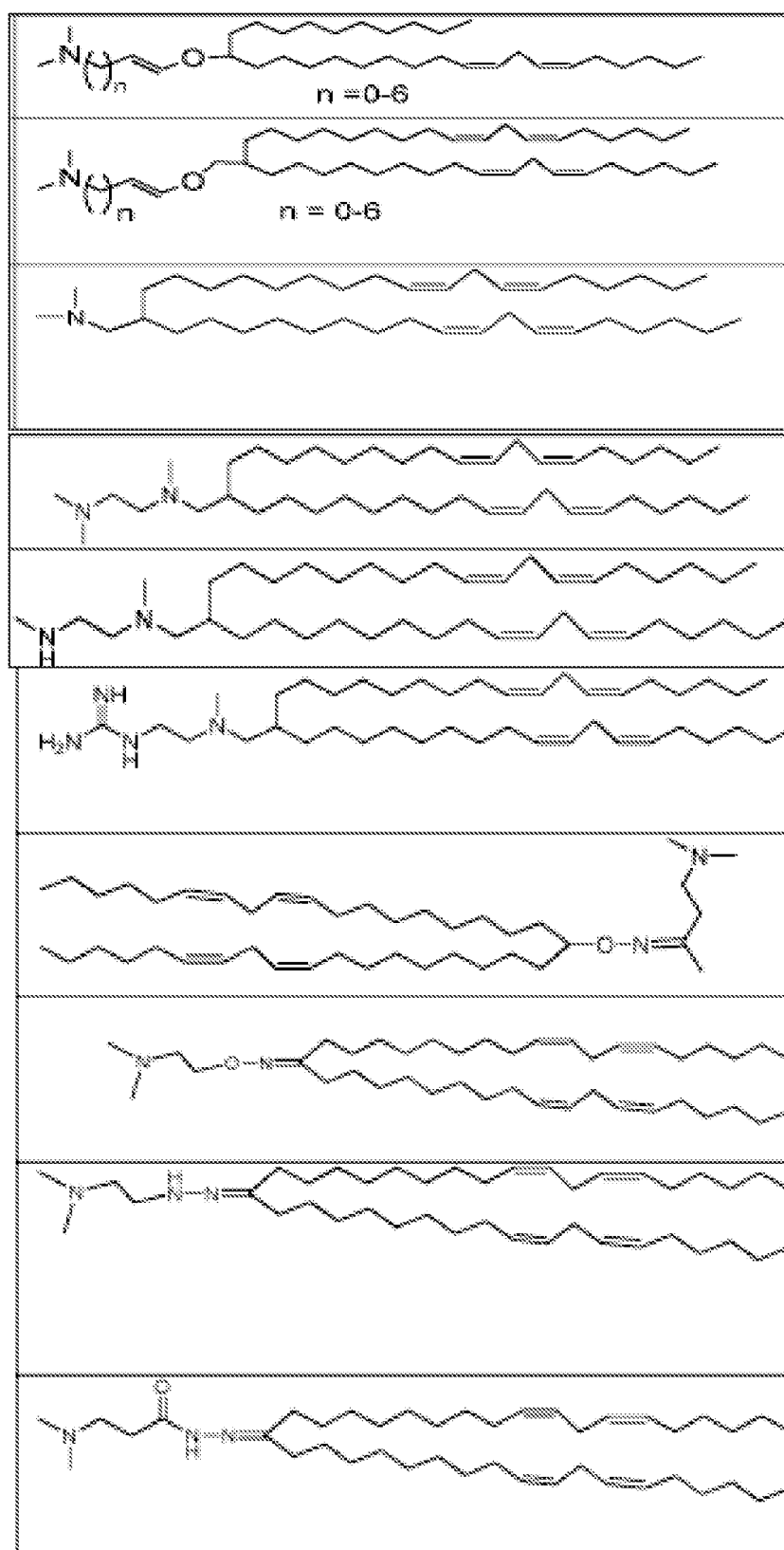


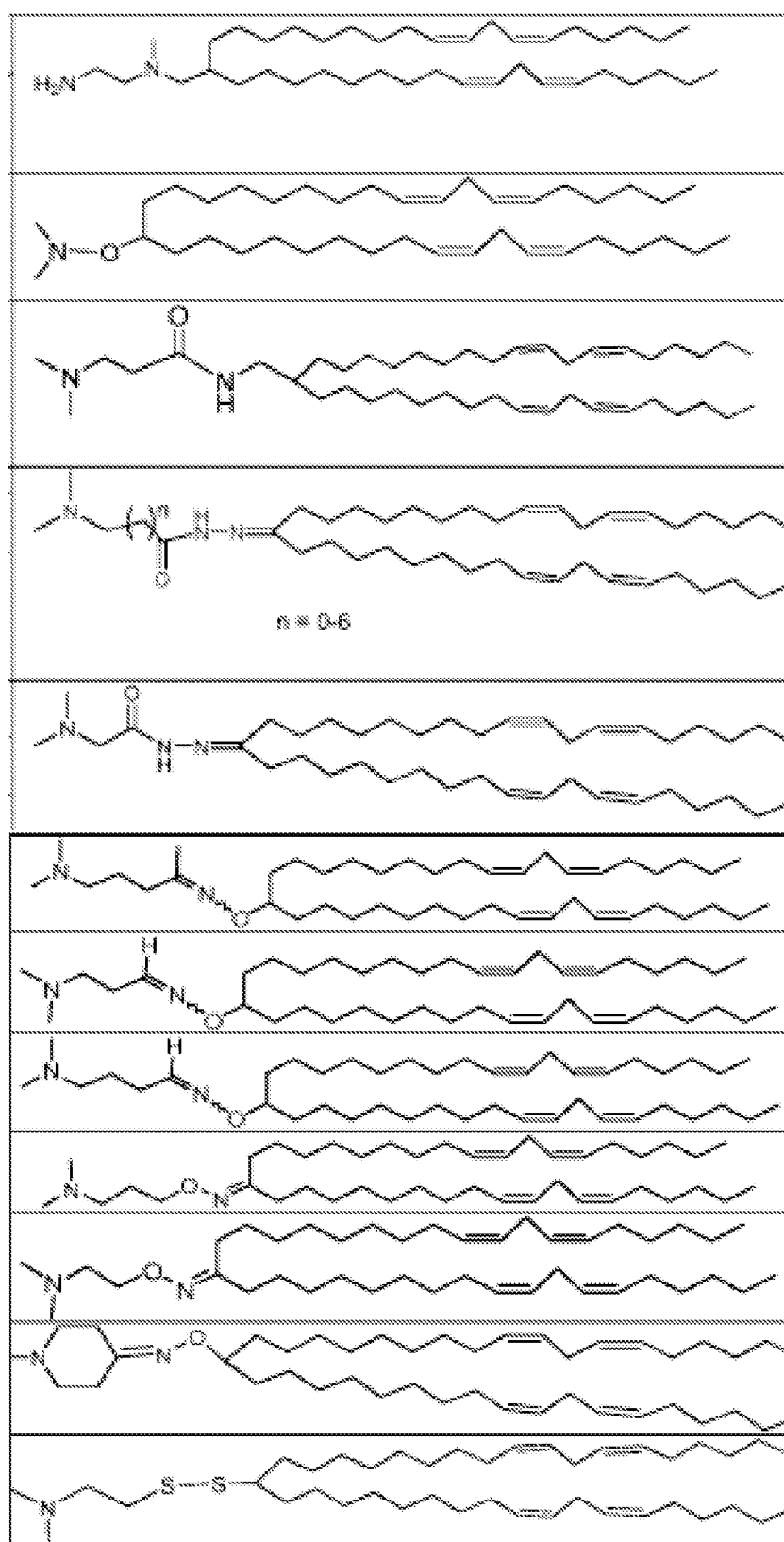


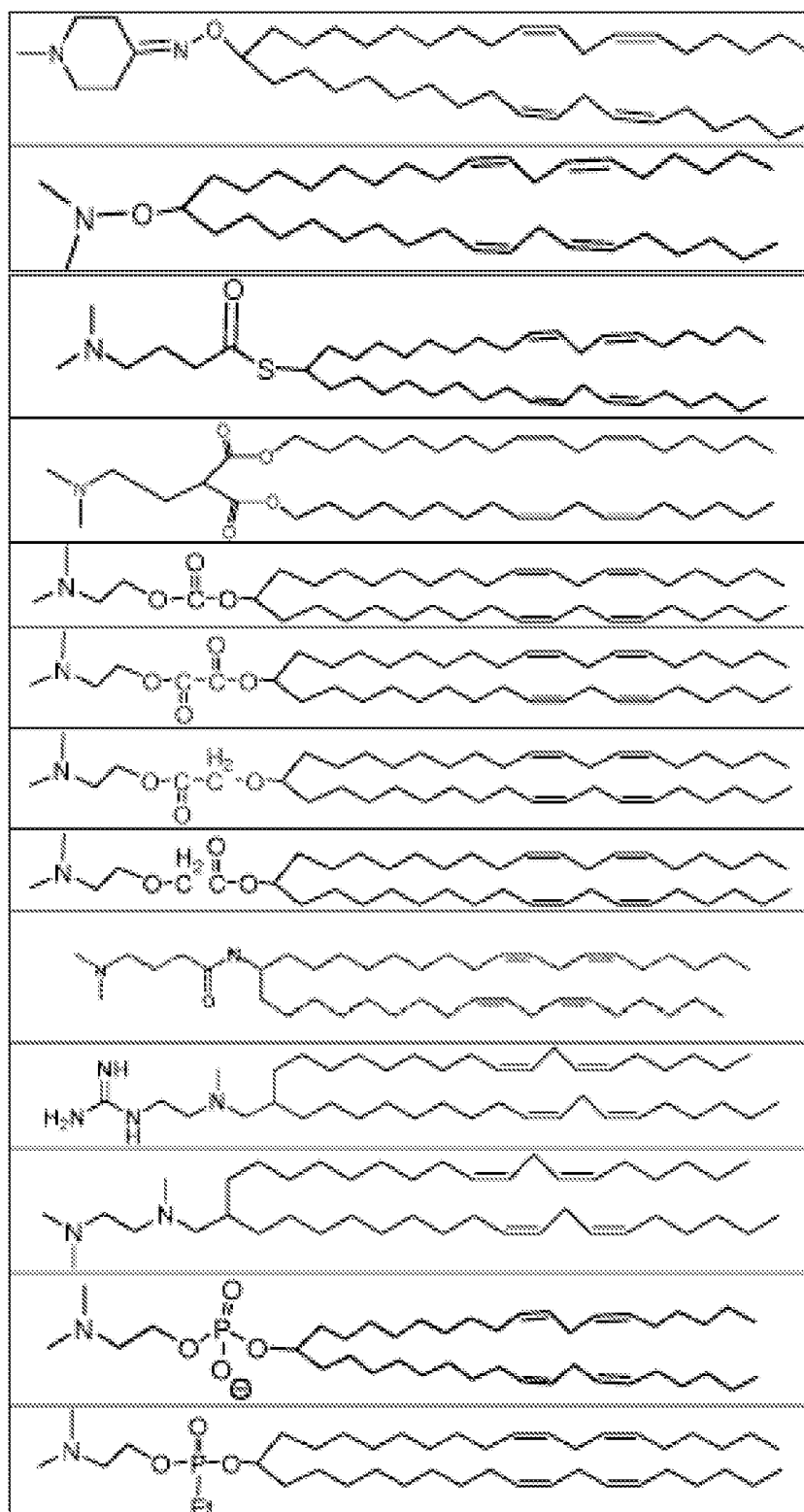


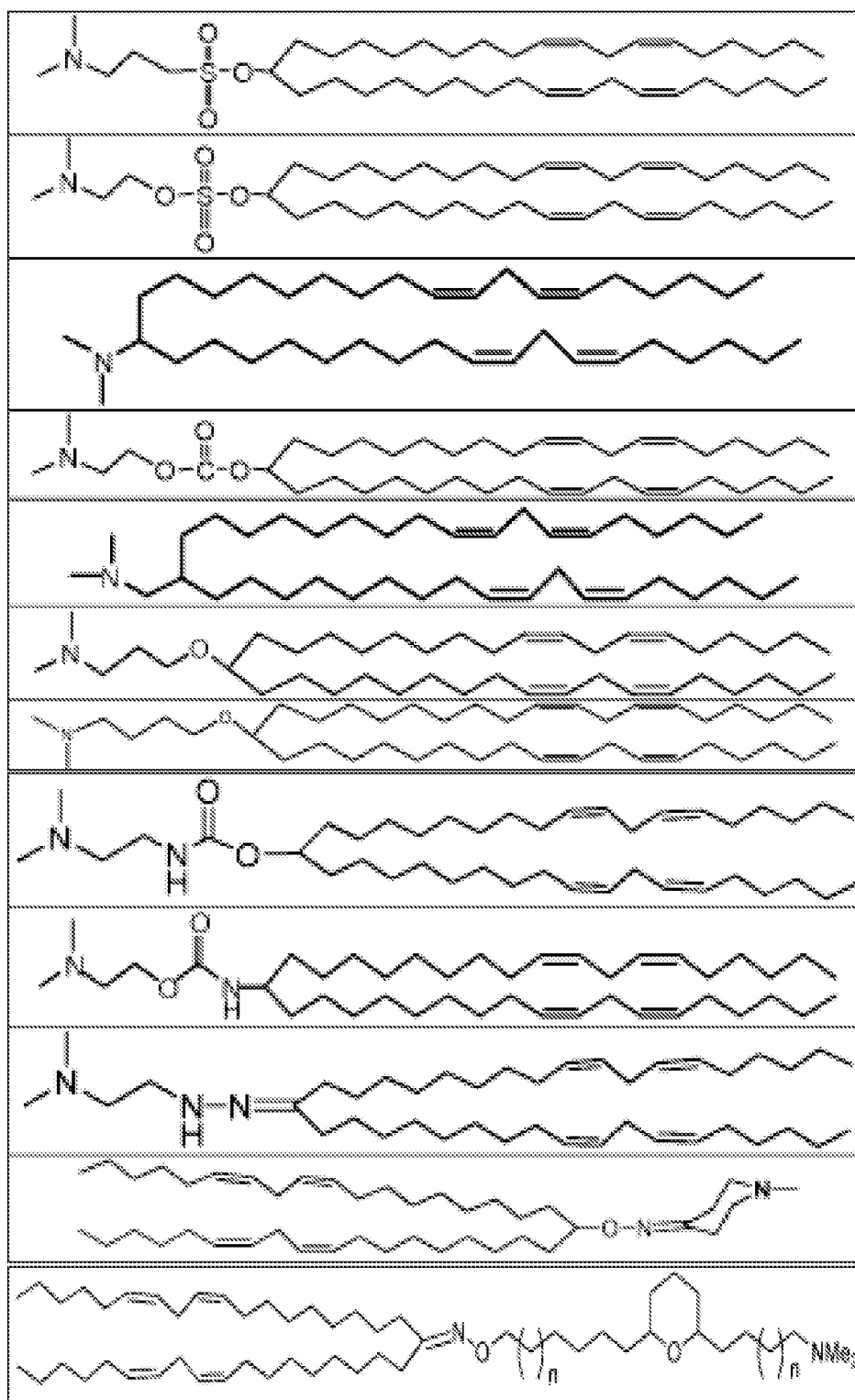


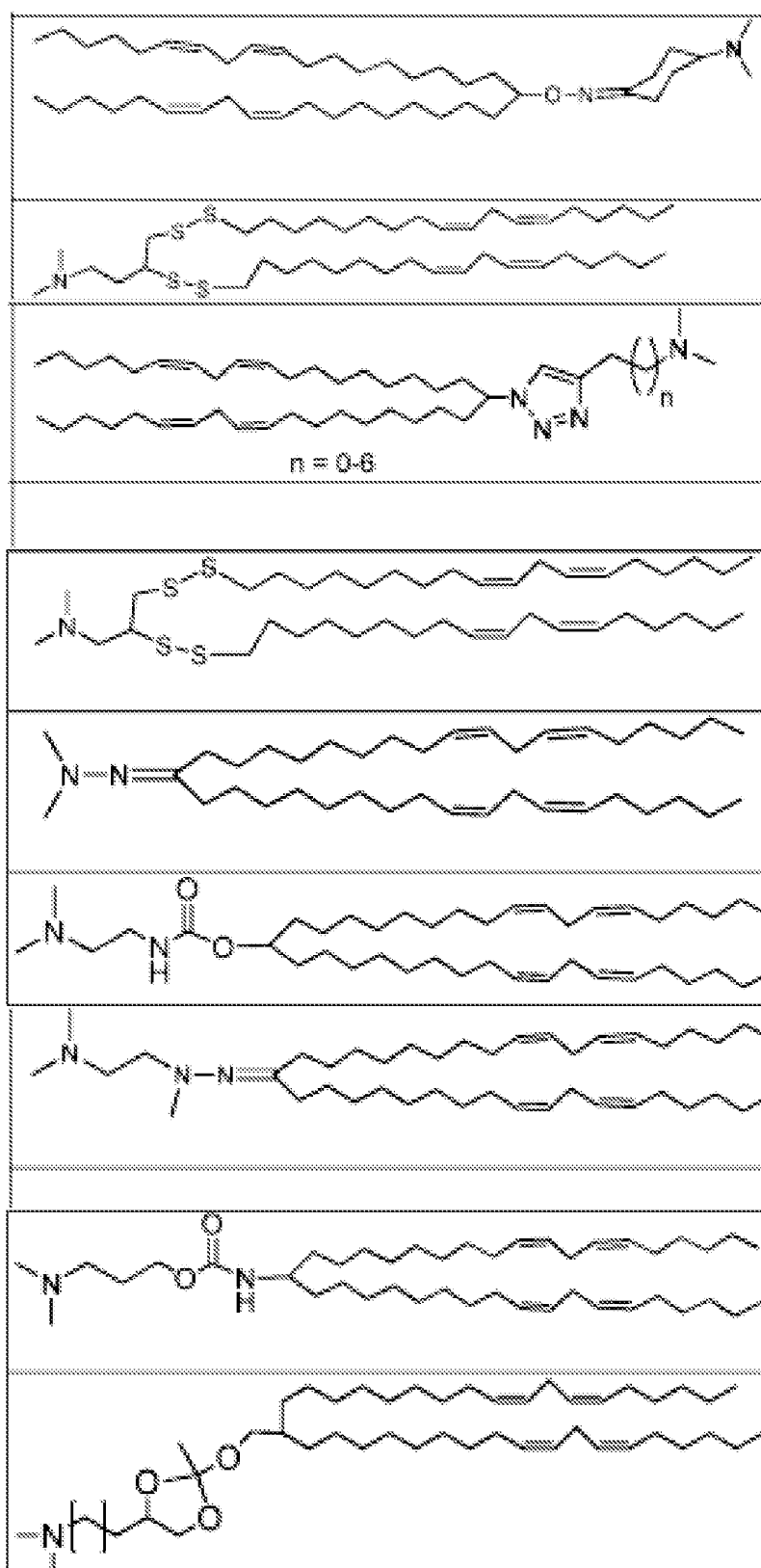


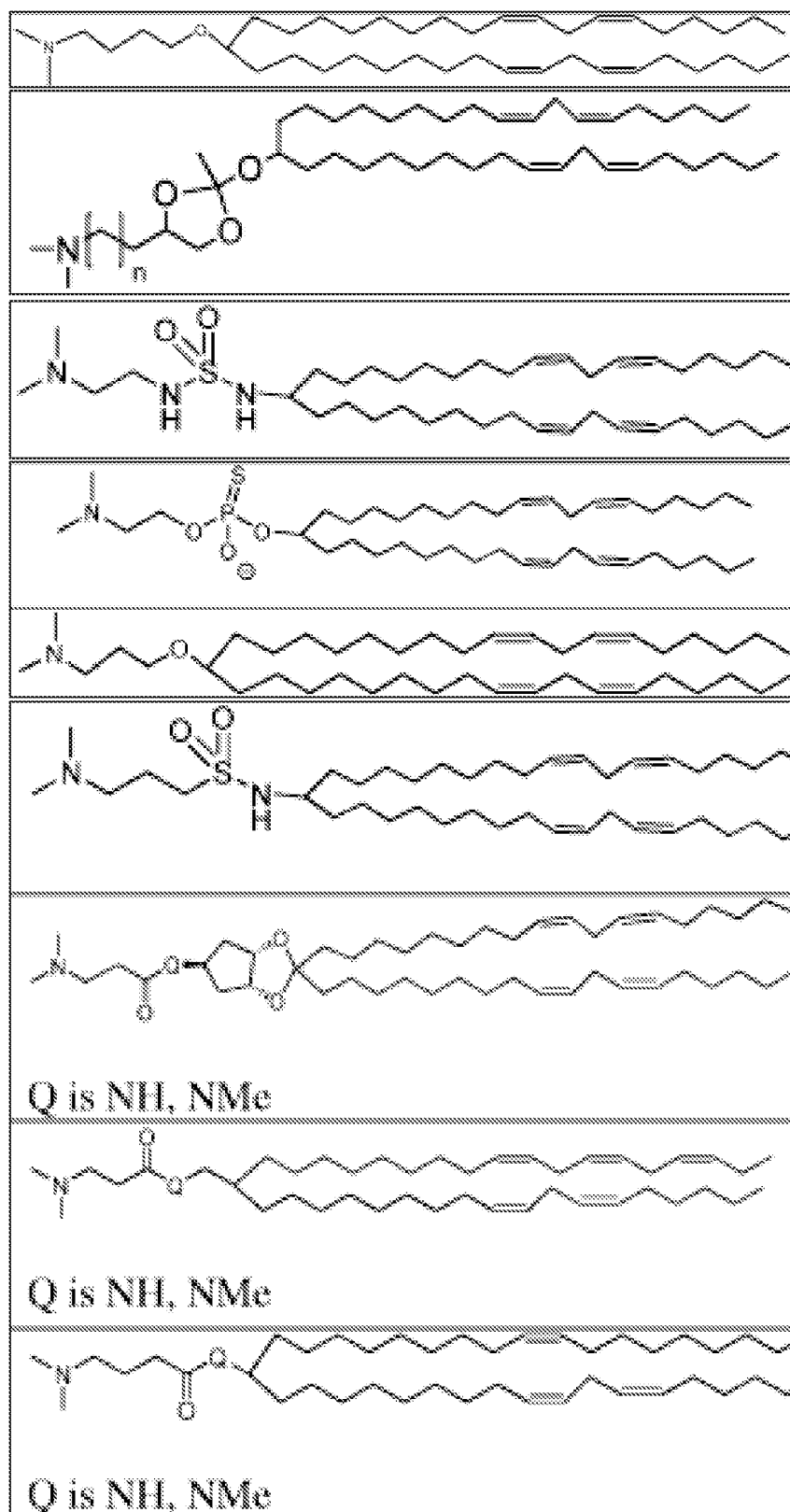


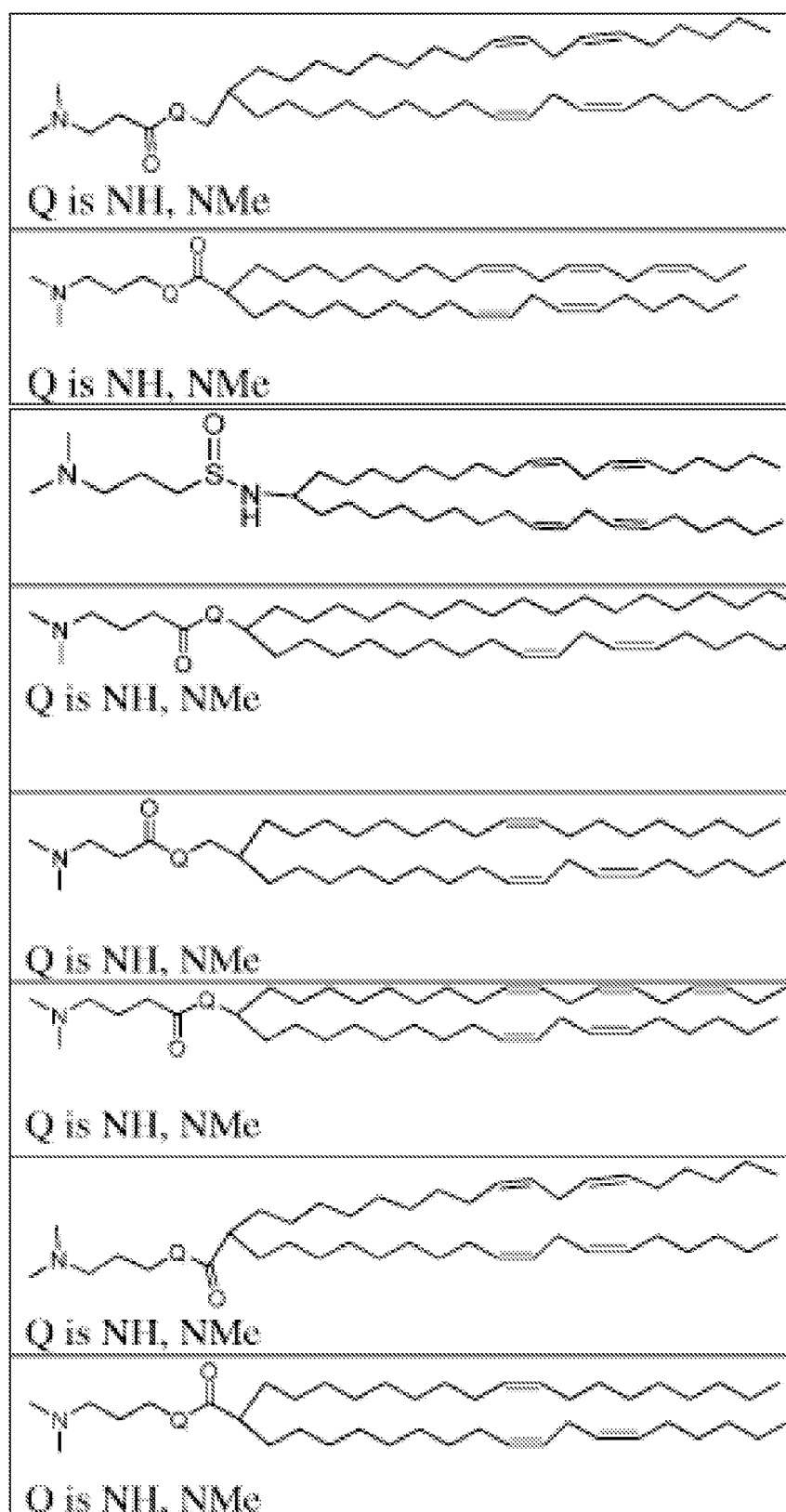


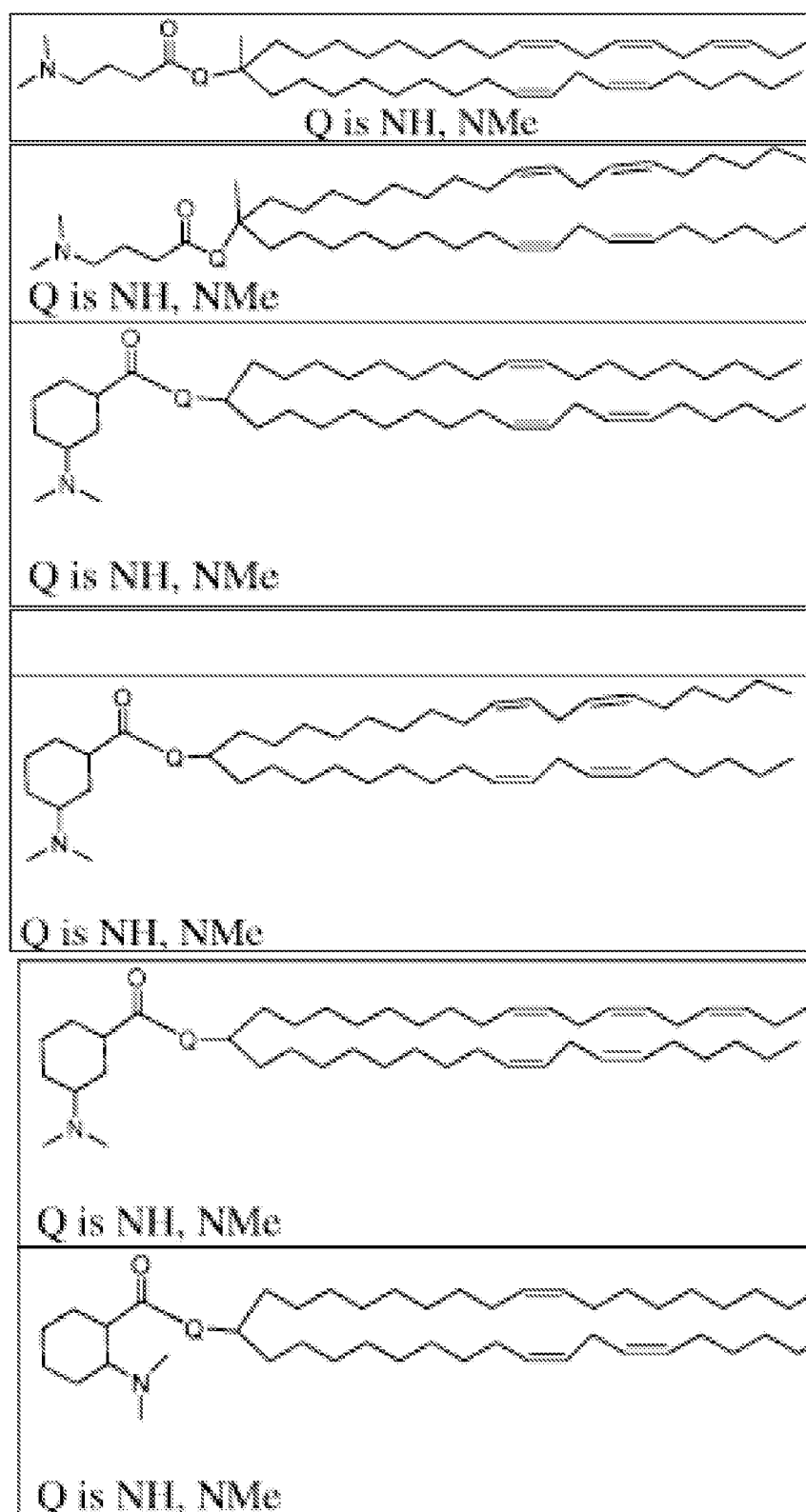


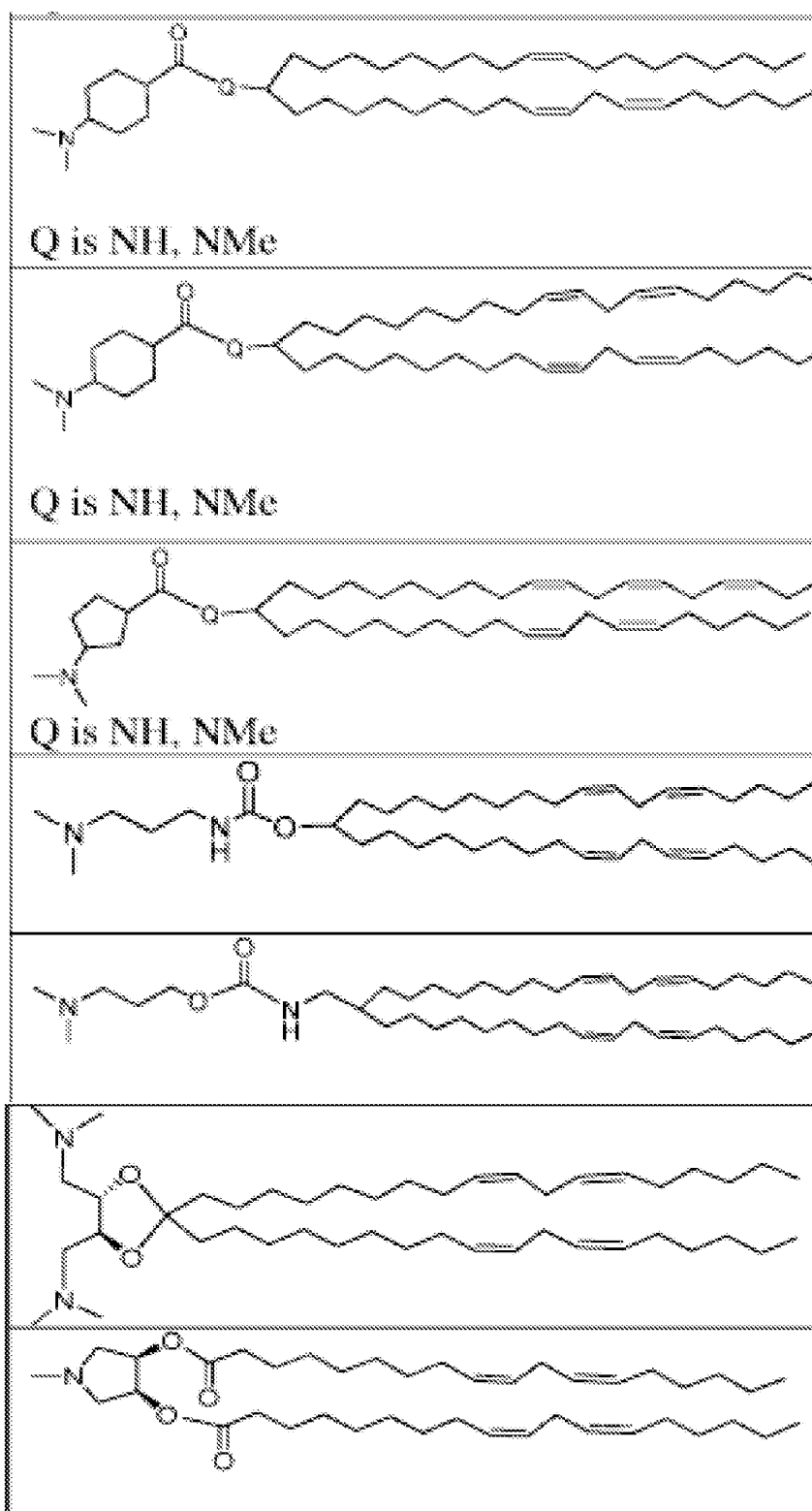


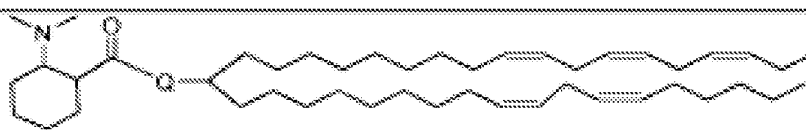




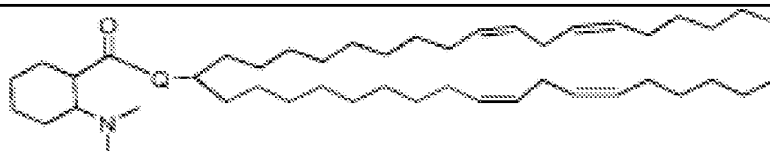




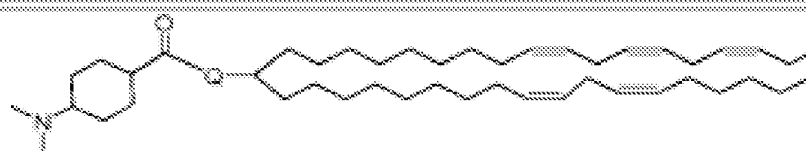




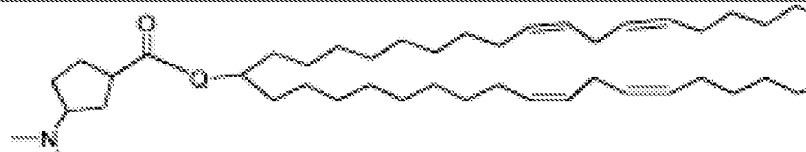
Q is NH, NMe



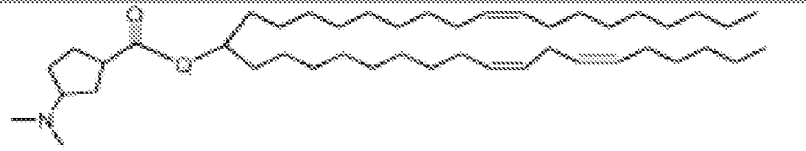
Q is NH, NMe



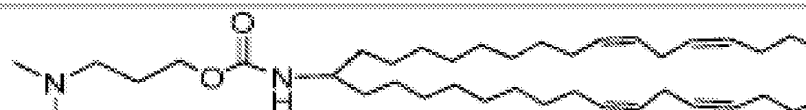
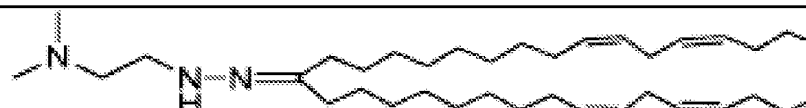
Q is NH, NMe

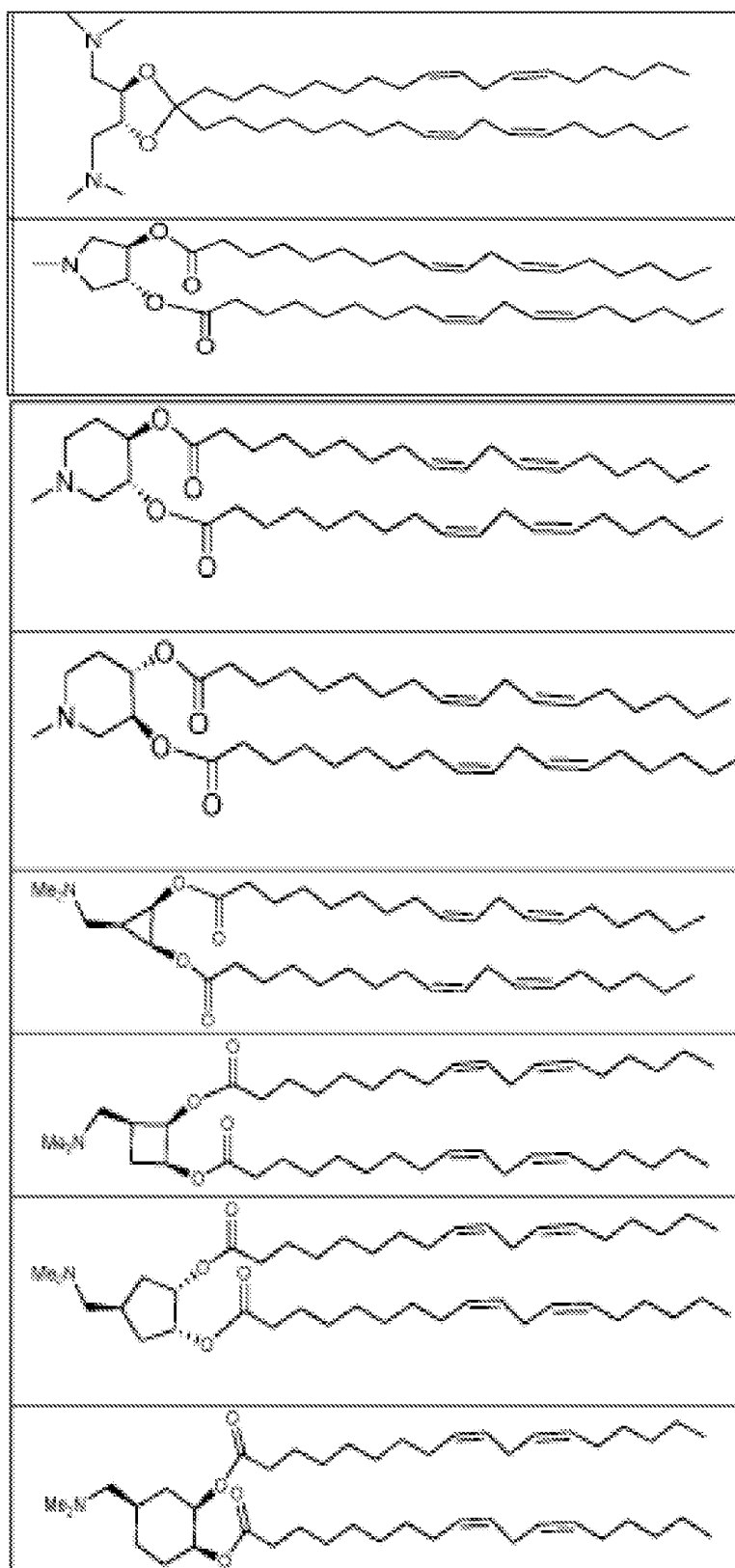


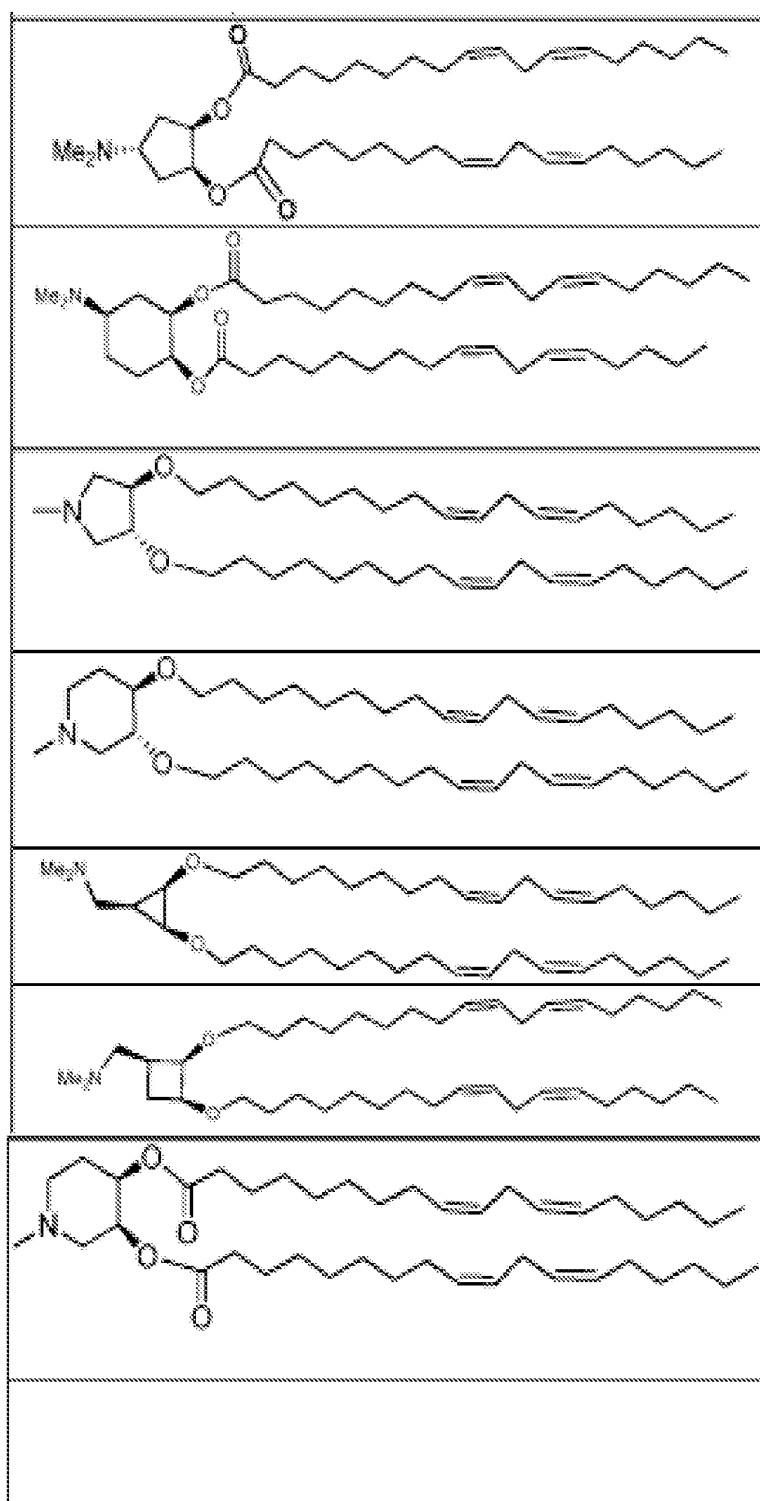
Q is NH, NMe

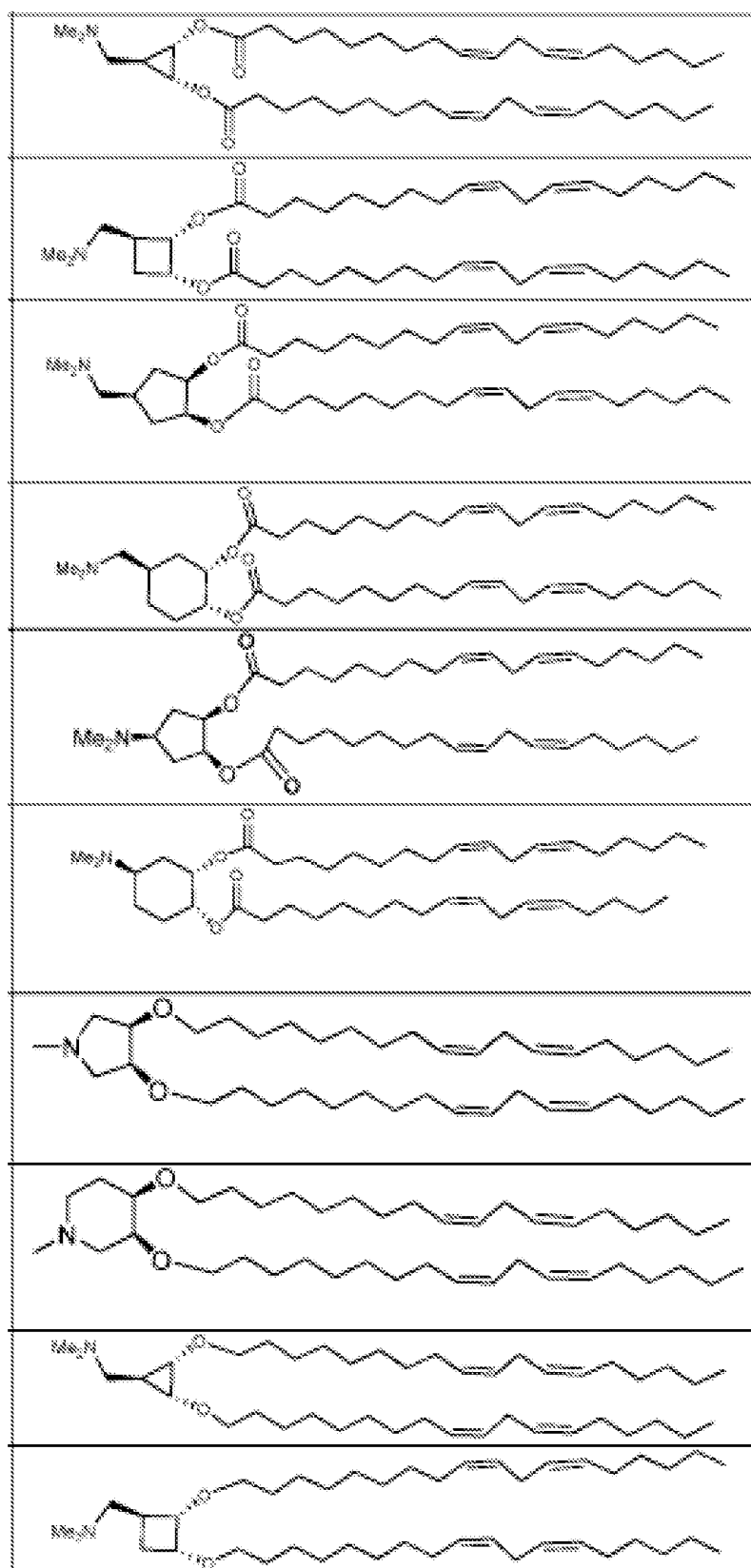


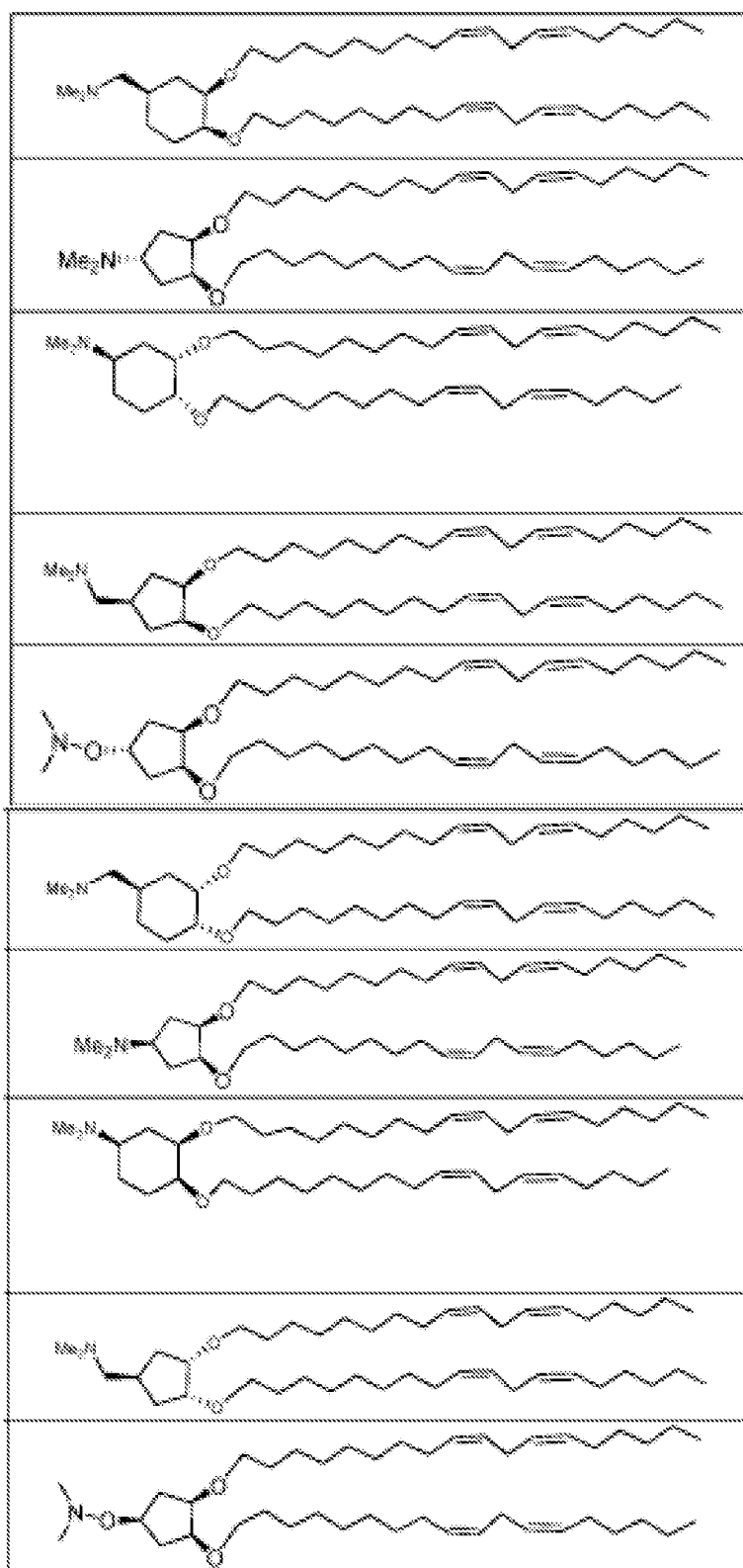
Q is NH, NMe





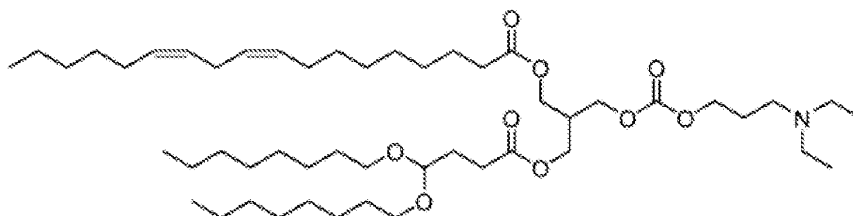






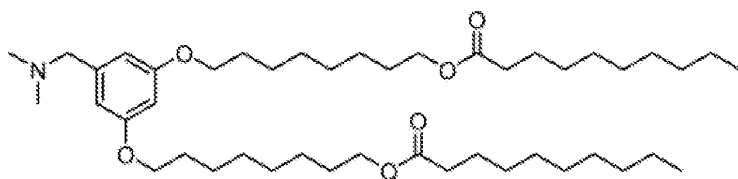
[0467] In some embodiments, the transfer vehicle comprises Lipid A, Lipid B, Lipid C, and/or Lipid D. In some embodiments, inclusion of Lipid A, Lipid B, Lipid C, and/or Lipid D improves encapsulation and/or endosomal escape. In some embodiments, Lipid A, Lipid B, Lipid C, and/or Lipid D are described in international patent application PCT/US2017/028981.

[0468] In some embodiments, an ionizable lipid is Lipid A, which is (9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl (9Z,12Z)-octadeca-9,12-dienoate. Lipid A can be depicted as:



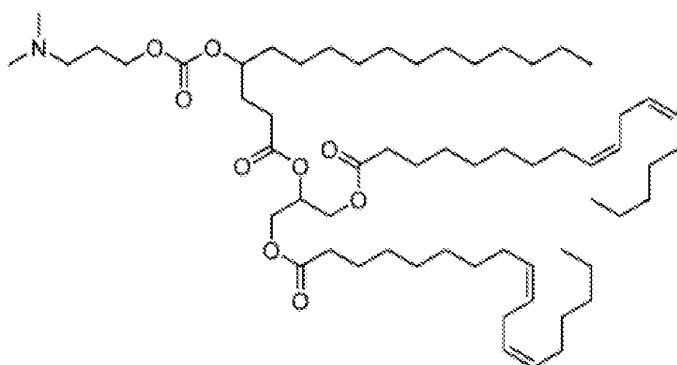
[0469] Lipid A may be synthesized according to WO2015/095340 (e.g., pp. 84-86), incorporated by reference in its entirety.

[0470] In some embodiments, an ionizable lipid is Lipid B, which is ((5-((dimethylamino)methyl)-1,3-phenylene)bis(oxy))bis(octane-8,1-diyl)bis(decanoate). Lipid B can be depicted as:



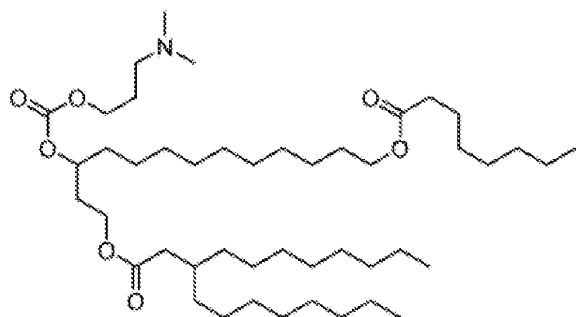
[0471] Lipid B may be synthesized according to WO2014/136086 (e.g., pp. 107-09), incorporated by reference in its entirety.

[0472] In some embodiments, an ionizable lipid is Lipid C, which is 2-(((4-(((3-(dimethylamino)propoxy)carbonyl)oxy)hexadecanoyl)oxy)propane-1,3-diyl(9Z,9'Z,12Z,12'Z)-bis(octadeca-9,12-dienoate). Lipid C can be depicted as:



[0473] In some embodiments, an ionizable lipid is Lipid D, which is 3-(((3-(dimethylamino)propoxy)carbonyl)oxy)-13-(octanoyloxy)tridecyl 3-octylundecanoate.

Lipid D can be depicted as:



[0474] Lipid C and Lipid D may be synthesized according to WO2015/095340, incorporated by reference in its entirety.

[0475] In some embodiments, an ionizable lipid is described in US patent publication number 20190321489. In some embodiments, an ionizable lipid is described in international patent publication WO 2010/053572, incorporated herein by reference. In some embodiments, an ionizable lipid is C12-200, described at paragraph [00225] of WO 2010/053572.

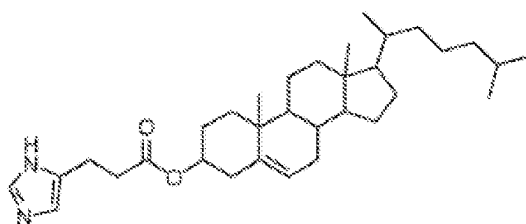
[0476] Several ionizable lipids have been described in the literature, many of which are commercially available. In certain embodiments, such ionizable lipids are included in the transfer vehicles described herein. In some embodiments, the ionizable lipid N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium chloride or "DOTMA" is used. (Felgner *et al.* Proc. Nat'l Acad. Sci. 84, 7413 (1987); U.S. Pat. No. 4,897,355). DOTMA can be formulated alone or can be combined with a neutral lipid, dioleoylphosphatidylethanolamine or "DOPE" or other cationic or non-cationic lipids into a lipid nanoparticle. Other suitable cationic lipids include, for example, ionizable cationic lipids as described in U.S. provisional patent application 61/617,468, filed Mar. 29, 2012 (incorporated herein by reference), such

as, e.g., (15Z,18Z)-N,N-dimethyl-6-(9Z,12Z)-octadeca-9,12-dien-1-yl)tetracos-15,18-dien-1-amine (HGT5000), (15Z,18Z)-N,N-dimethyl-6-((9Z,12Z)-octadeca-9,12-dien-1-yl)tetracos-4,15,18-trien-1-amine (HGT5001), and (15Z,18Z)-N,N-dimethyl-6-((9Z,12Z)-octadeca-9,12-dien-1-yl)tetracos-5,15,18-trien-1-amine (HGT5002), C12-200 (described in WO 2010/053572), 2-(2,2-di((9Z,12Z)-octadeca-9,12-dien-1-yl)-1,3-dioxolan-4-yl)-N,N-dimethylethanamine (DlinKC2-DMA)) (See, WO 2010/042877; Semple *et al.*, Nature Biotech. 28:172-176 (2010)), 2-(2,2-di((9Z,2Z)-octadeca-9,12-dien-1-yl)-1,3-dioxolan-4-yl)-N,N-dimethylethanamine (Dlin-KC2-DMA), (3S,10R,13R,17R)-10,13-dimethyl-17-((R)-6-methylheptan-2-yl)-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-yl 3-(1H-imidazol-4-yl)propanoate (ICE), (15Z,18Z)-N,N-dimethyl-6-(9Z,12Z)-octadeca-9,12-dien-1-yl)tetracos-15,18-dien-1-amine (HGT5000), (15Z,18Z)-N,N-dimethyl-6-((9Z,12Z)-octadeca-9,12-dien-1-yl)tetracos-4,15,18-trien-1-amine (HGT5001), (15Z,18Z)-N,N-dimethyl-6-((9Z,12Z)-octadeca-9,12-dien-1-yl)tetracos-5,15,18-trien-1-amine (HGT5002), 5-carboxyspermylglycine-dioctadecylamide (DOGS), 2,3-dioleoyloxy-N-[2(spermine-carboxamido)ethyl]-N,N-dimethyl-1-propanaminium (DOSPA) (Behr *et al.* Proc. Nat'l Acad. Sci. 86, 6982 (1989); U.S. Pat. No. 5,171,678; 5,334,761), 1,2-Dioleoyl-3-Dimethylammonium-Propane (DODAP), 1,2-Dioleoyl-3-Trimethylammonium-Propane or (DOTAP). Contemplated ionizable lipids also include 1,2-distearoyloxy-N,N-dimethyl-3-aminopropane (DSDMA), 1,2-dioleoyloxy-N,N-dimethyl-3-aminopropane (DODMA), 1,2-dilinoleoyloxy-N,N-dimethyl-3-aminopropane (DlinDMA), 1,2-dilinolenyloxy-N,N-dimethyl-3-aminopropane (DLenDMA), N-dioleoyl-N,N-dimethylammonium chloride (DODAC), N,N-distearyl-N,N-dimethylammonium bromide (DDAB), N-(1,2-dimyristyloxyprop-3-yl)-N,N-dimethyl-N-hydroxyethyl ammonium bromide (DMRIE), 3-dimethylamino-2-(cholest-5-en-3-beta-oxybutan-4-oxy)-1-(cis,cis-9,12-octadecadienoxy)propane (ClinDMA), 2-[5'-(cholest-5-en-3-beta-oxy)-3'-oxapentoxyl]-3-dimethyl-1-(cis,cis-9',12'-octadecadienoxy)propane (CpLinDMA), N,N-dimethyl-3,4-dioleoyloxybenzylamine (DMOBA), 1,2-N,N'-dioleoylcarbonyl-3-dimethylaminopropane (DOcarbDAP), 2,3-Dilinoleoyloxy-N,N-dimethylpropylamine (DlinDAP), 1,2-N,N'-Dilinoleoylcarbonyl-3-dimethylaminopropane (DlincarbDAP), 1,2-Dilinoleoylcarbonyl-3-dimethylaminopropane (DlinCDAP), 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (Dlin-K-DMA), 2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (Dlin-K-XTC2-DMA) or GL67, or mixtures thereof. (Heyes, J., *et al.*, J Controlled Release 107: 276-287 (2005); Morrissey, D V., *et al.*, Nat. Biotechnol. 23(8): 1003-1007 (2005); PCT Publication

WO2005/121348A1). The use of cholesterol-based ionizable lipids to formulate the transfer vehicles (*e.g.*, lipid nanoparticles) is also contemplated by the present invention. Such cholesterol-based ionizable lipids can be used, either alone or in combination with other lipids. Suitable cholesterol-based ionizable lipids include, for example, DC-Cholesterol (N,N-dimethyl-N-ethylcarboxamidocholesterol), and 1,4-bis(3-N-oleylamino-propyl)piperazine (Gao, *et al.*, Biochem. Biophys. Res. Comm. 179, 280 (1991); Wolf *et al.* BioTechniques 23, 139 (1997); U.S. Pat. No. 5,744,335).

[0477] Also contemplated are cationic lipids such as dialkylamino-based, imidazole-based, and guanidinium-based lipids. For example, also contemplated is the use of the ionizable lipid (3S,10R, 13R, 17R)-10,13-dimethyl-17-((R)-6-methylheptan-2-yl)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-yl 3-(1H-imidazol-4-yl)propanoate (ICE), as disclosed in International Application No. PCT/US2010/058457, incorporated herein by reference.

[0478] Also contemplated are ionizable lipids such as the dialkylamino-based, imidazole-based, and guanidinium-based lipids. For example, certain embodiments are directed to a composition comprising one or more imidazole-based ionizable lipids, for example, the imidazole cholesterol ester or “ICE” lipid, (3S, 10R, 13R, 17R)-10, 13-dimethyl-17-((R)-6-methylheptan-2-yl)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-yl 3-(1H-imidazol-4-yl)propanoate, as represented by structure (XIII) below. In an embodiment, a transfer vehicle for delivery of circRNA may comprise one or more imidazole-based ionizable lipids, for example, the imidazole cholesterol ester or “ICE” lipid (3S, 10R, 13R, 17R)-10, 13-dimethyl-17-((R)-6-methylheptan-2-yl)-2, 3, 4, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-yl 3-(1H-imidazol-4-yl)propanoate, as represented by structure (XIII).



(XIII)

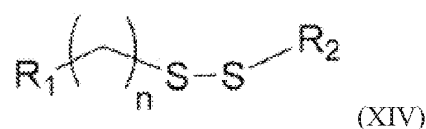
[0479] Without wishing to be bound by a particular theory, it is believed that the fusogenicity of the imidazole-based cationic lipid ICE is related to the endosomal disruption which is facilitated by the imidazole group, which has a lower pKa relative to traditional ionizable lipids. The endosomal disruption in turn promotes osmotic swelling and the

disruption of the liposomal membrane, followed by the transfection or intracellular release of the nucleic acid(s) contents loaded therein into the target cell.

[0480] The imidazole-based ionizable lipids are also characterized by their reduced toxicity relative to other ionizable lipids.

[0481] In some embodiments, an ionizable lipid is described by US patent publication number 20190314284. In certain embodiments, the an ionizable lipid is described by structure 3, 4, 5, 6, 7, 8, 9, or 10 (*e.g.*, HGT4001, HGT4002, HGT4003, HGT4004 and/or HGT4005). In certain embodiments, the one or more cleavable functional groups (*e.g.*, a disulfide) allow, for example, a hydrophilic functional head-group to dissociate from a lipophilic functional tail-group of the compound (*e.g.*, upon exposure to oxidative, reducing or acidic conditions), thereby facilitating a phase transition in the lipid bilayer of the one or more target cells. For example, when a transfer vehicle (*e.g.*, a lipid nanoparticle) comprises one or more of the lipids of structures 3-10, the phase transition in the lipid bilayer of the one or more target cells facilitates the delivery of the circRNA into the one or more target cells.

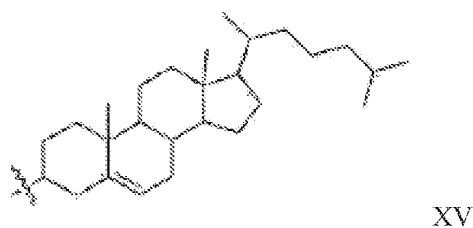
[0482] In certain embodiments, the ionizable lipid is described by structure (XIV),

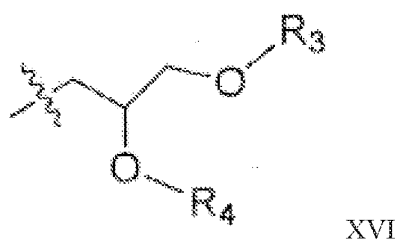


wherein:

R_1 is selected from the group consisting of imidazole, guanidinium, amino, imine, enamine, an optionally-substituted alkyl amino (*e.g.*, an alkyl amino such as dimethylamino) and pyridyl;

R_2 is selected from the group consisting of structure XV and structure XVI,





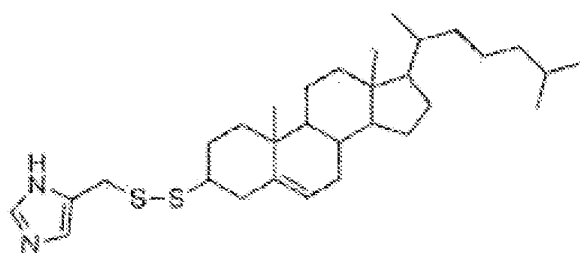
wherein R₃ and R₄ are each independently selected from the group consisting of an optionally substituted, variably saturated or unsaturated C₆-C₂₀ alkyl and an optionally substituted, variably saturated or unsaturated C₆-C₂₀ acyl; and wherein n is zero or any positive integer (*e.g.*, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty or more). In certain embodiments, R₃ and R₄ are each an optionally substituted, polyunsaturated C₁₈ alkyl, while in other embodiments R₃ and R₄ are each an unsubstituted, polyunsaturated C₁₈ alkyl. In certain embodiments, one or more of R₃ and R₄ are (9Z,12Z)-octadeca-9,12-dien.

[0483] Also disclosed herein are pharmaceutical compositions that comprise the compound of structure XIV, wherein R₁ is selected from the group consisting of imidazole, guanidinium, amino, imine, enamine, an optionally-substituted alkyl amino (*e.g.*, an alkyl amino such as dimethylamino) and pyridyl; wherein R₂ is structure XV; and wherein n is zero or any positive integer. Further disclosed herein are pharmaceutical compositions comprising the compound of structure XIV, wherein R₁ is selected from the group consisting of imidazole, guanidinium, amino, imine, enamine, an optionally-substituted alkyl amino (*e.g.*, an alkyl amino such as dimethylamino) and pyridyl; wherein R₂ is structure XVI; wherein R₃ and R₄ are each independently selected from the group consisting of an optionally substituted, variably saturated or unsaturated C₆-C₂₀ alkyl and an optionally substituted, variably saturated or unsaturated C₆-C₂₀ acyl; and wherein n is zero or any positive integer. In certain embodiments, R₃ and R₄ are each an optionally substituted, polyunsaturated C₁₈ alkyl, while in other embodiments R₃ and R₄ are each an unsubstituted, polyunsaturated C₁₈ alkyl (*e.g.*, octadeca-9,12-dien).

[0484] In certain embodiments, the R₁ group or head-group is a polar or hydrophilic group (*e.g.*, one or more of the imidazole, guanidinium and amino groups) and is bound to the R₂ lipid group by way of the disulfide (S—S) cleavable linker group, for example as depicted in structure XIV. Other contemplated cleavable linker groups may include compositions that comprise one or more disulfide (S—S) linker group bound (*e.g.*, covalently bound) to, for example an alkyl group (*e.g.*, C₁ to C₁₀ alkyl). In certain embodiments, the R₁ group is

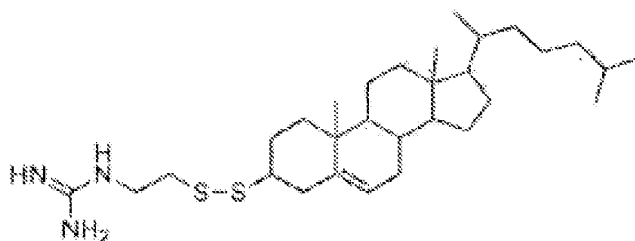
covalently bound to the cleavable linker group by way of a C₁-C₂₀ alkyl group (*e.g.*, where n is one to twenty), or alternatively may be directly bound to the cleavable linker group (*e.g.*, where n is zero). In certain embodiments, the disulfide linker group is cleavable *in vitro* and/or *in vivo* (*e.g.*, enzymatically cleavable or cleavable upon exposure to acidic or reducing conditions).

[0485] In certain embodiments, the inventions relate to the compound 5-(((10,13-dimethyl-17-(6-methylheptan-2-yl)-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-yl)disulfanyl)methyl)-1H-imidazole, having structure XVII (referred to herein as “HGT4001”).



XVII

[0486] In certain embodiments, the inventions relate to the compound 1-(2-(((3S,10R,13R)-10,13-dimethyl-17-((R)-6-methylheptan-2-yl)-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-yl)disulfanyl)ethyl)guanidine, having structure XVIII (referred to herein as “HGT4002”).



XVIII

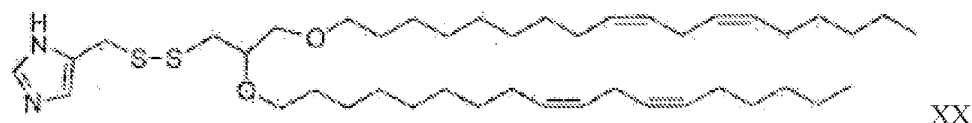
[0487] In certain embodiments, the inventions relate to the compound 2-((2,3-bis((9Z,12Z)-octadeca-9,12-dien-1-yloxy)propyl)disulfanyl)-N,N-dimethylethanamine, having structure XIX (referred to herein as “HGT4003”).



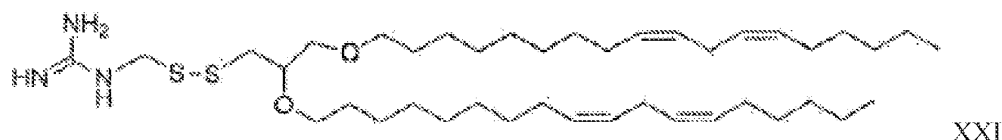
XIX

[0488] In other embodiments, the inventions relate to the compound 5-(((2,3-bis((9Z,12Z)-octadeca-9,12-dien-1-yloxy)propyl)disulfanyl)methyl)-1H-imidazole having the

structure of structure XX (referred to herein as “HGT4004”).



[0489] In still other embodiments, the inventions relate to the compound 1-(((2,3-bis((9Z,12Z)-octadeca-9,12-dien-1-yloxy)propyl)disulfanyl)methyl)guanidine having structure XXI (referred to herein as “HGT4005”).



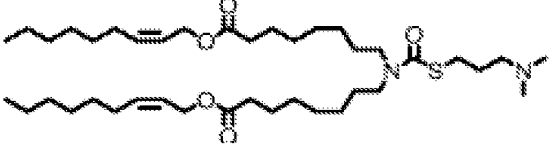
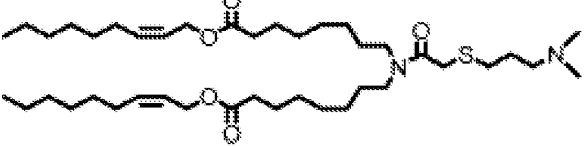
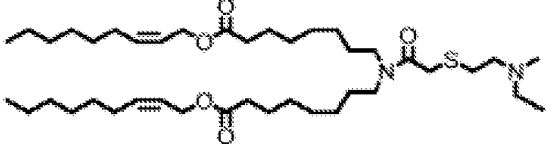
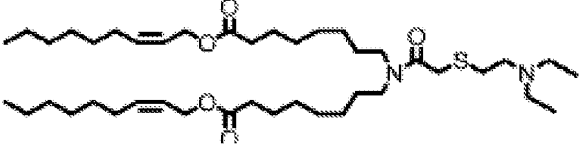
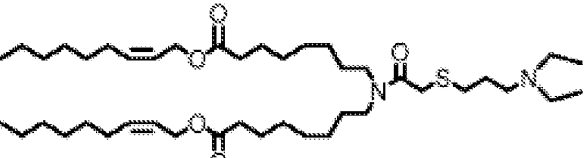
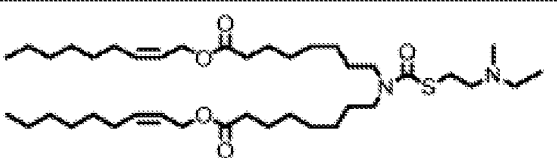
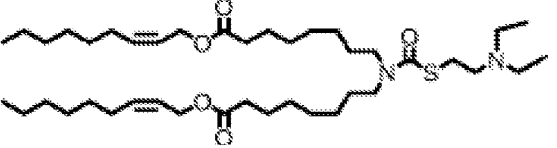
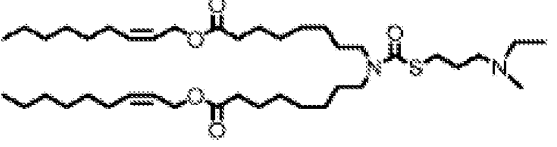
[0490] In certain embodiments, the compounds described as structures 3-10 are ionizable lipids.

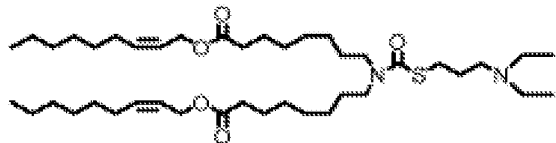
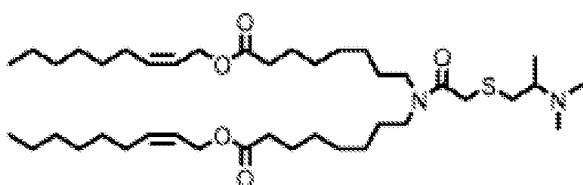
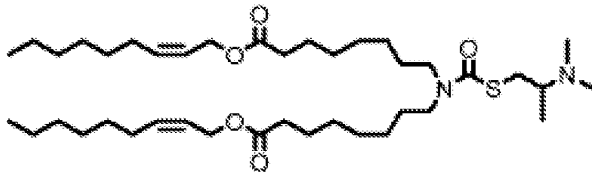
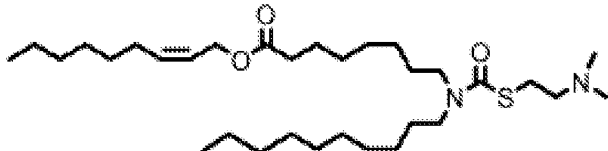
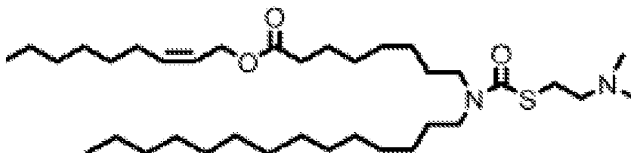
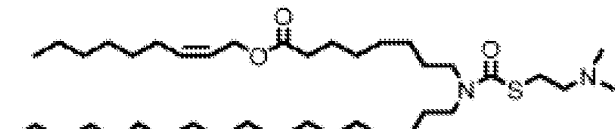
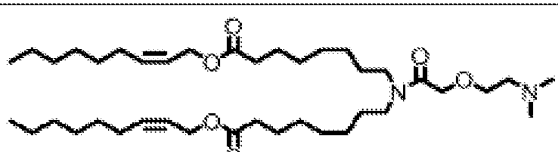
[0491] The compounds, and in particular the imidazole-based compounds described as structures 3-8 (*e.g.*, HGT4001 and HGT4004), are characterized by their reduced toxicity, in particular relative to traditional ionizable lipids. In some embodiments, the transfer vehicles described herein comprise one or more imidazole-based ionizable lipid compounds such that the relative concentration of other more toxic ionizable lipids in such pharmaceutical or liposomal composition may be reduced or otherwise eliminated.

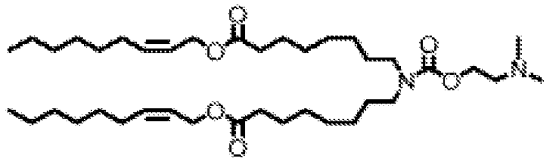
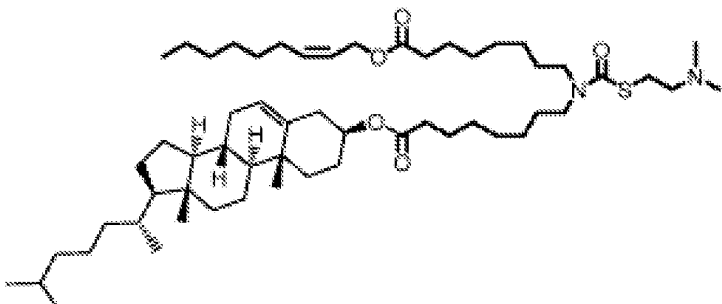
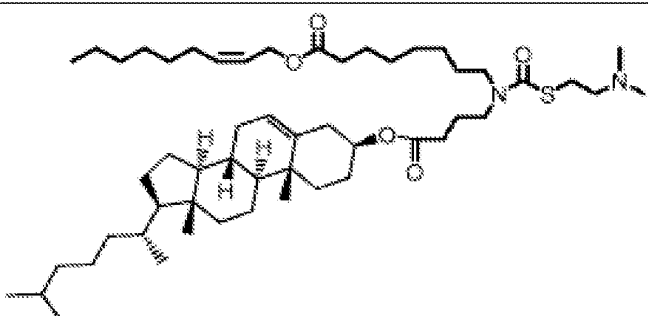
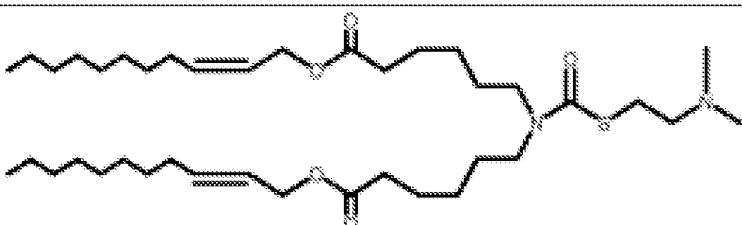
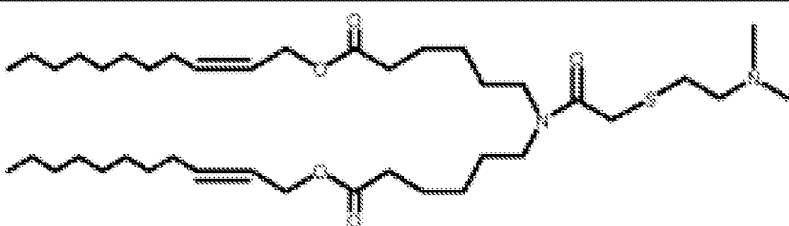
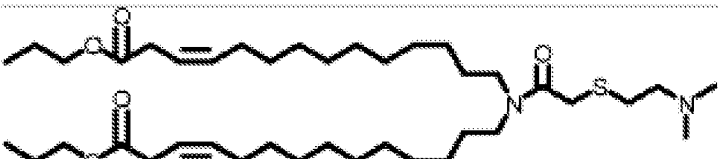
[0492] The ionizable lipids include those disclosed in international patent application PCT/US2019/025246, and US patent publications 2017/0190661 and 2017/0114010, incorporated herein by reference in their entirety. The ionizable lipids may include a lipid selected from the following tables 12, 13, 14, or 15.

Table 12

ATX-001	
ATX-002	

ATX-003	
ATX-004	
ATX-005	
ATX-006	
ATX-007	
ATX-008	
ATX-009	
ATX-010	
ATX-011	

ATX-012	
ATX-013	
ATX-014	
ATX-015	
ATX-016	
ATX-017	
ATX-018	
ATX-019	
ATX-020	
ATX-021	

ATX-022	
ATX-023	
ATX-024	
ATX-025	
ATX-026	
ATX-027	
ATX-028	

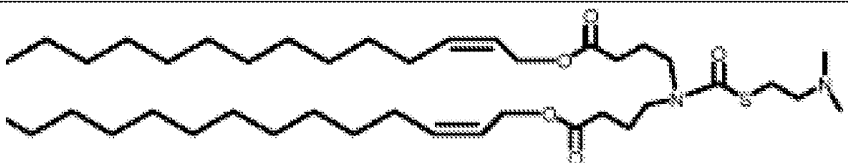
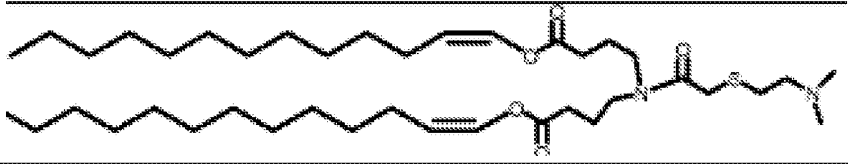
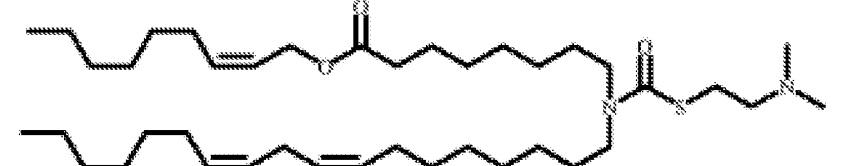
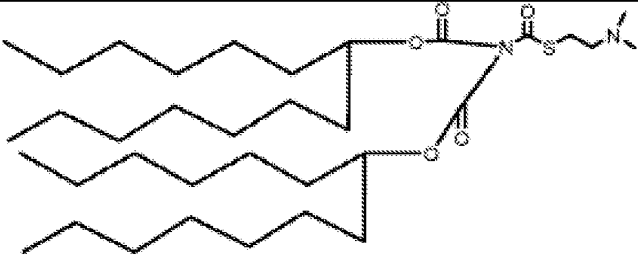
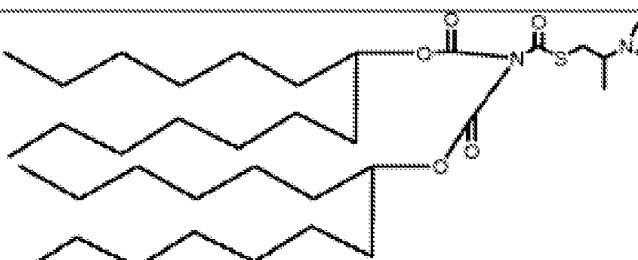
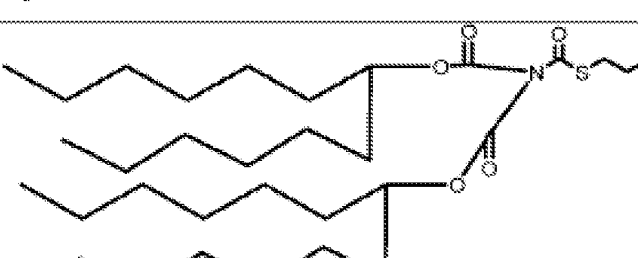
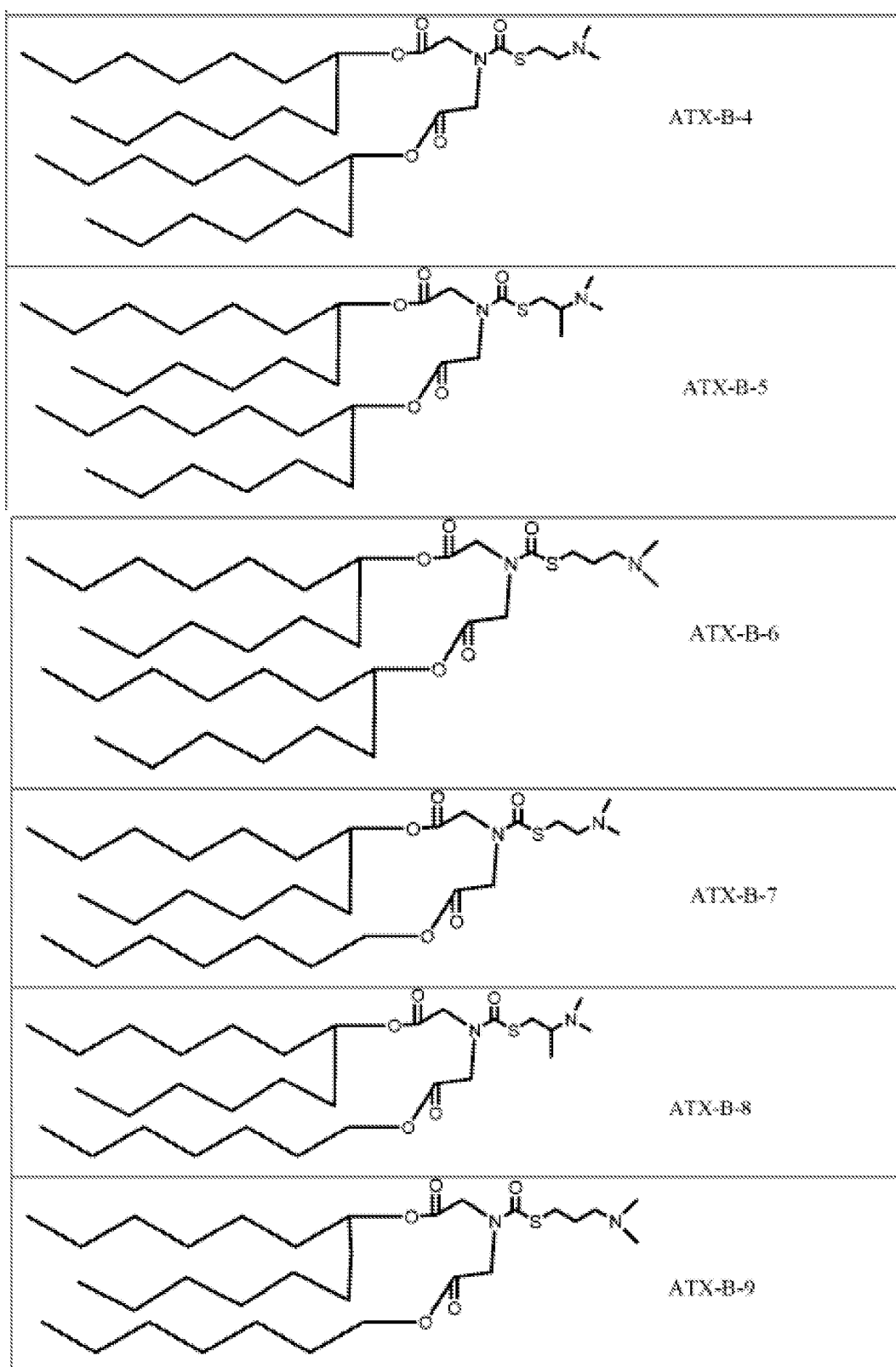
ATX-029	
ATX-030	
ATX-031	
ATX-032	

Table 13

	ATX-B-1
	ATX-B-2
	ATX-B-3



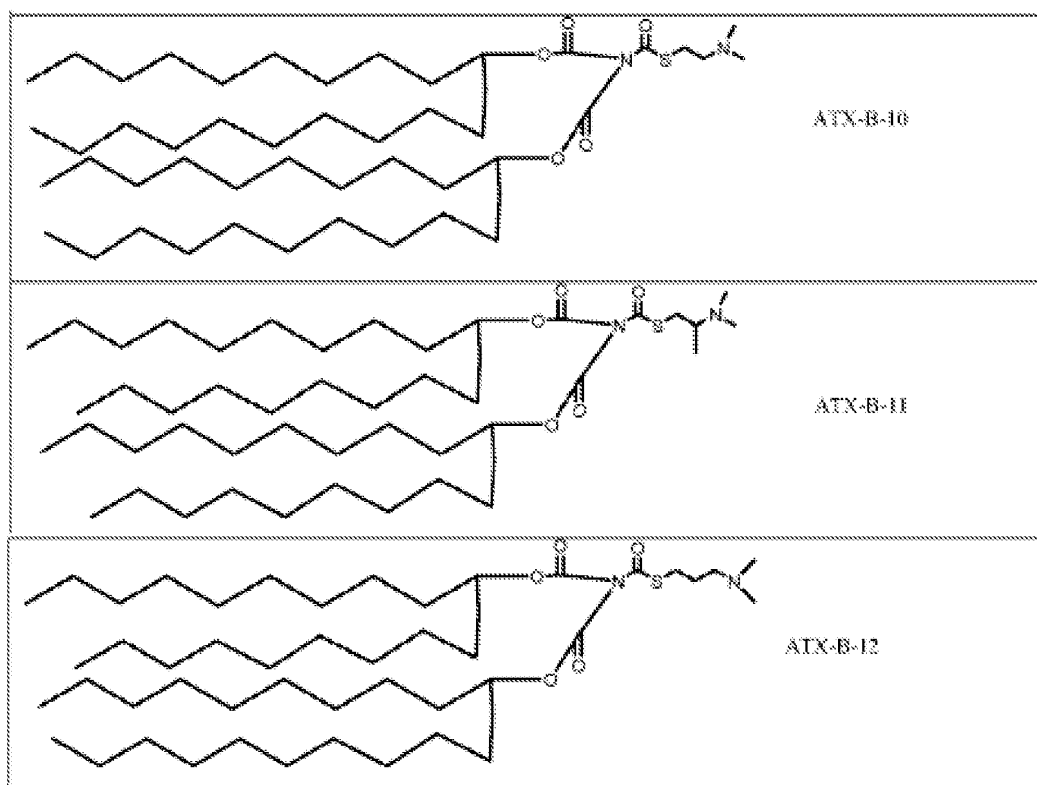
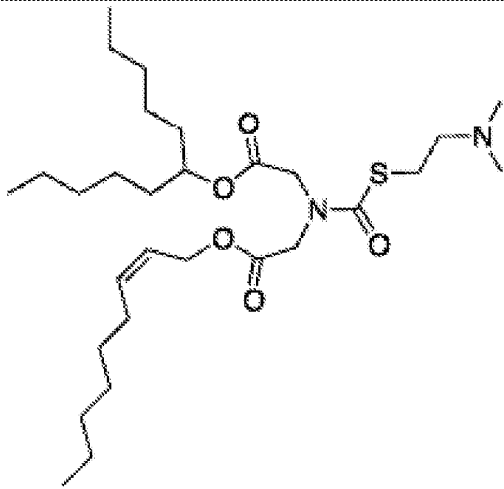
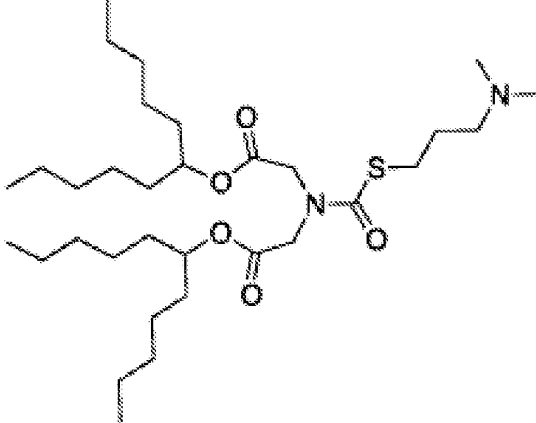
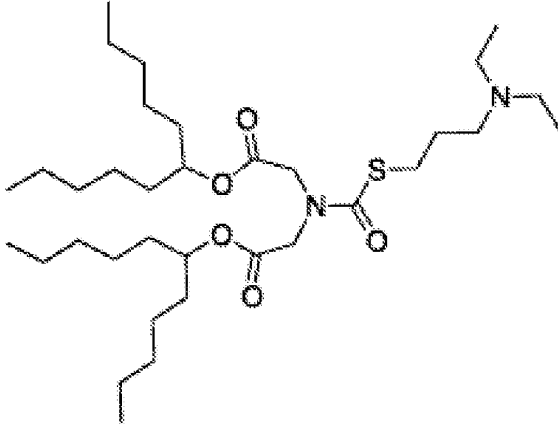
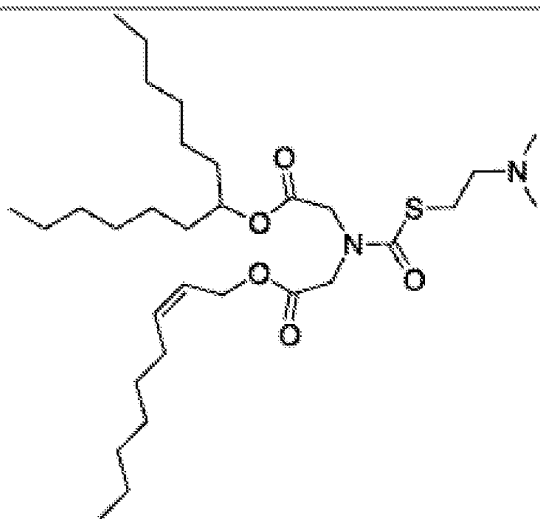
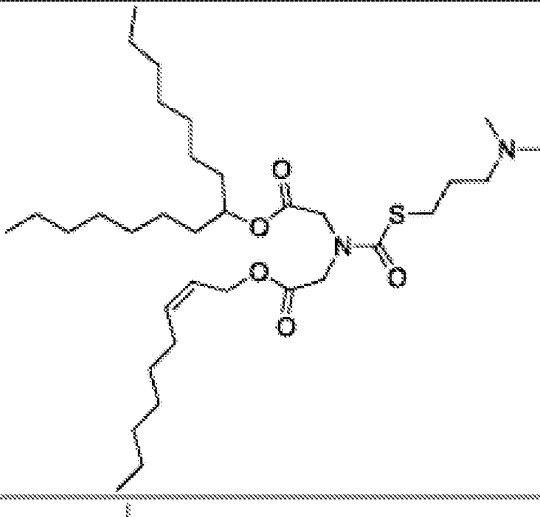
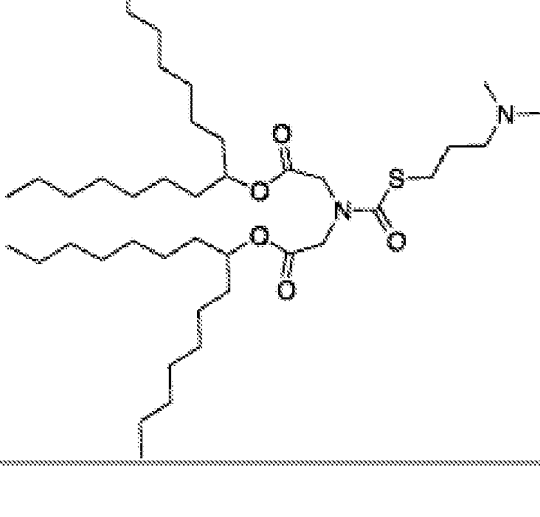
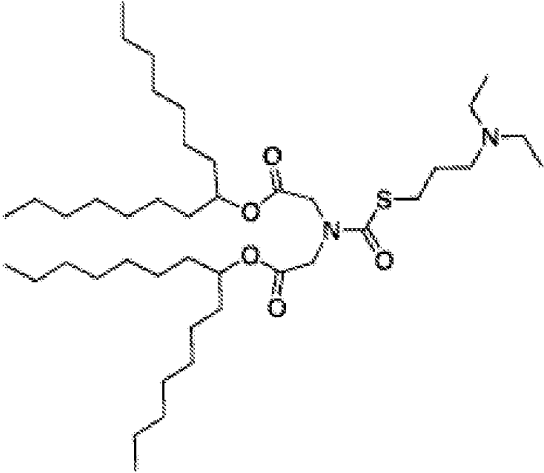
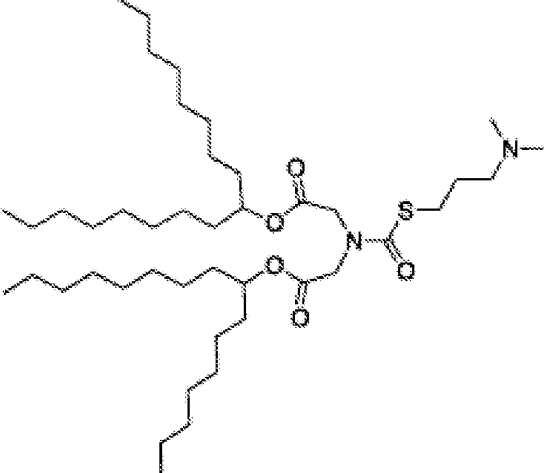
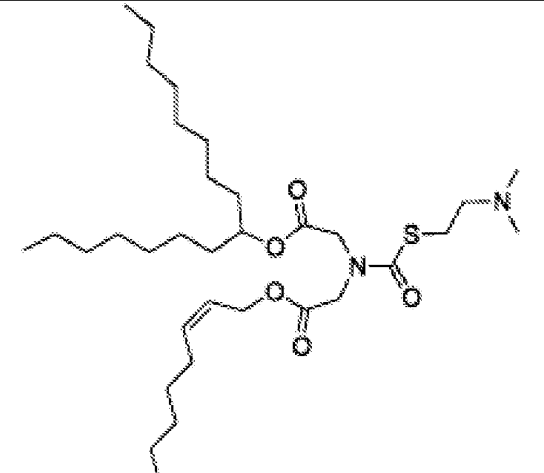
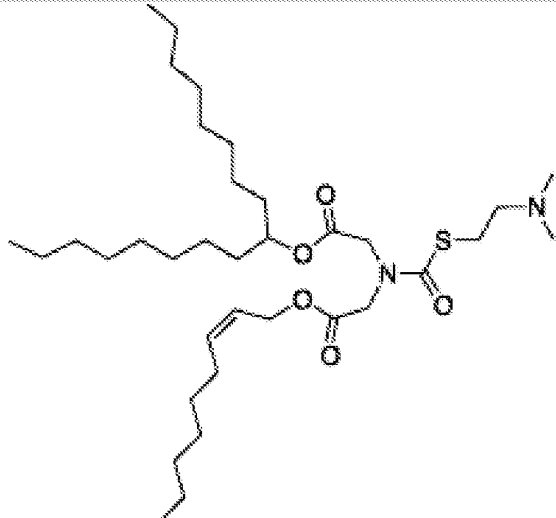
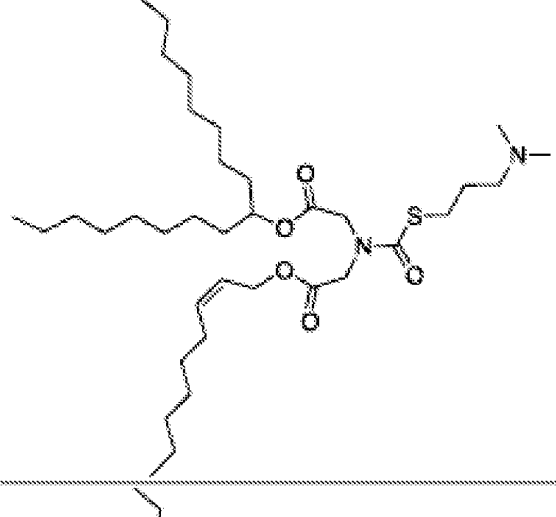
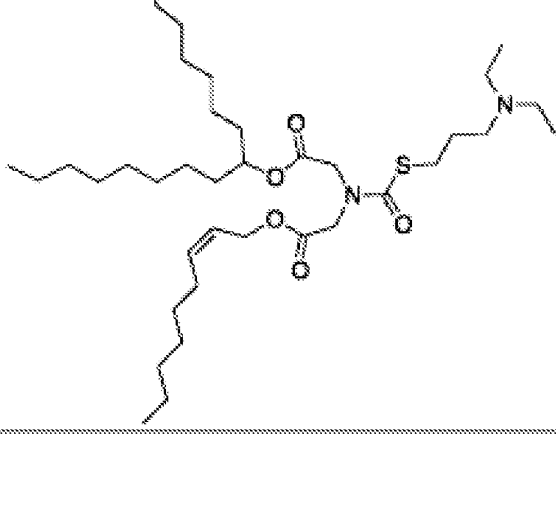


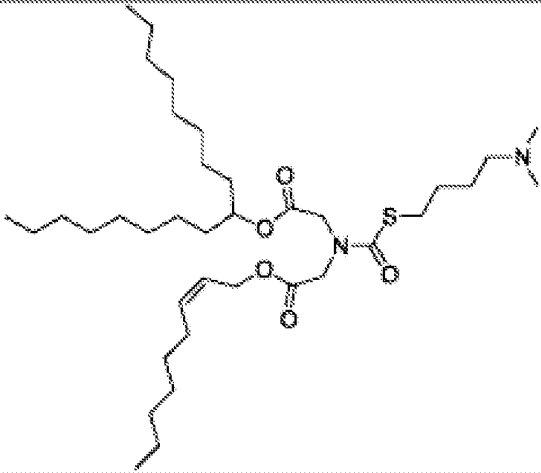
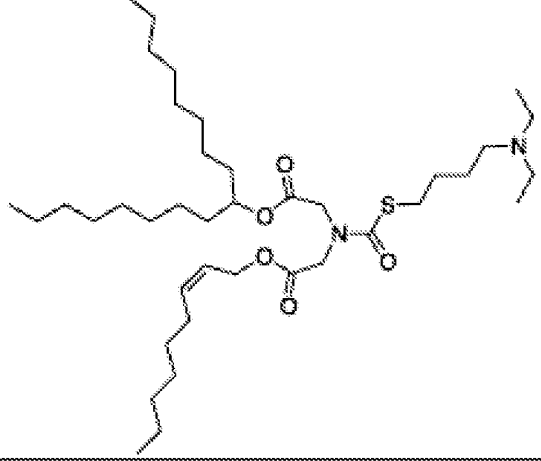
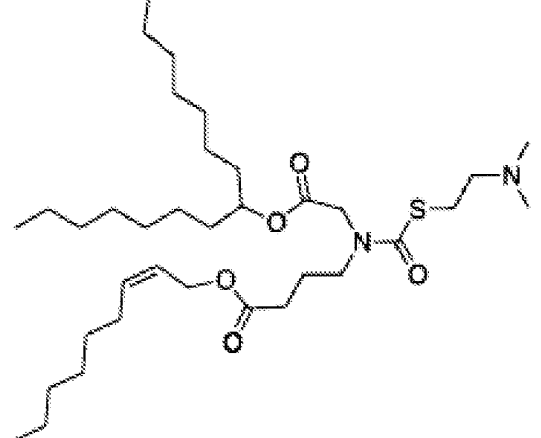
Table 14

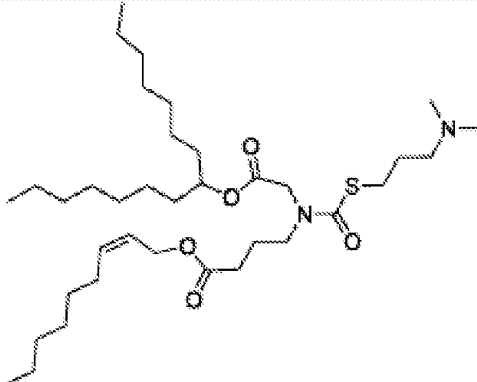
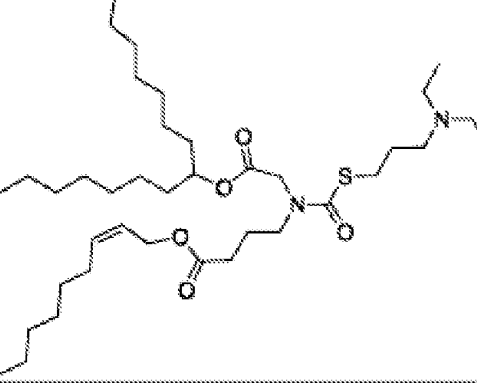
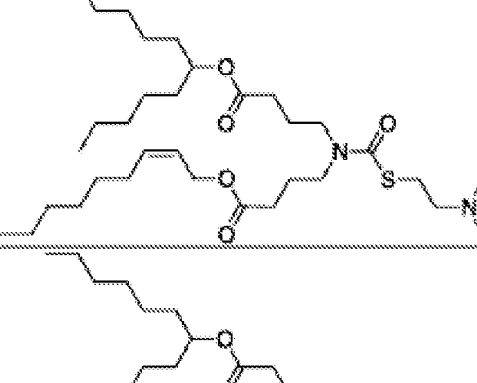
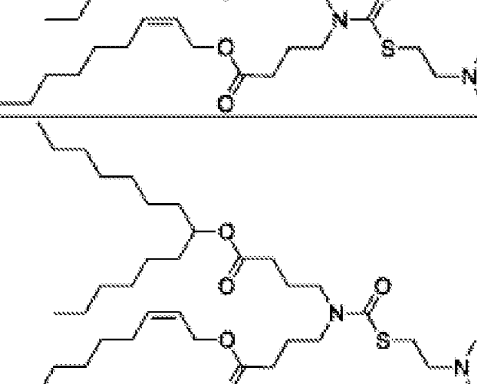
Compound	ATX-#
	0063
	0130
	0131

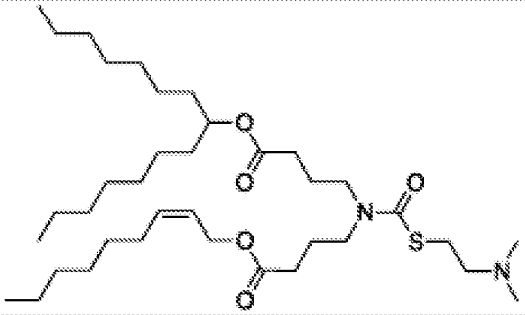
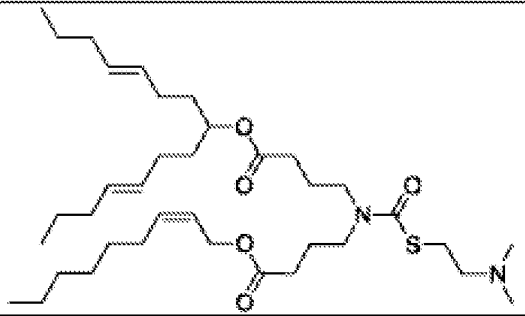
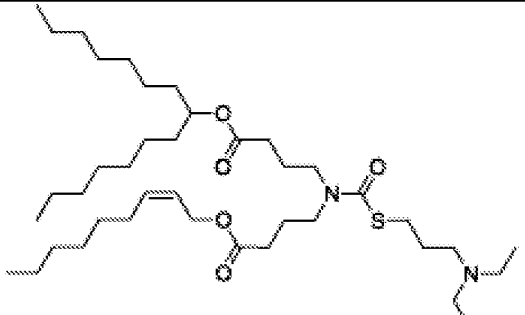
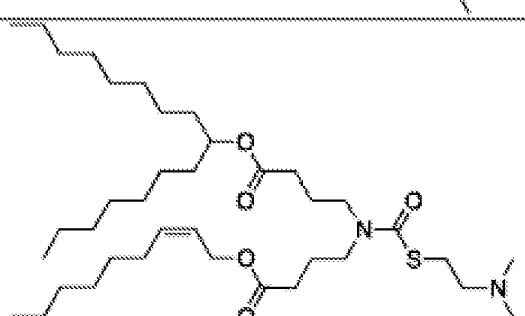
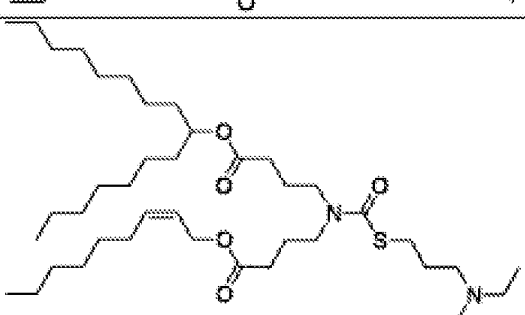
	0044
	0111
	0132

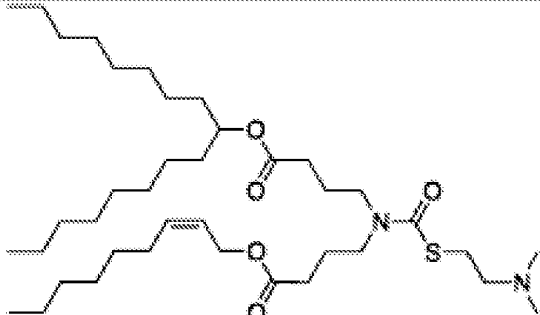
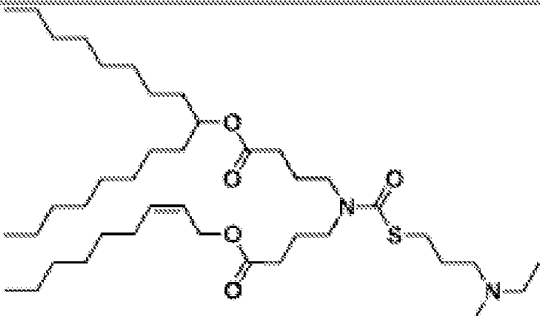
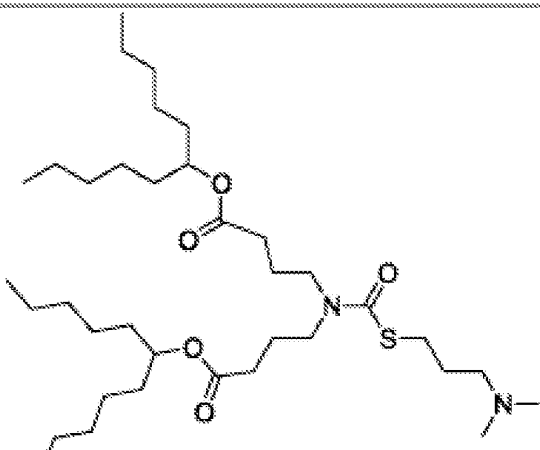
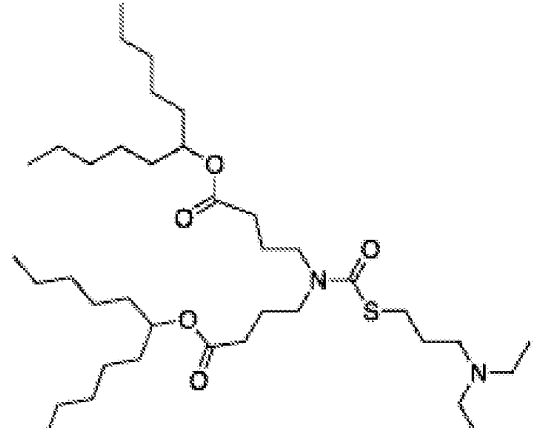
	0134
	0133
	0064

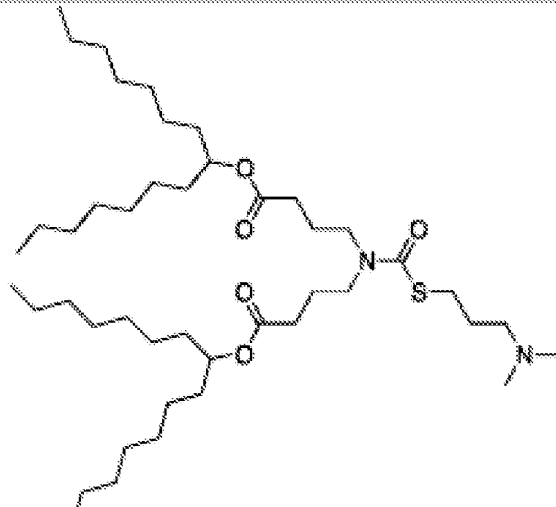
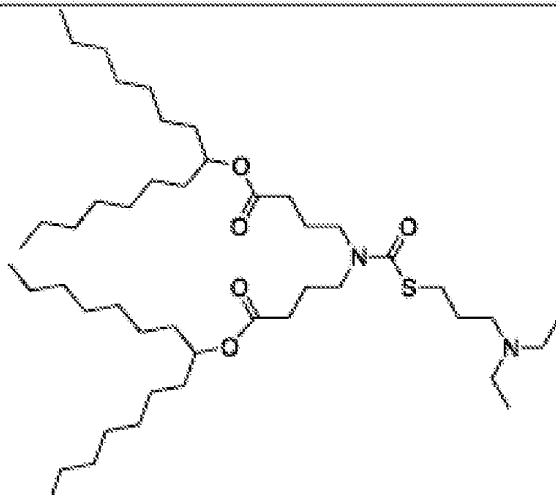
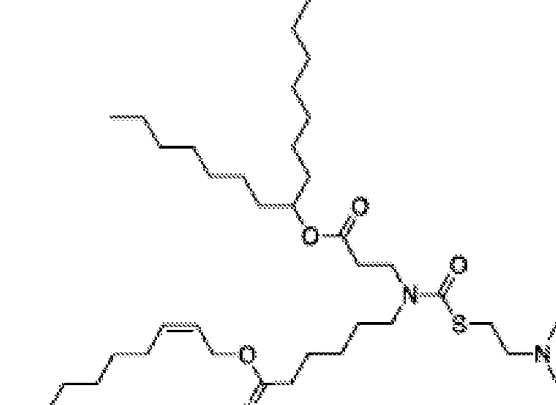
	0061
	0100
	0117

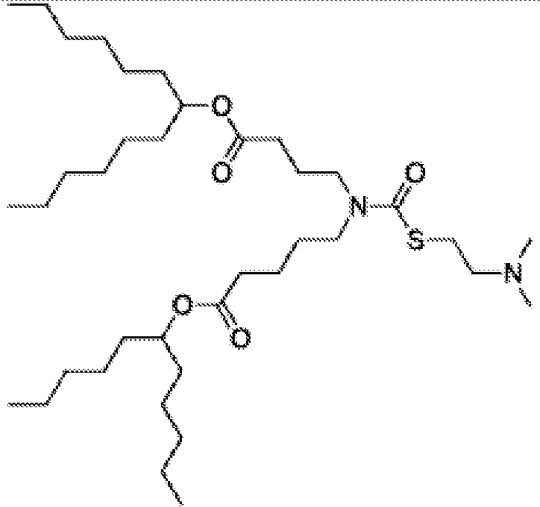
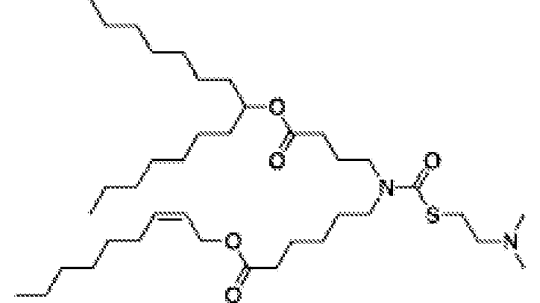
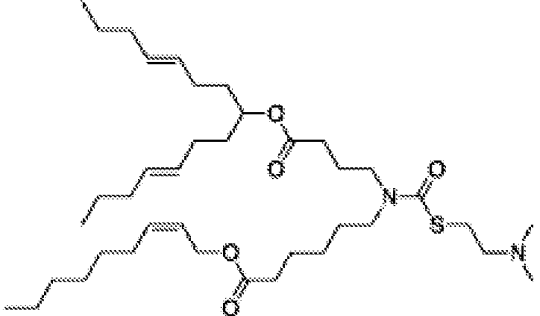
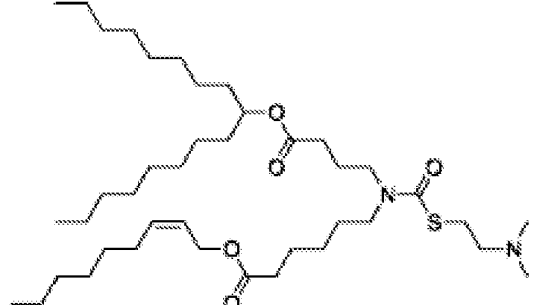
	0114
	0115
	0101

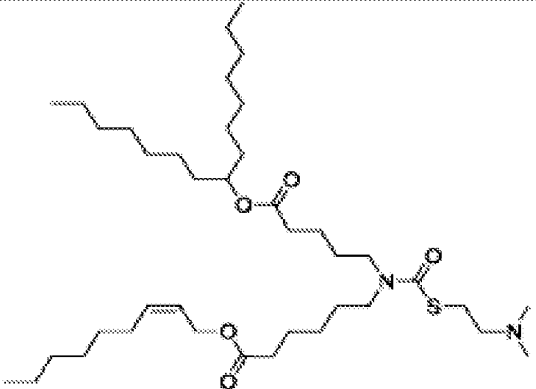
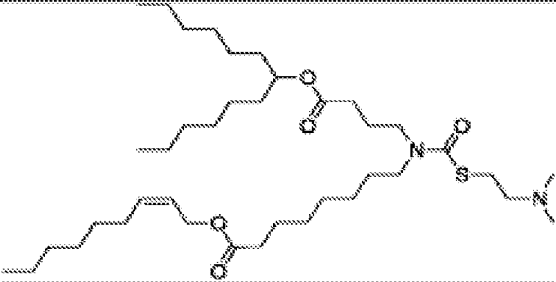
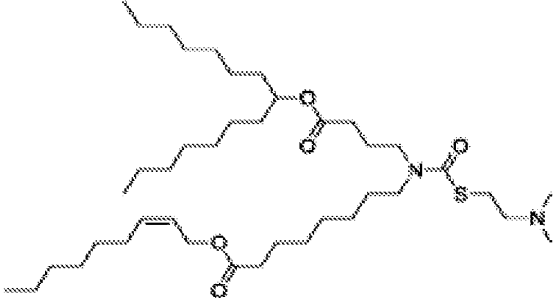
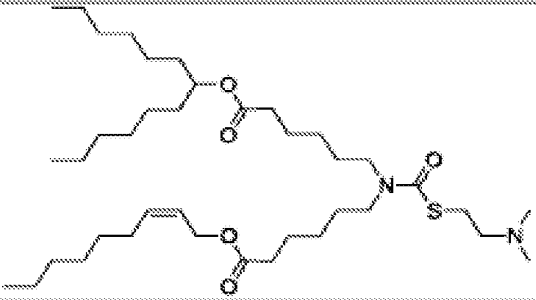
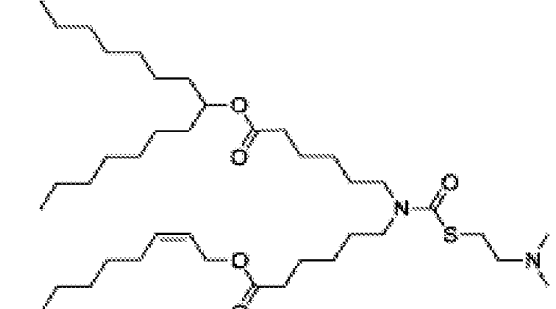
	0106
	0116
	0043
	0086
	0058

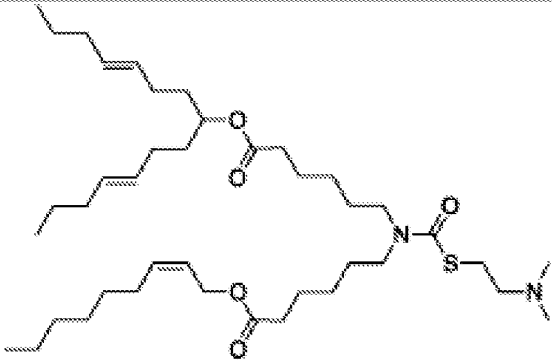
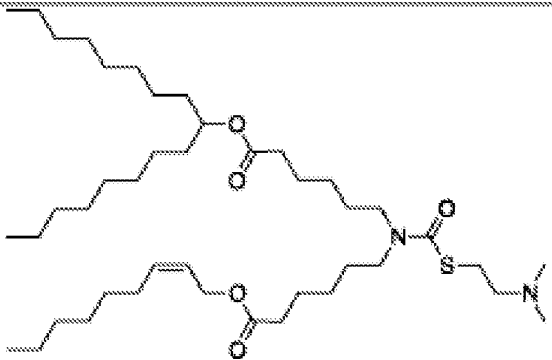
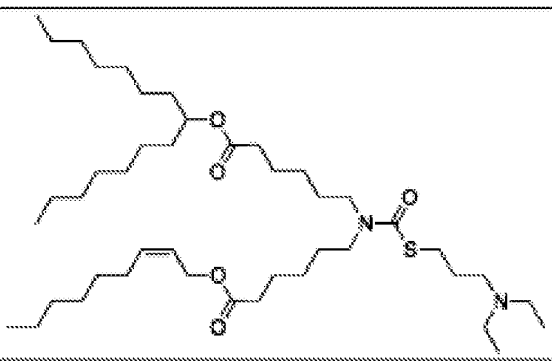
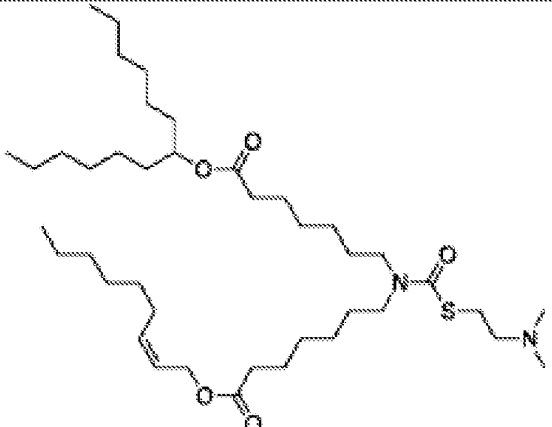
	0081
	0123
	0122
	0057
	0088

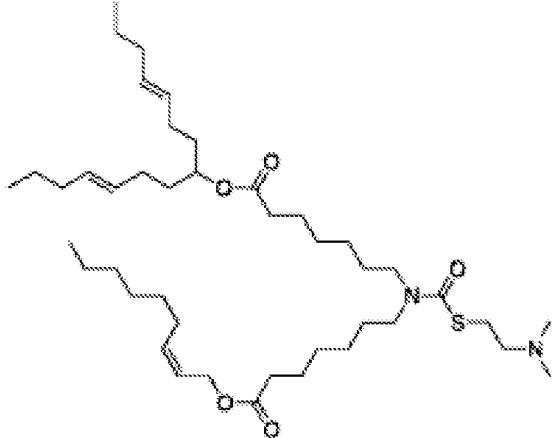
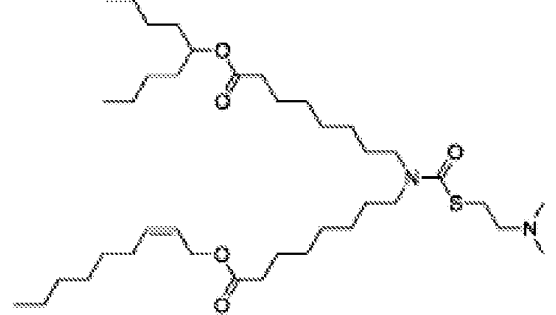
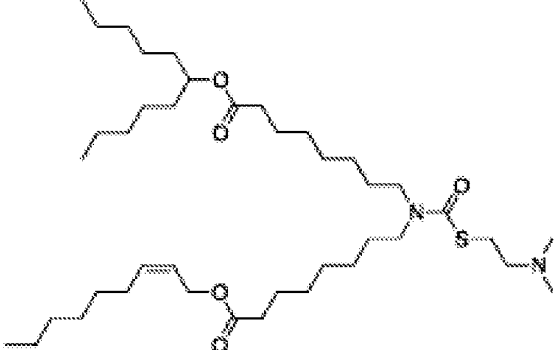
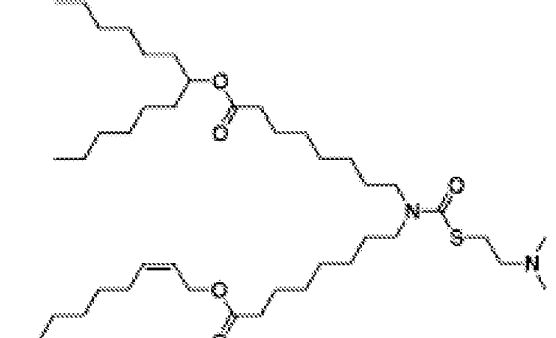
	0087
	0124
	0128
	0127

	0126
	0129
	0082

	0085
	0083
	0121
	0091

	0102
	0098
	0092
	0084
	0095

	0125
	0094
	0109
	0110

	0118
	0108
	0107
	0093

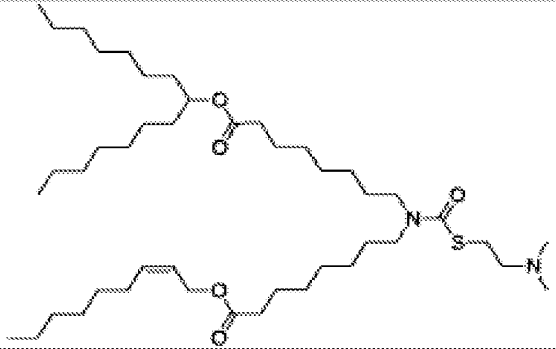
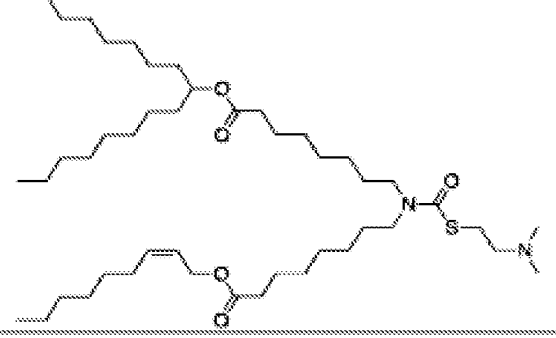
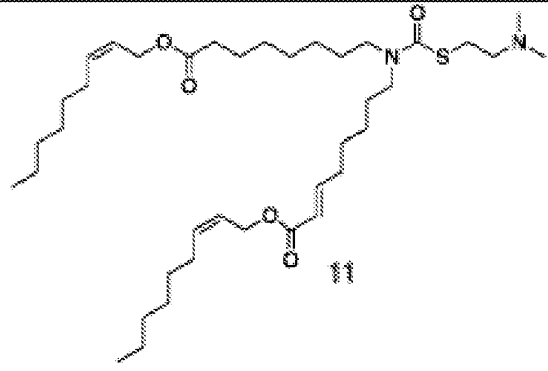
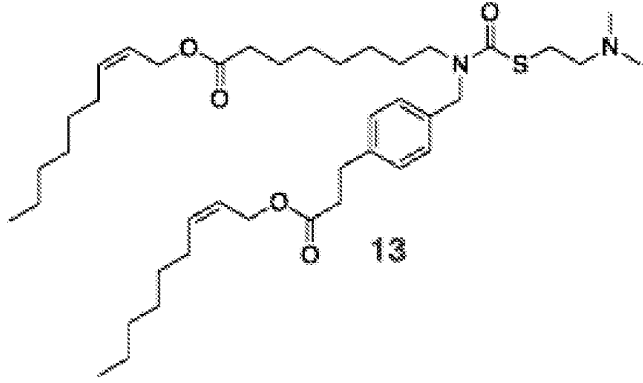
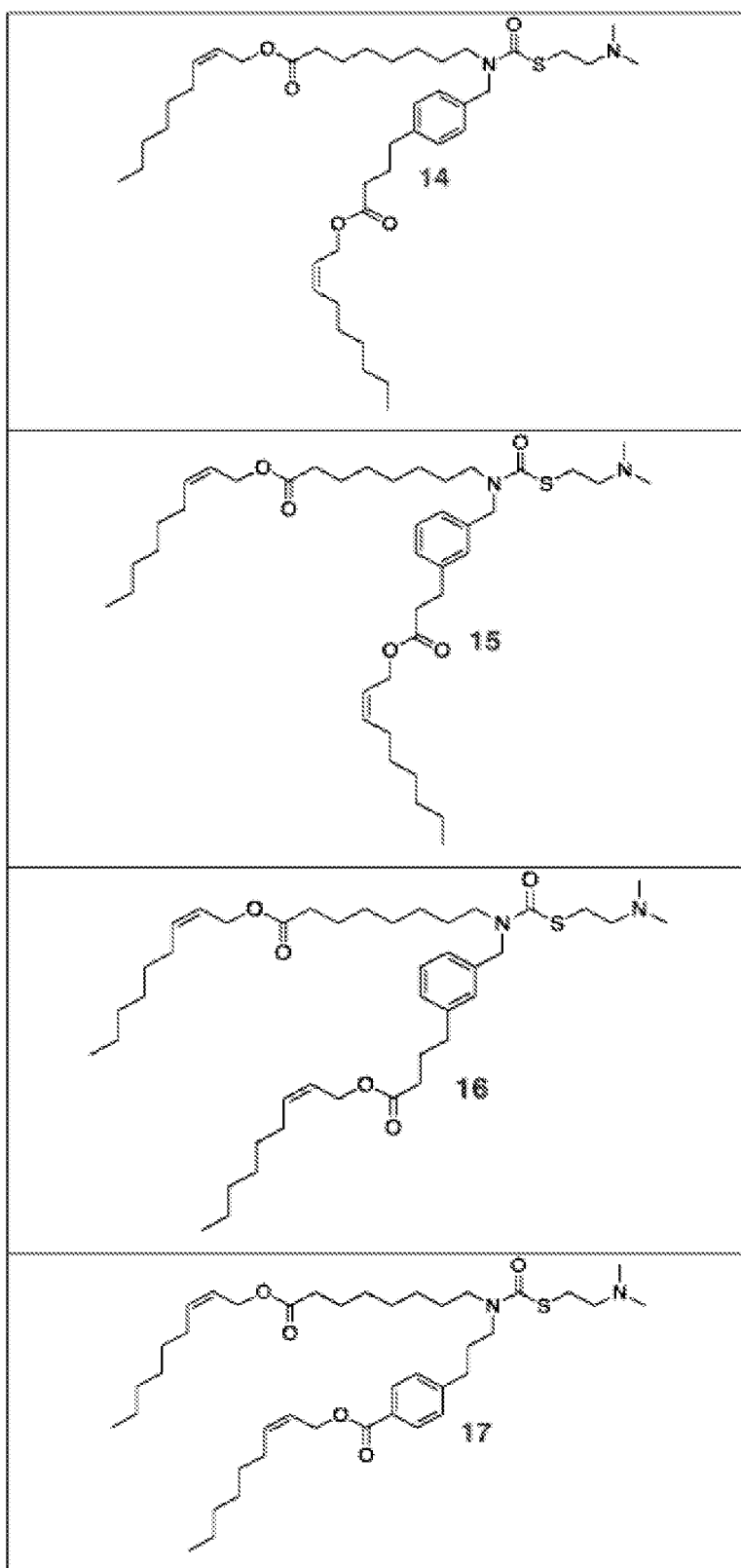
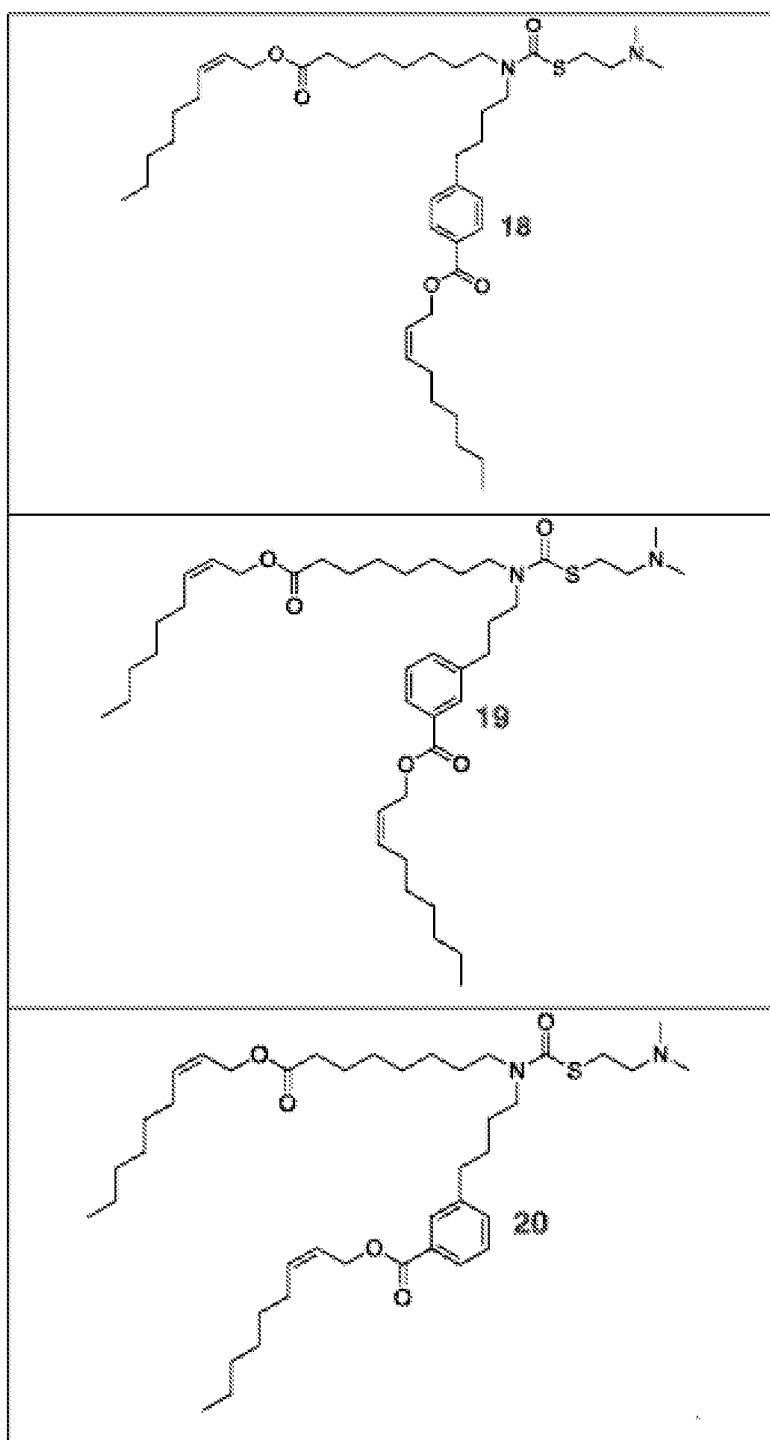
	0097
	0096

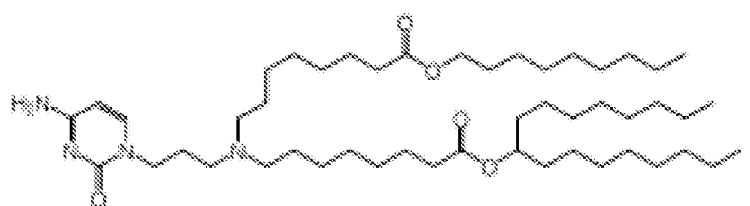
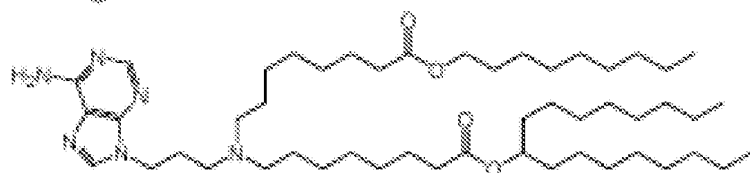
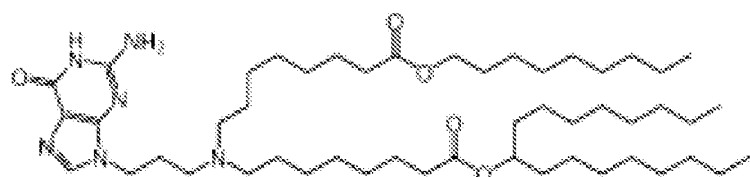
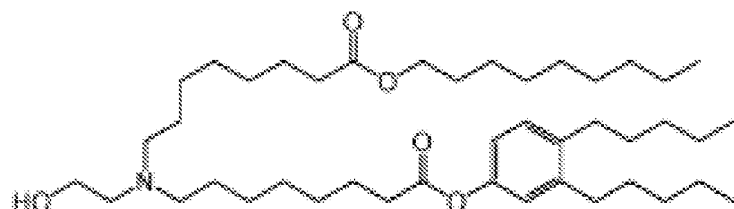
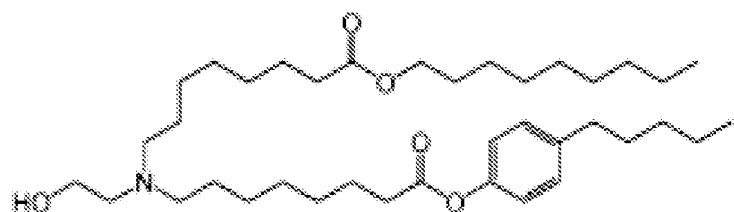
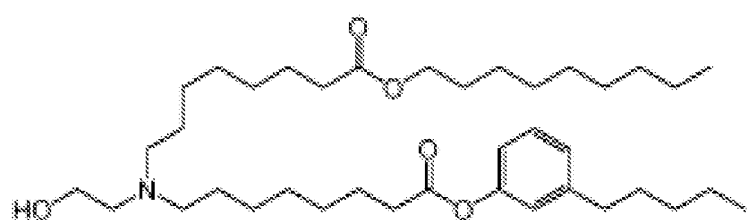
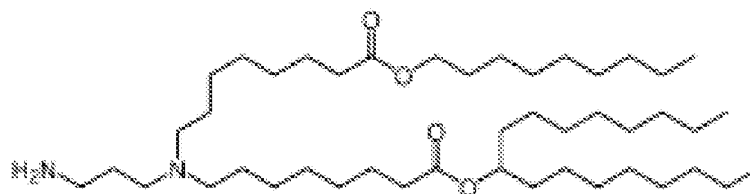
Table 15

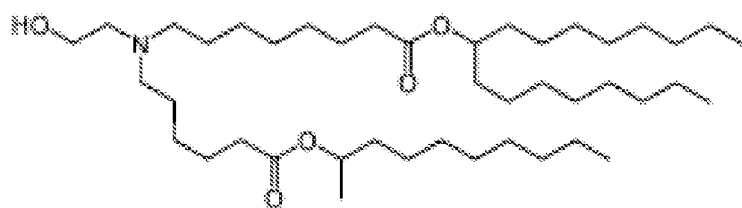
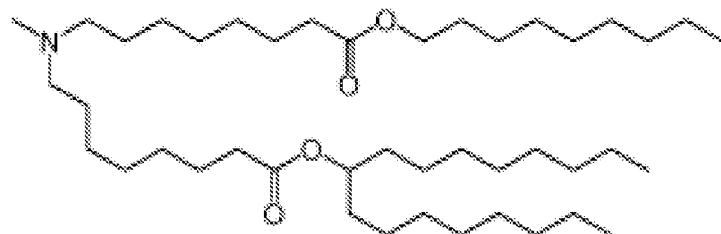
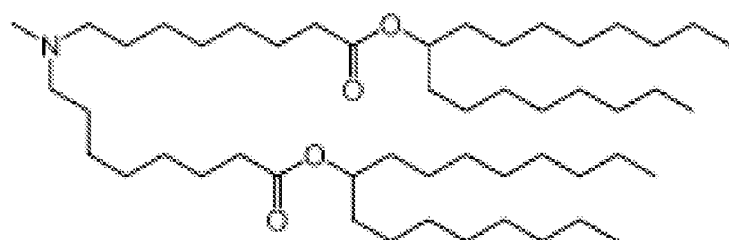
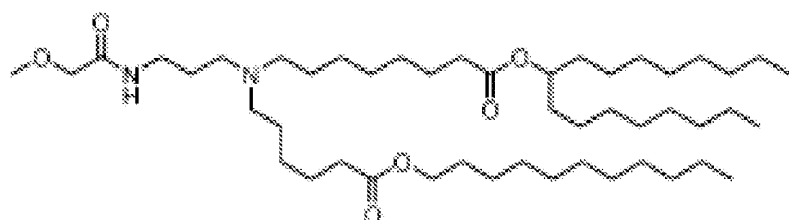
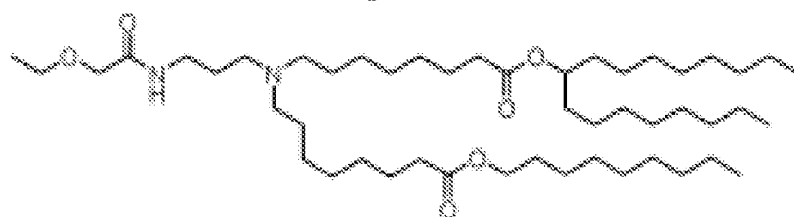
	11
	13

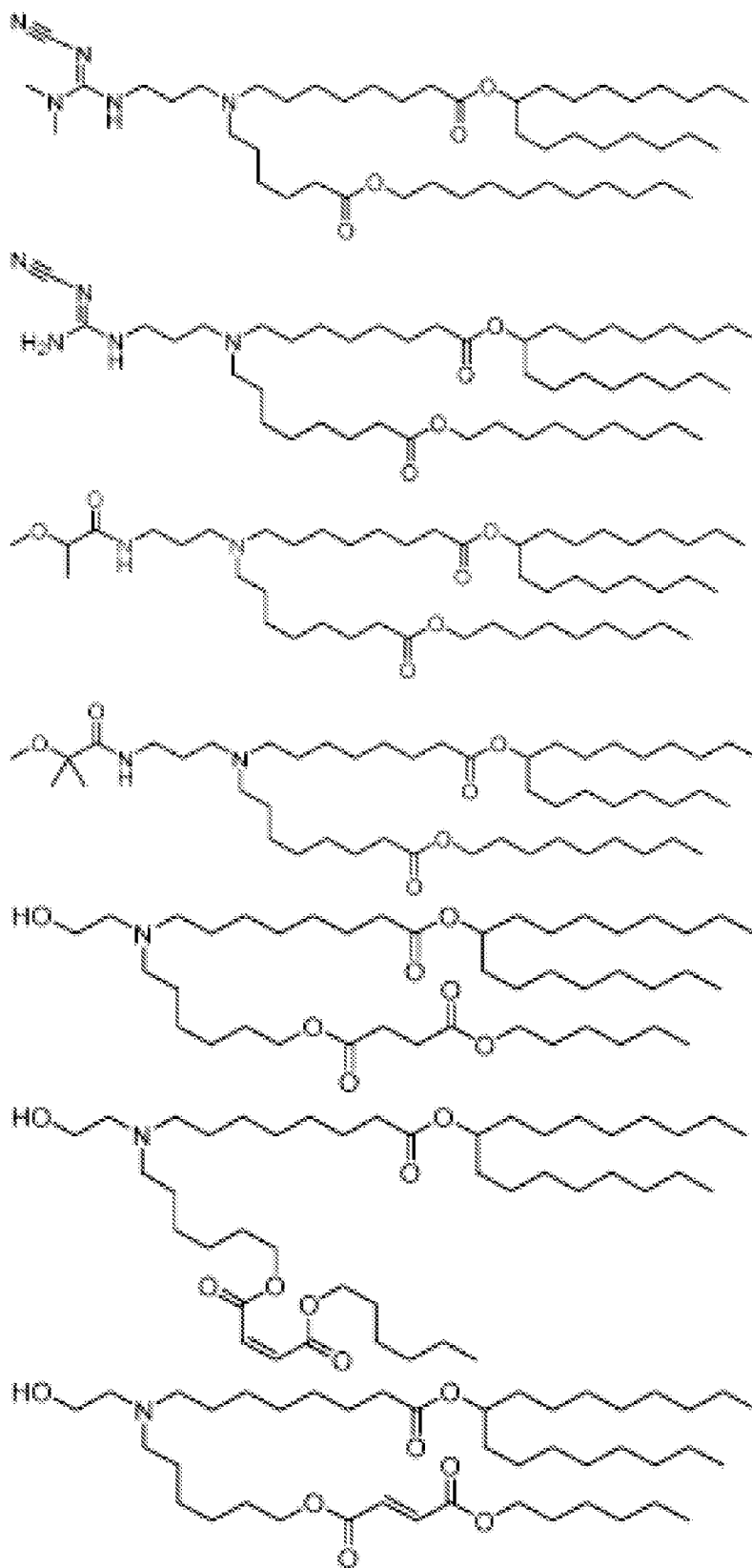


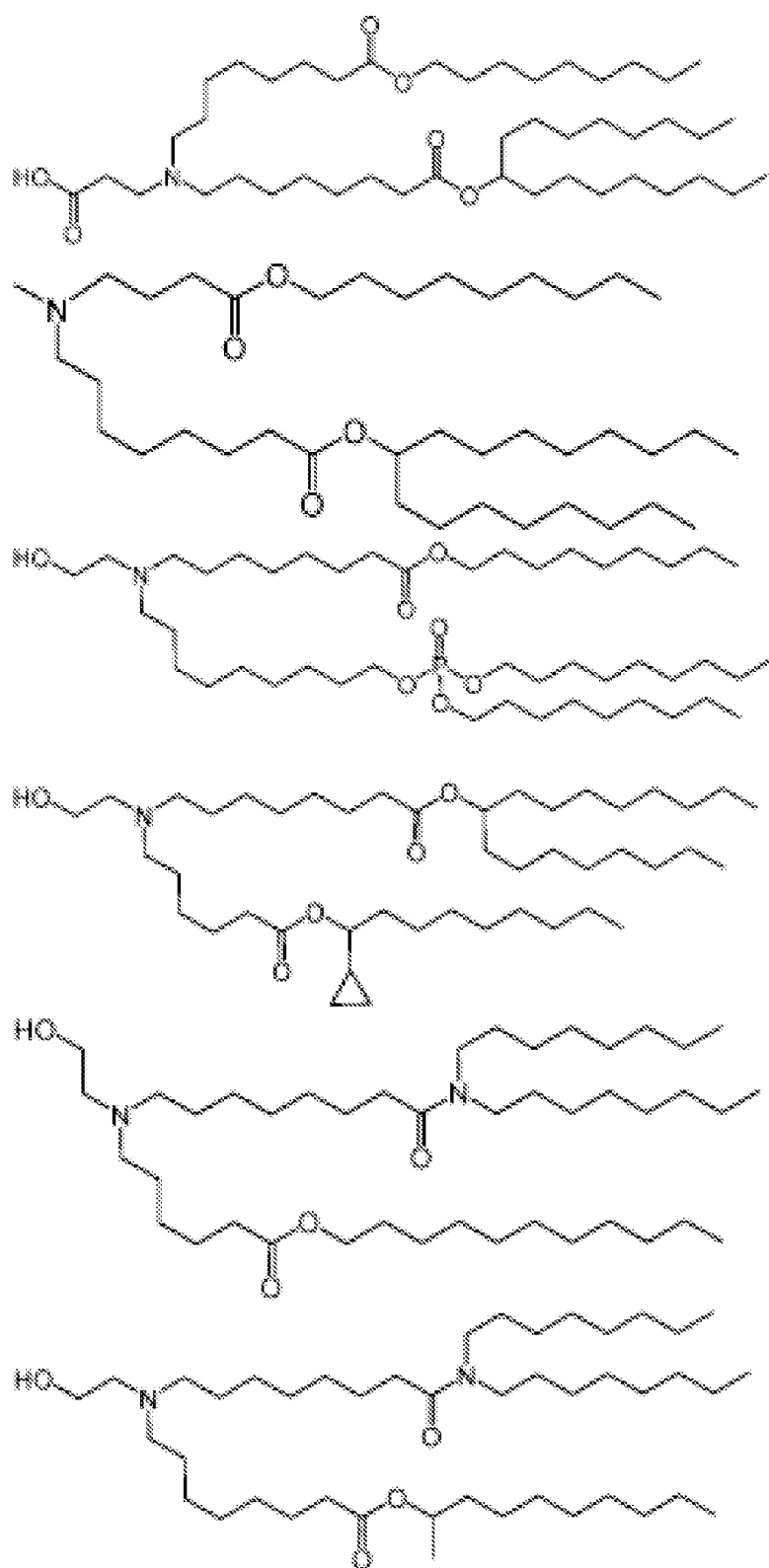


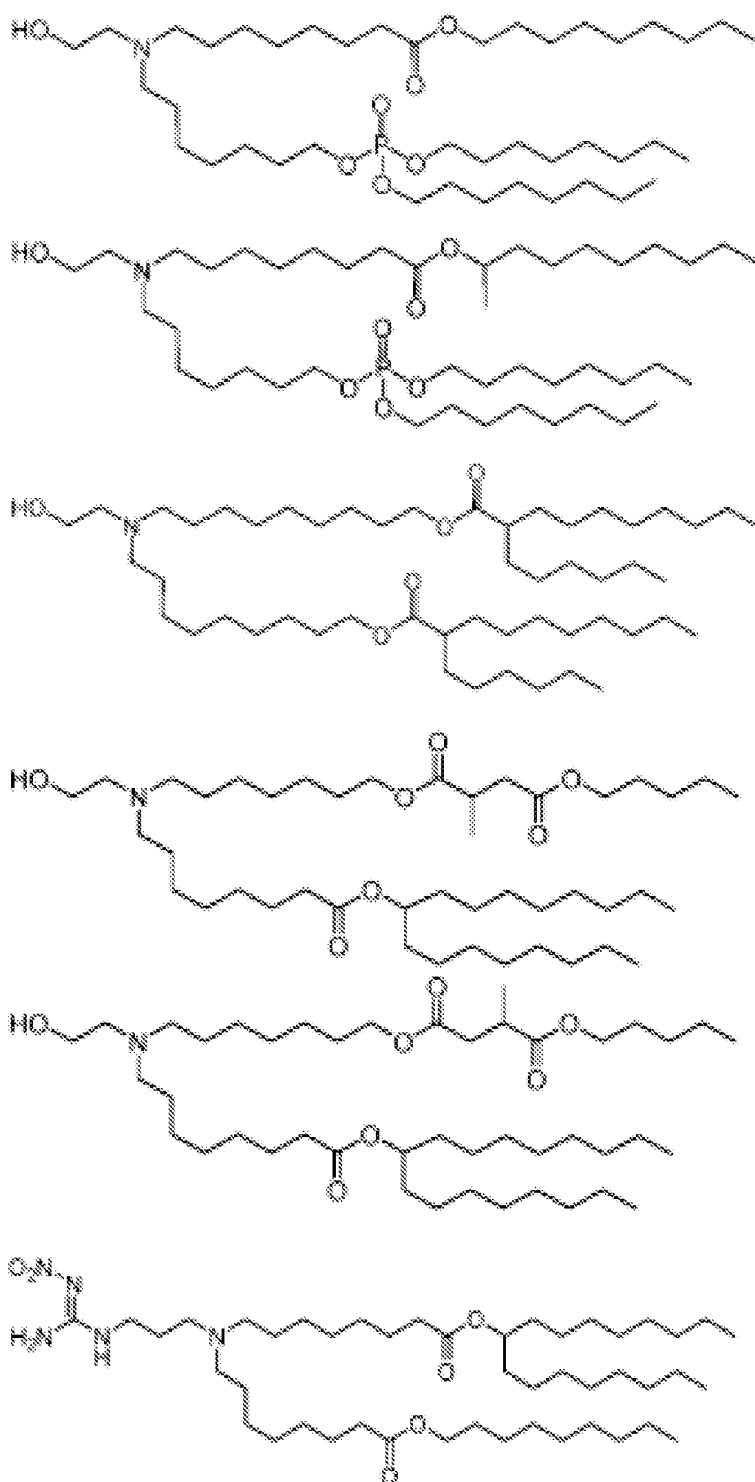
[0493] In some embodiments, an ionizable lipid is as described in international patent application PCT/US2019/015913. In some embodiments, an ionizable lipid is chosen from the following:

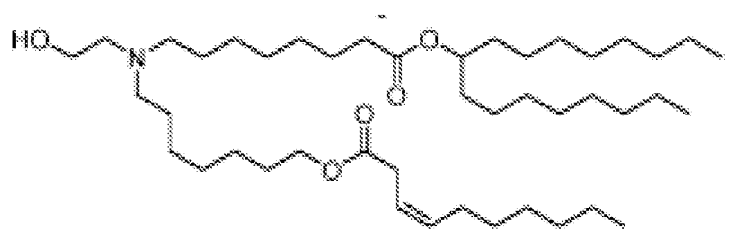
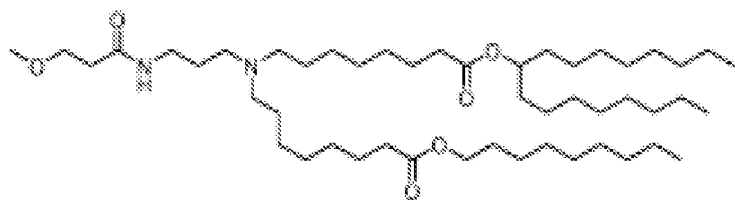
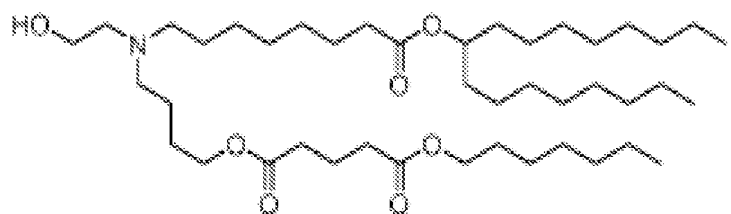
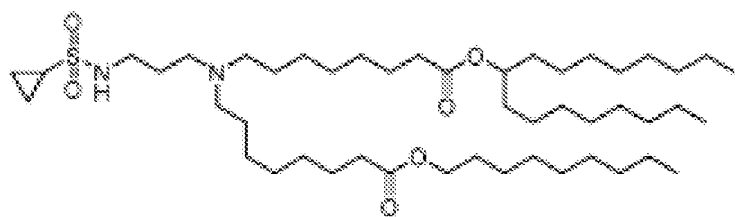
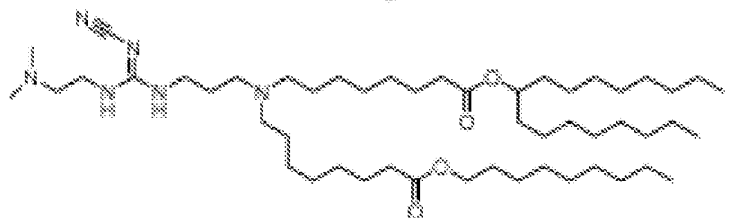
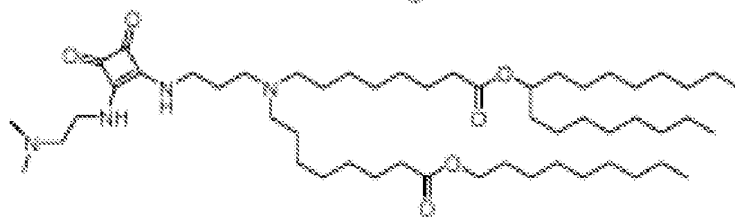


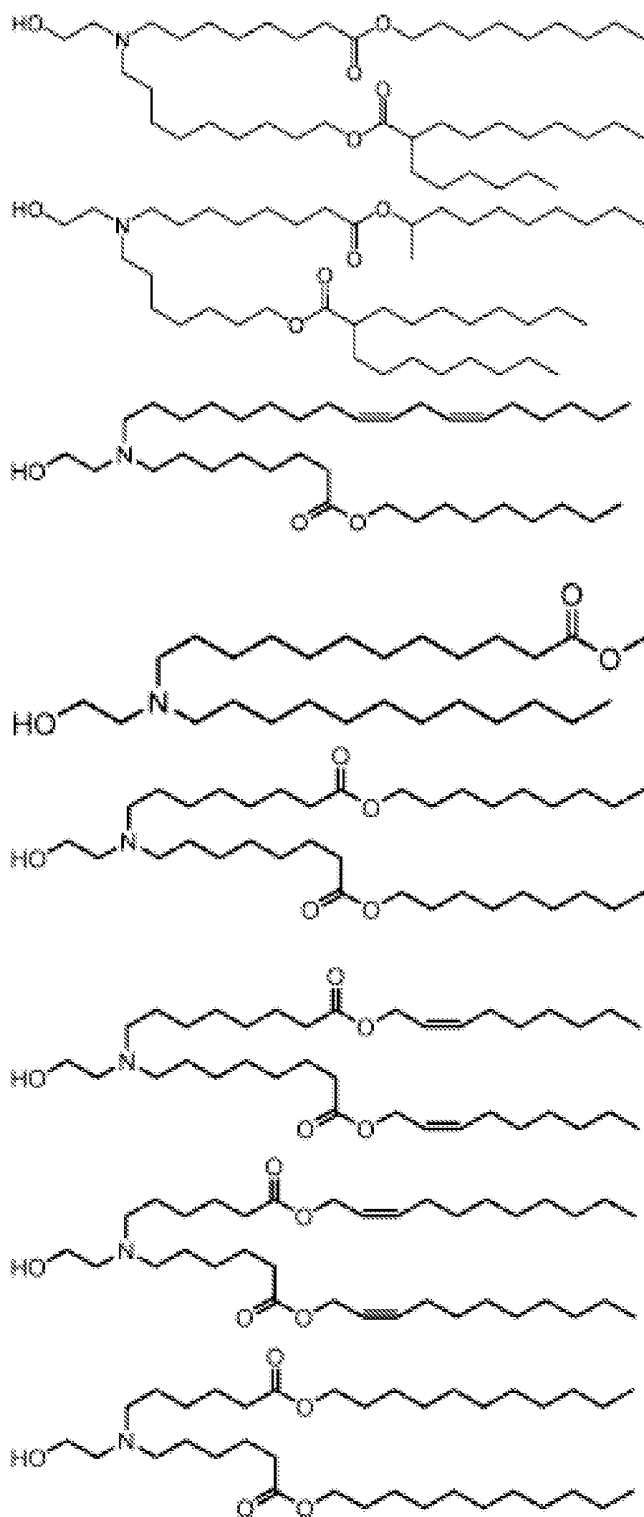


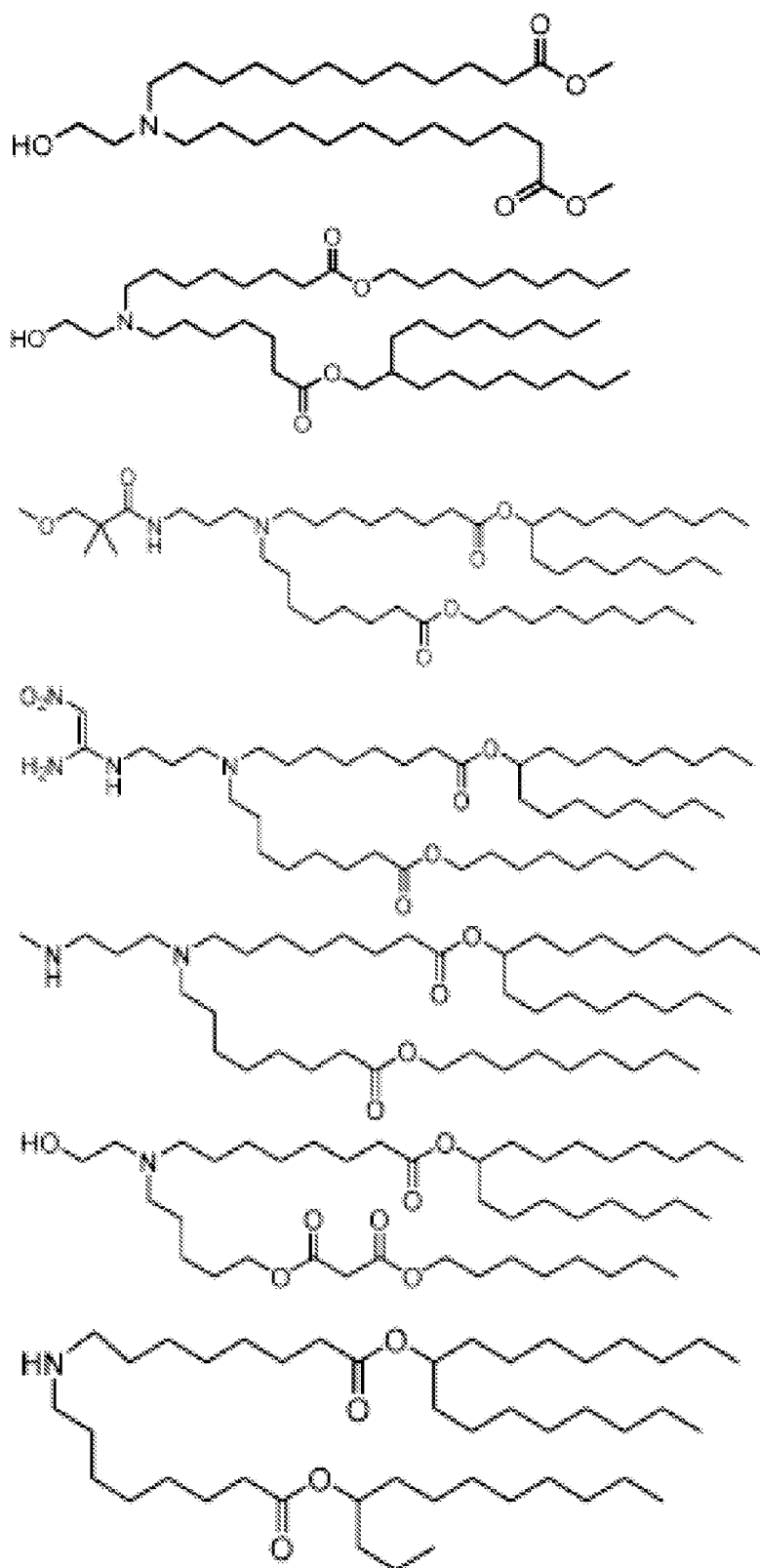


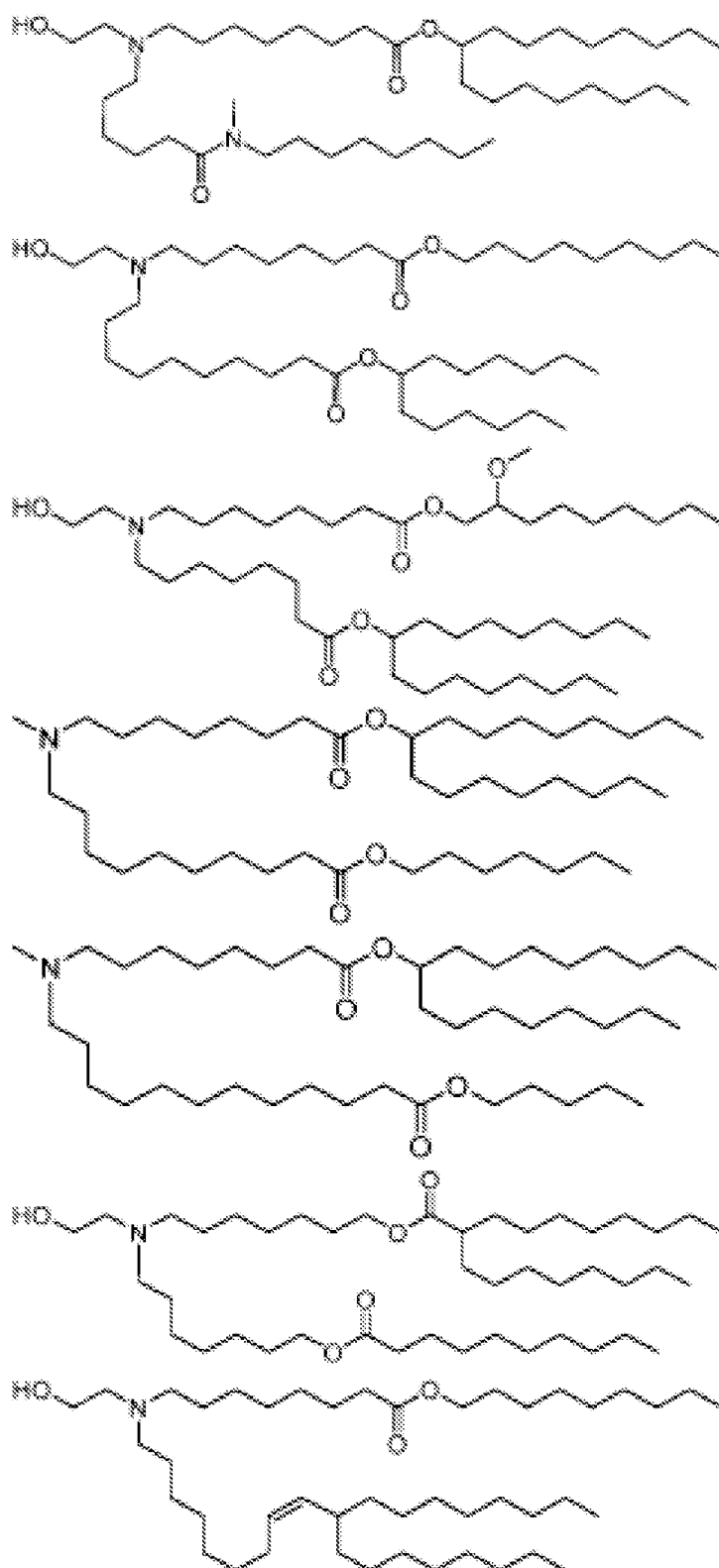


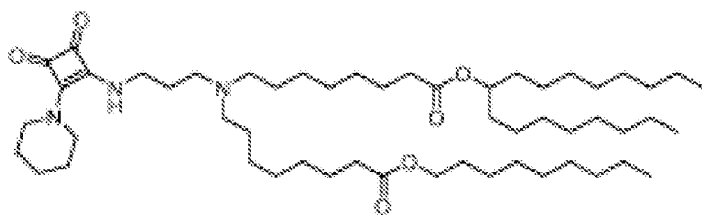
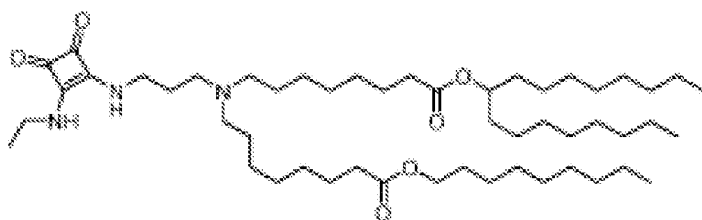
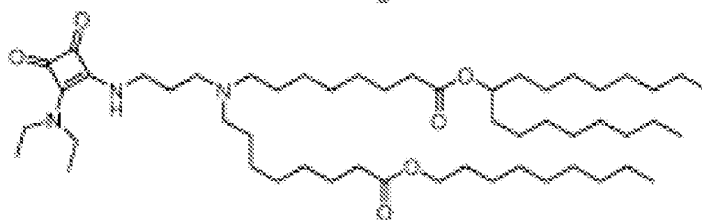
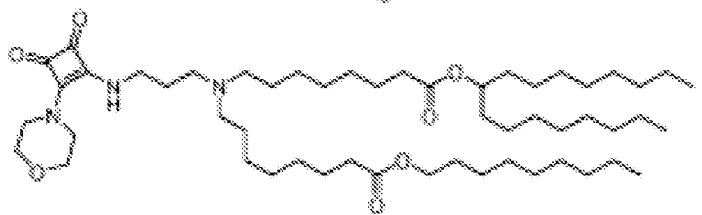
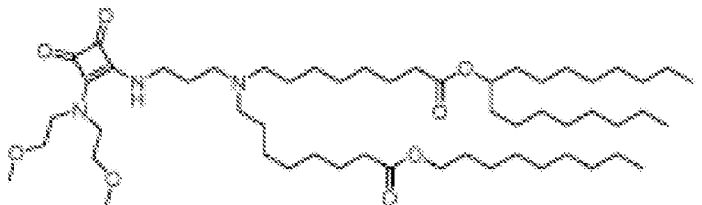
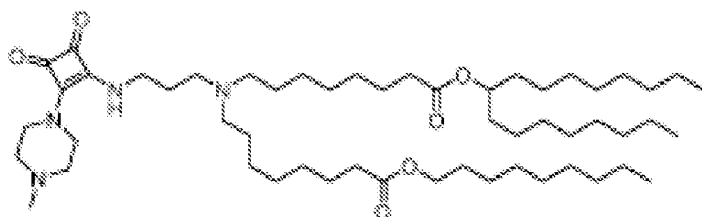
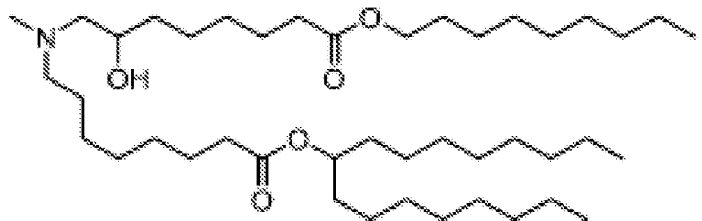


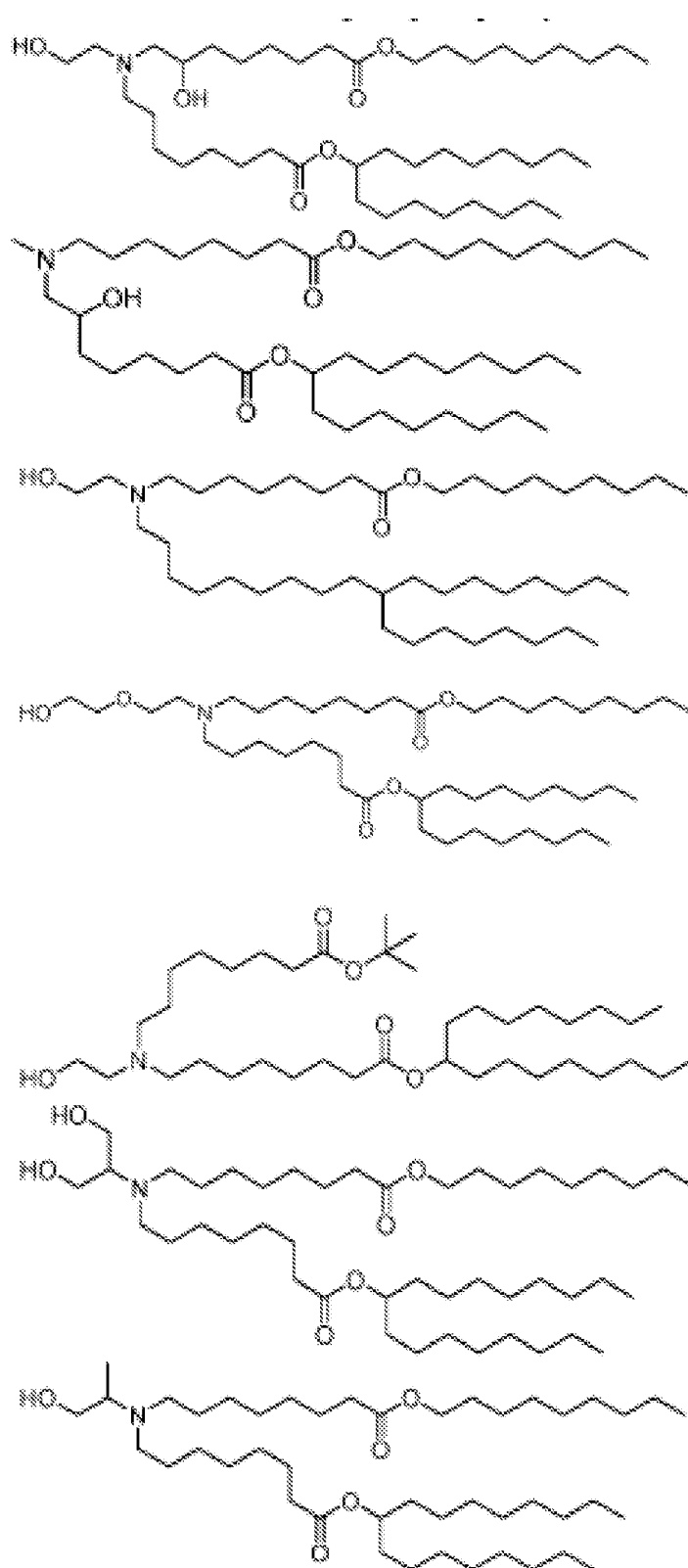


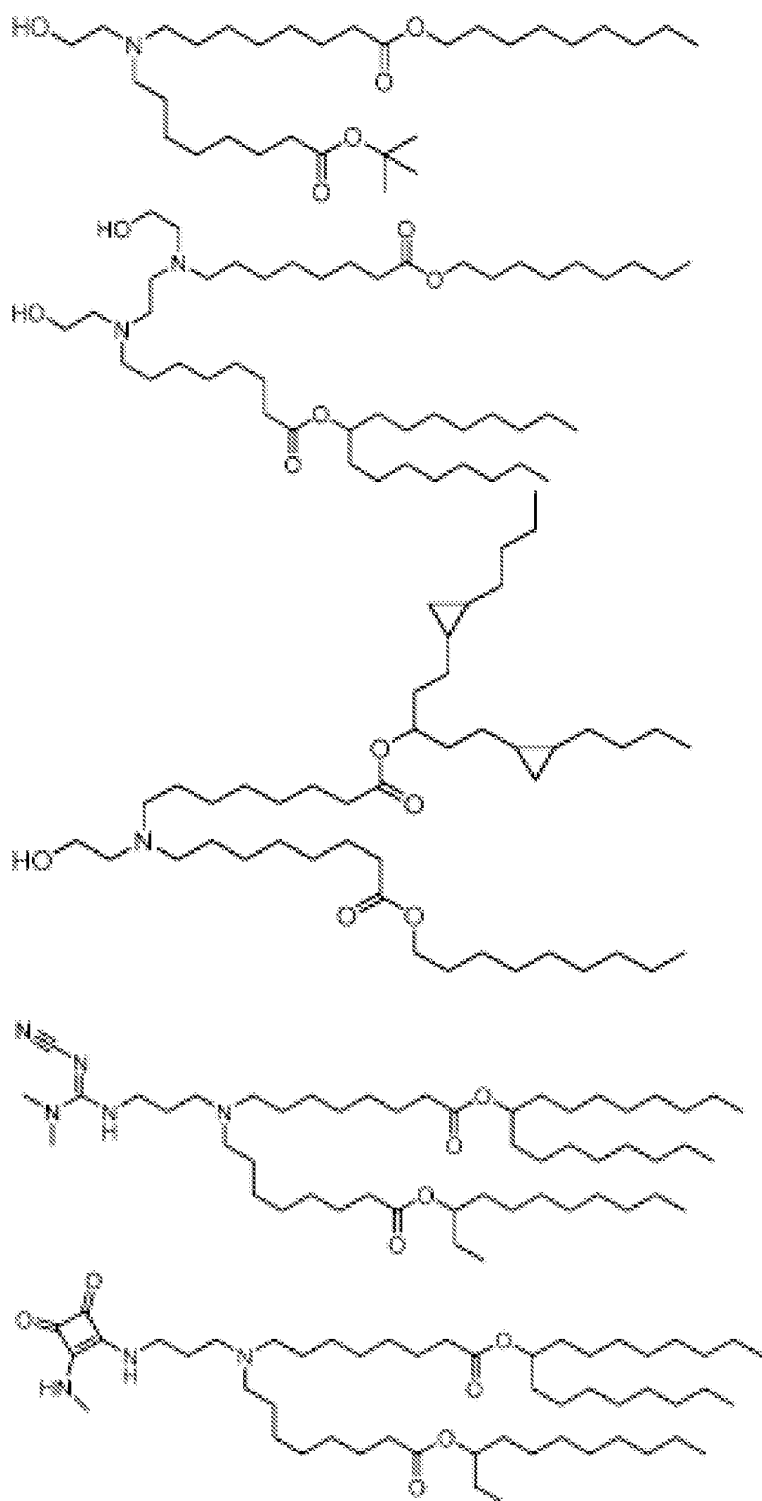


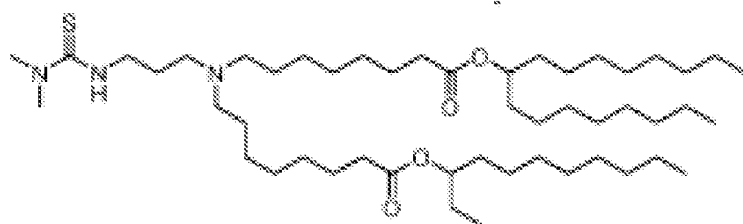
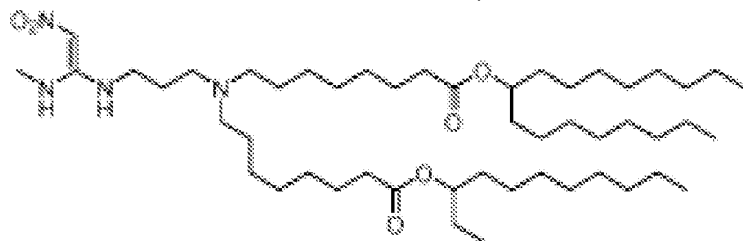
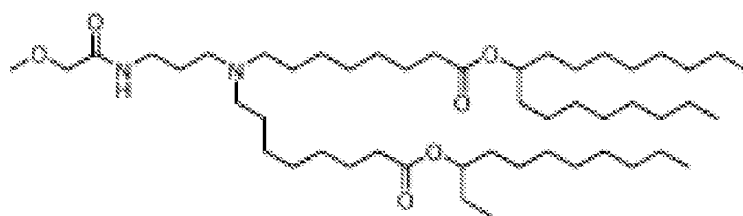
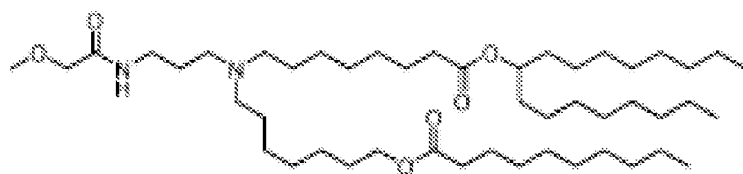
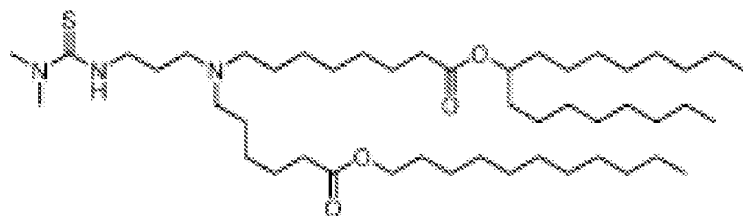
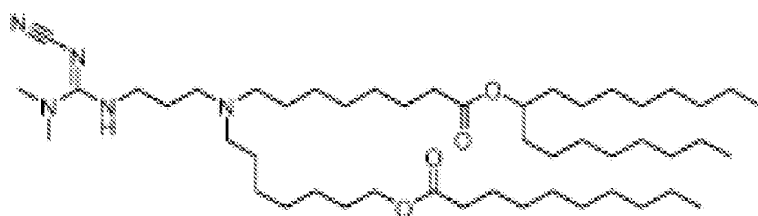


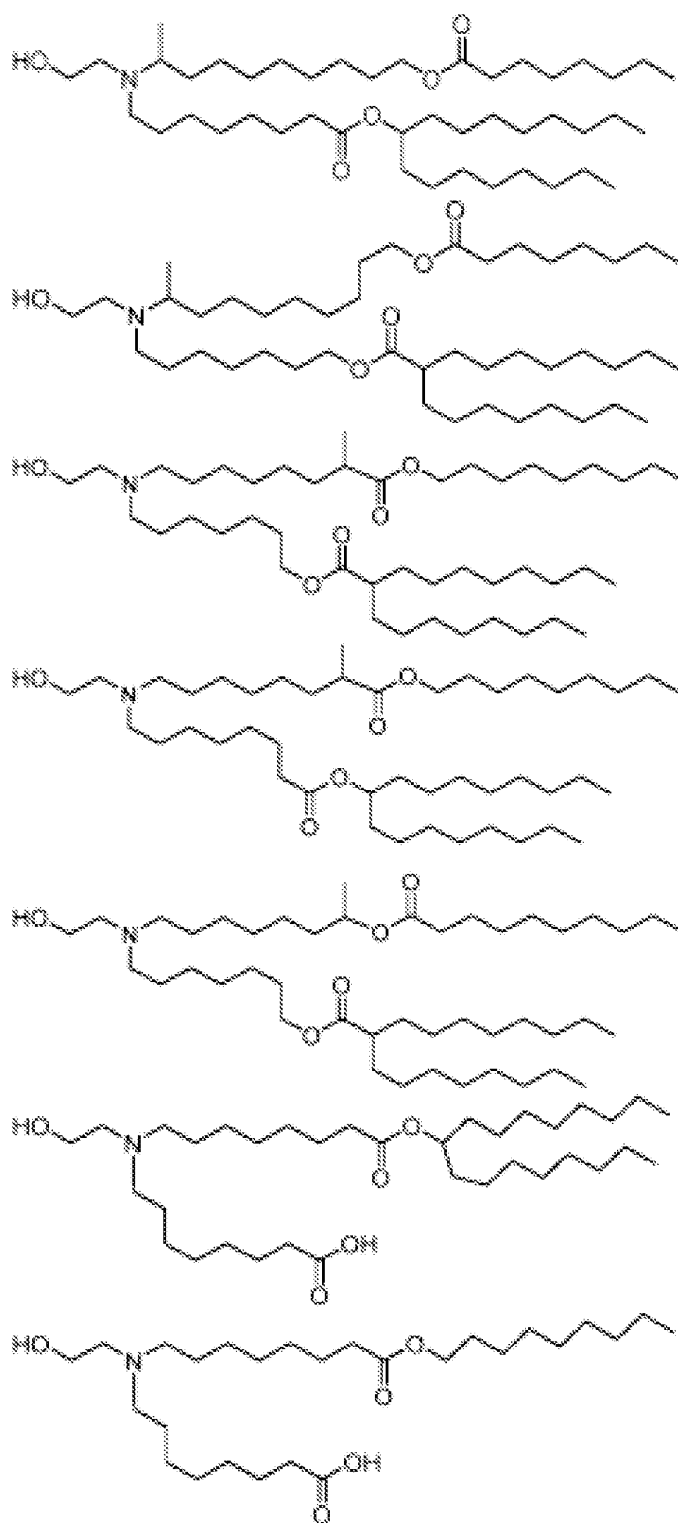


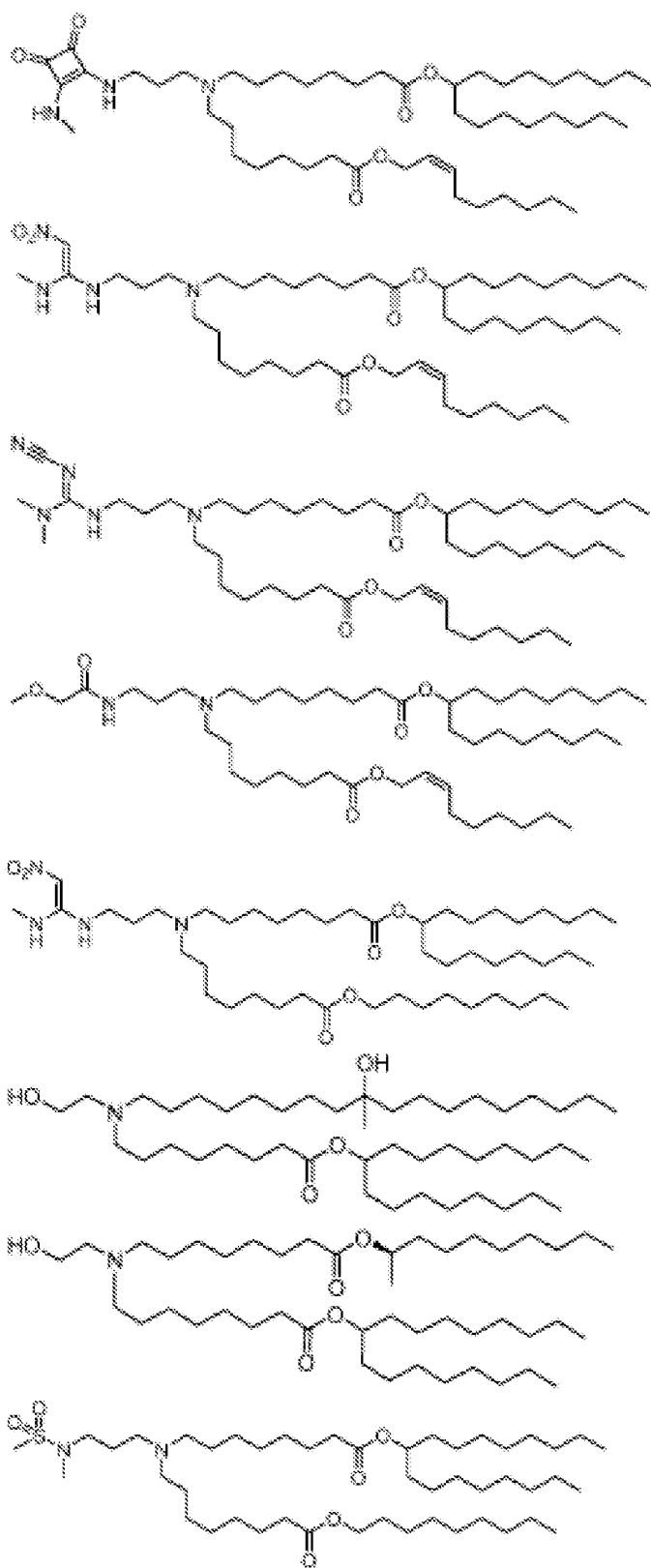


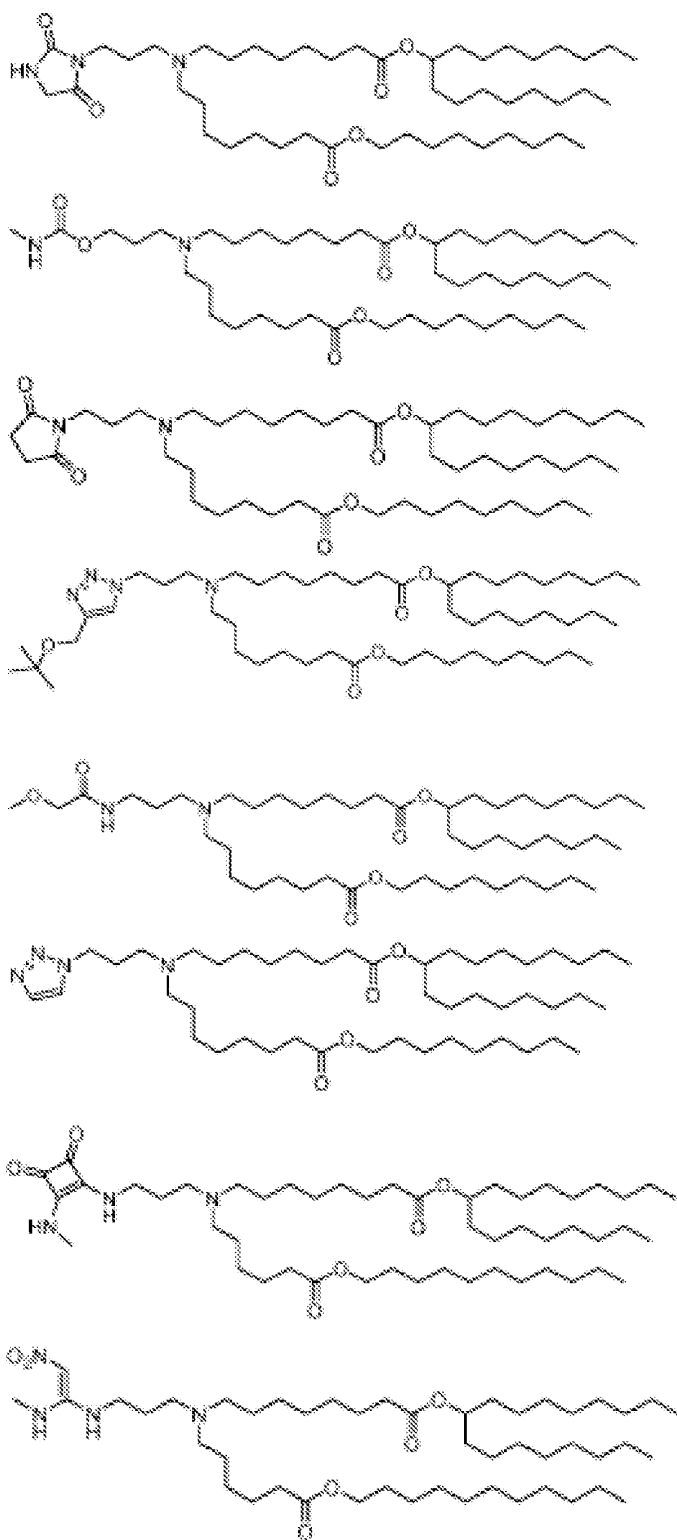


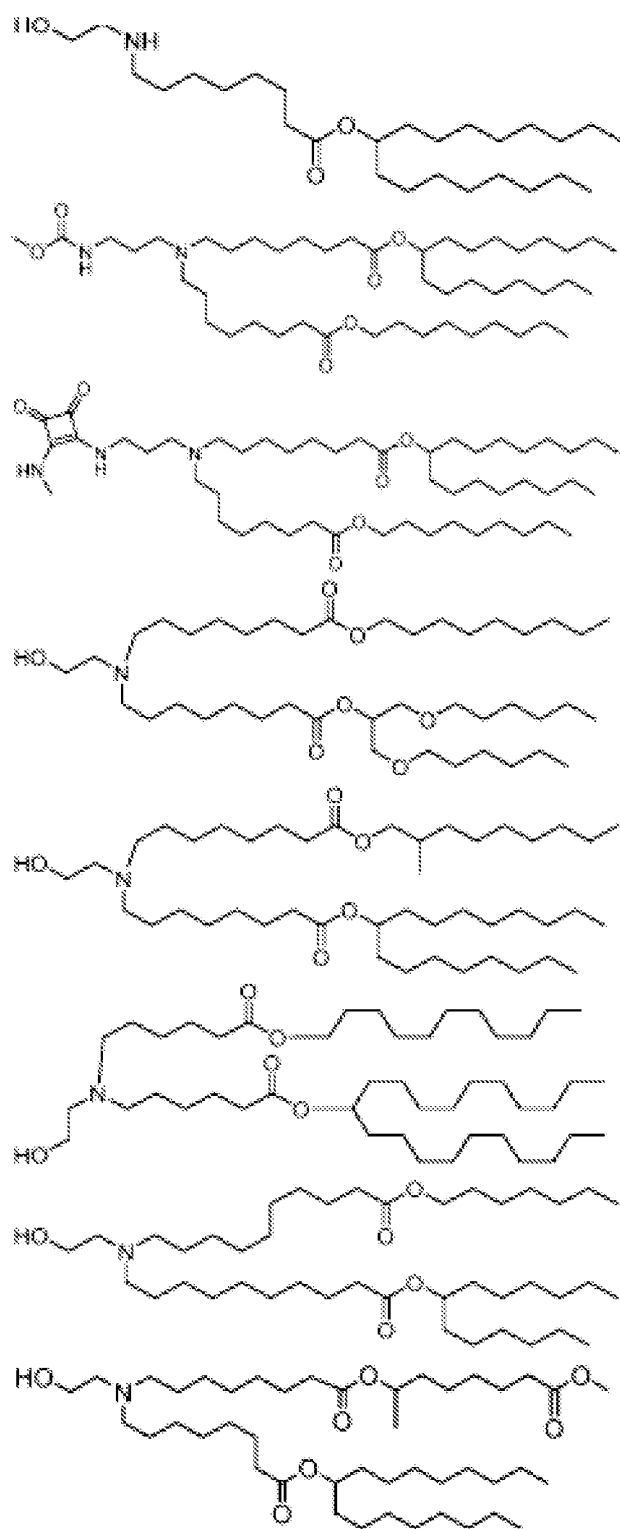


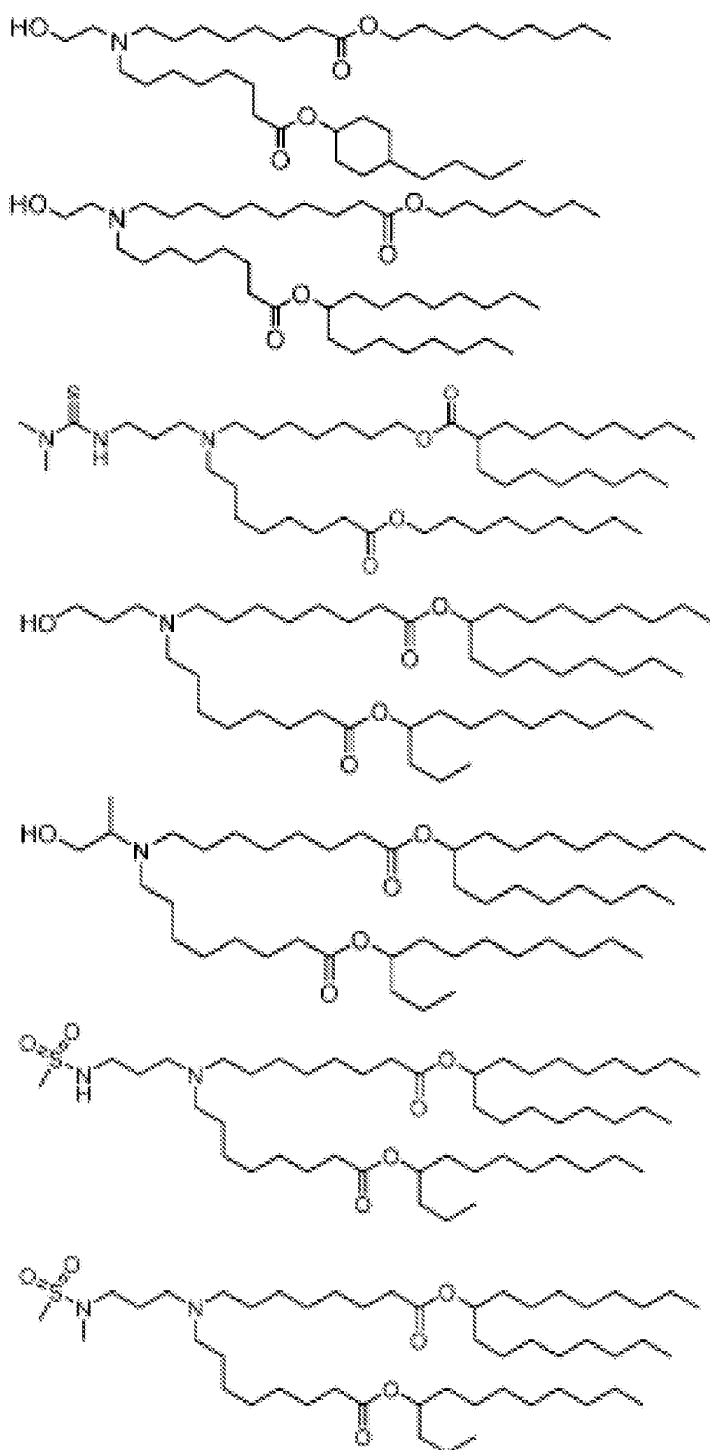


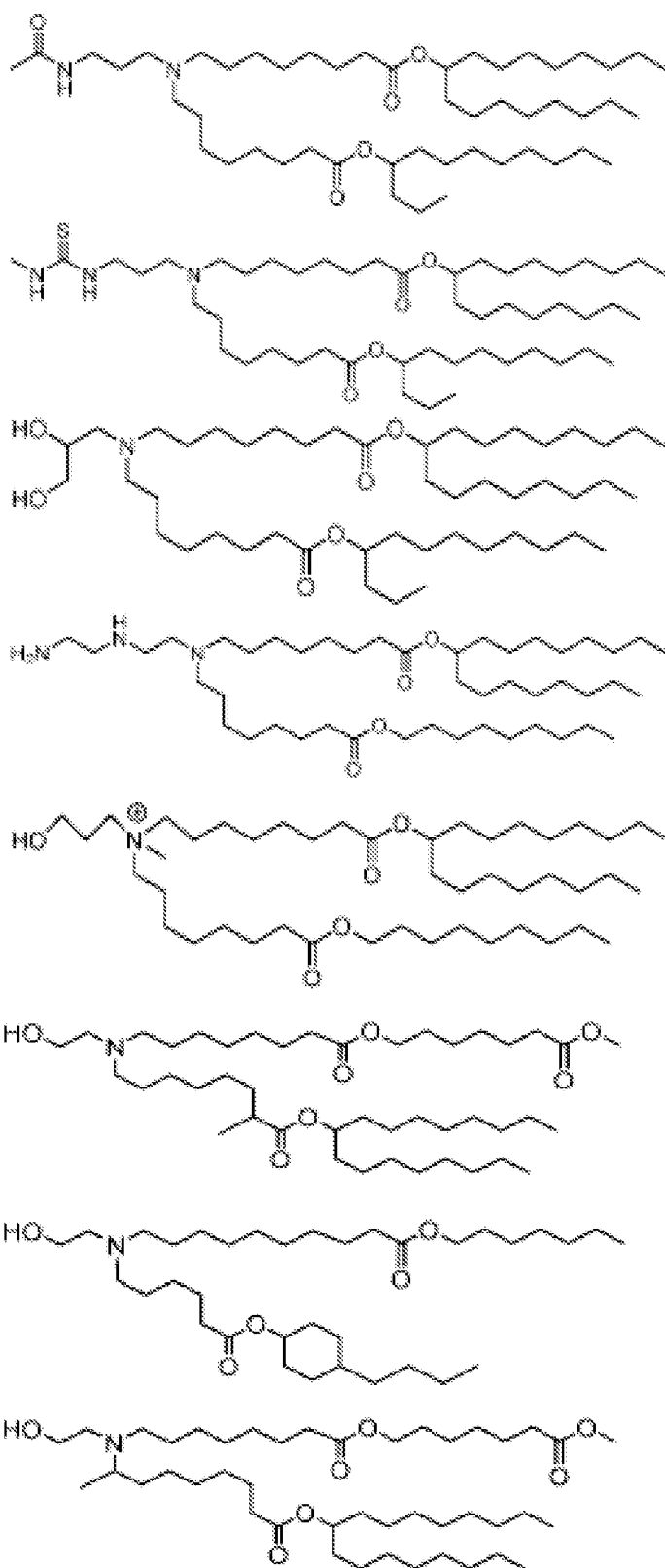


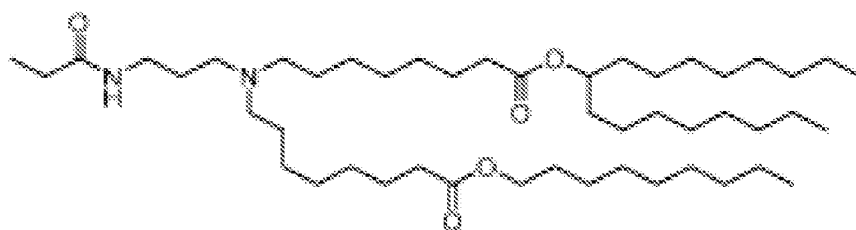
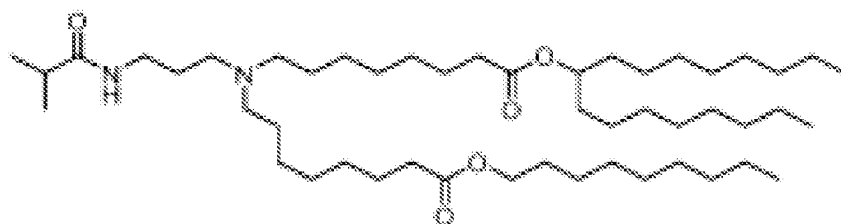
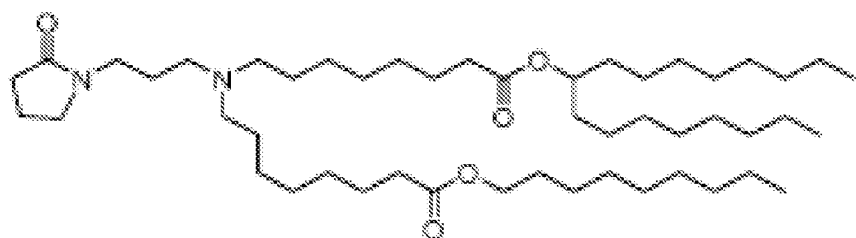
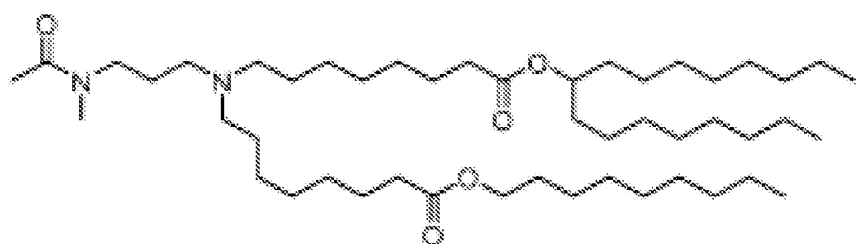
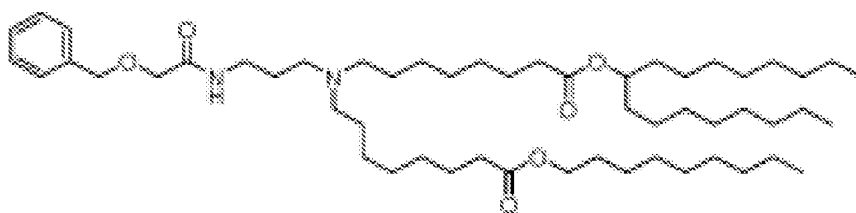
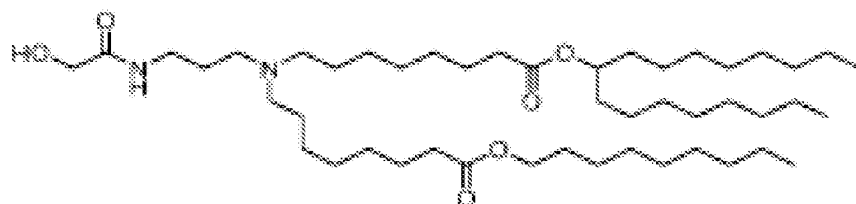


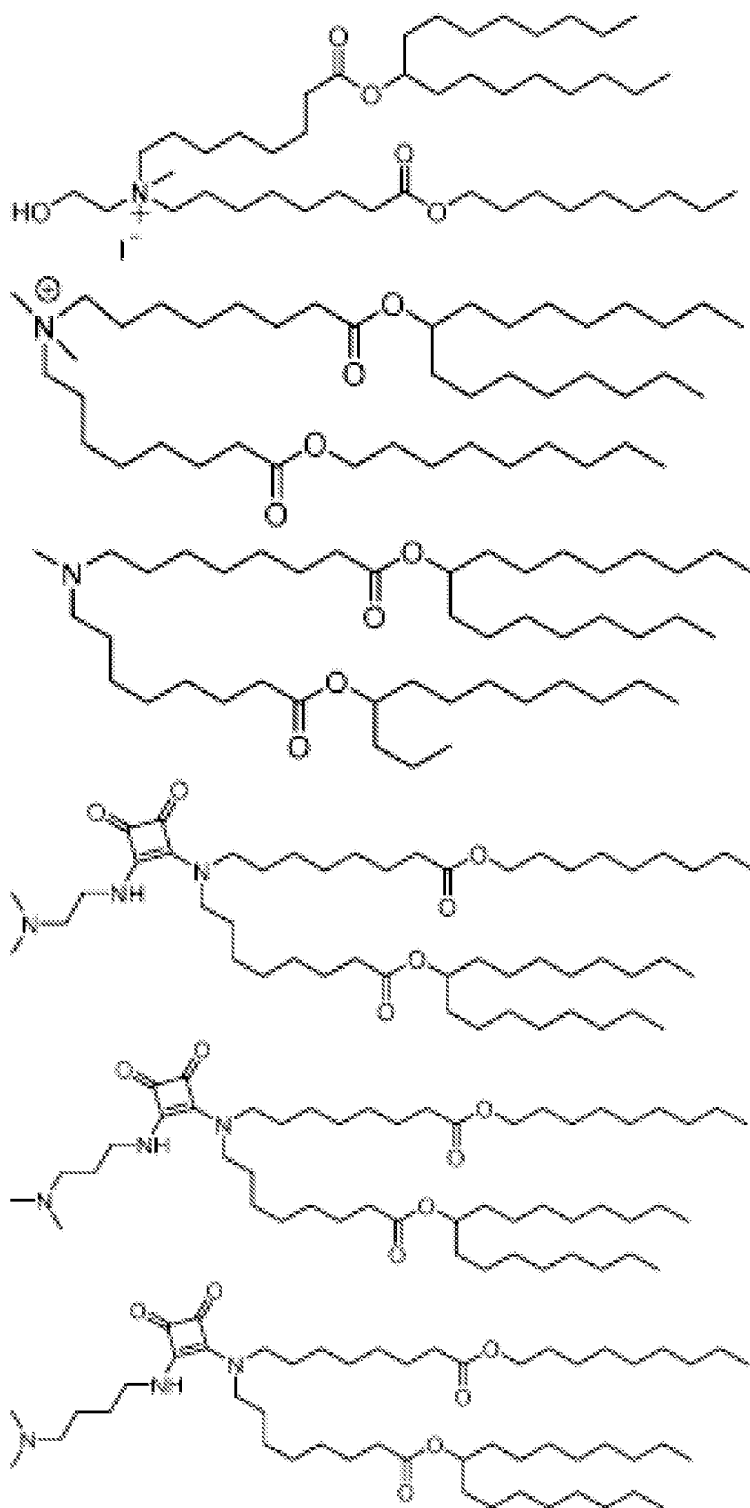


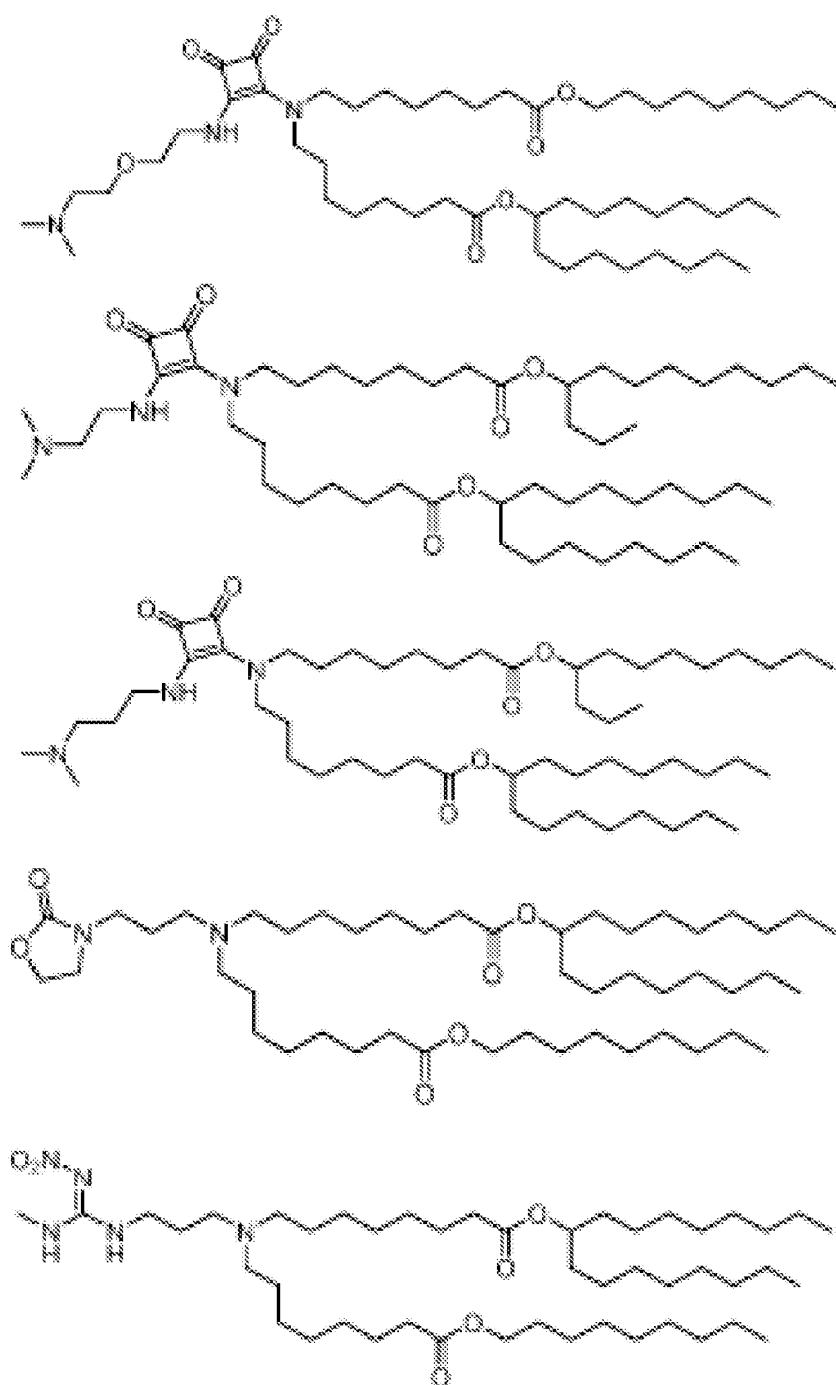


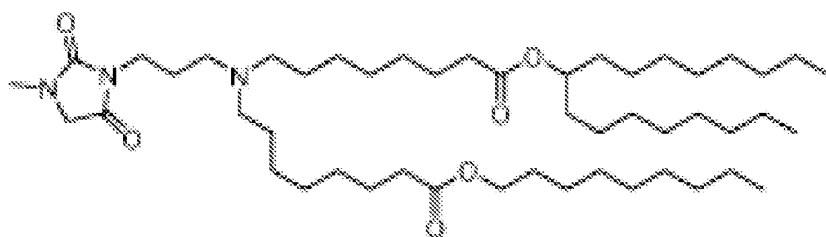
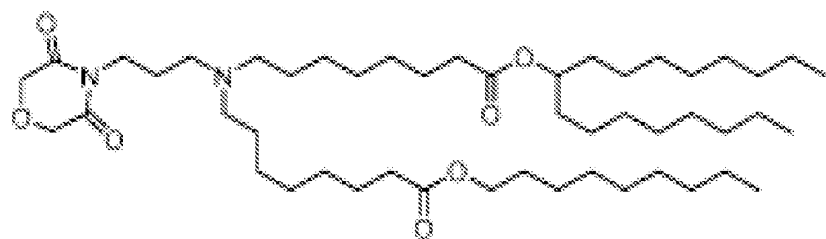
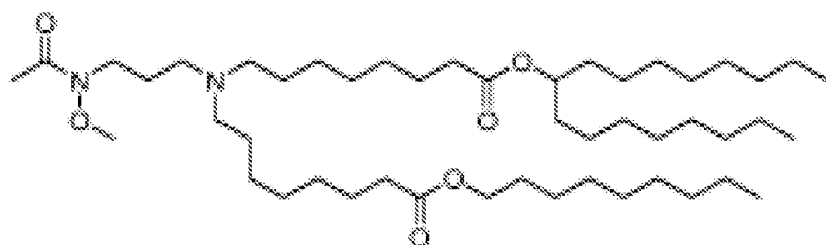
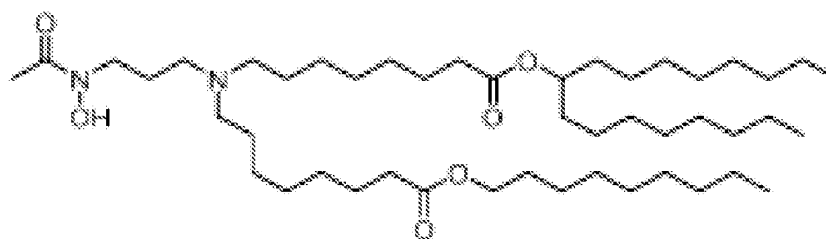
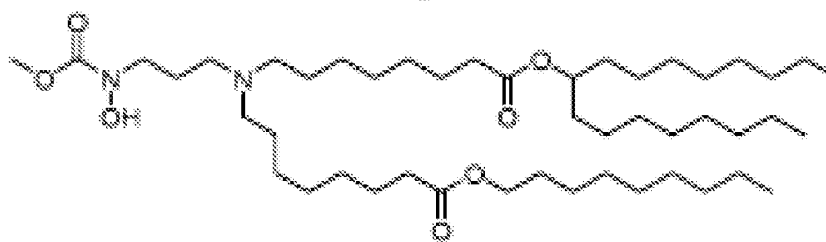
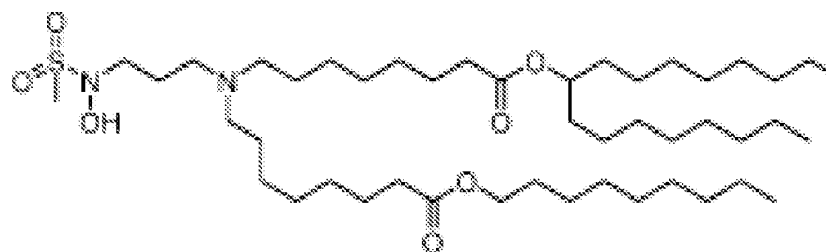


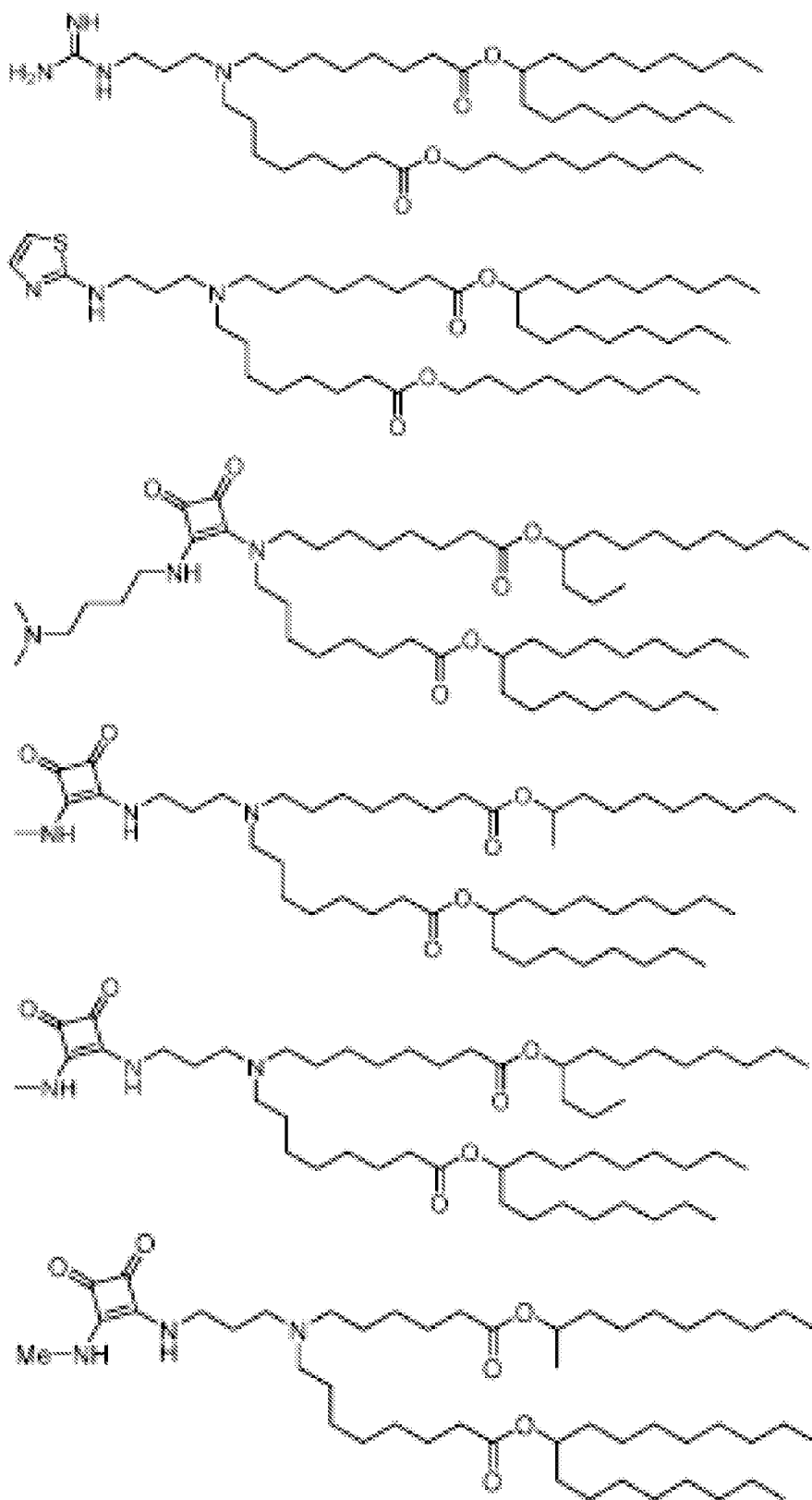


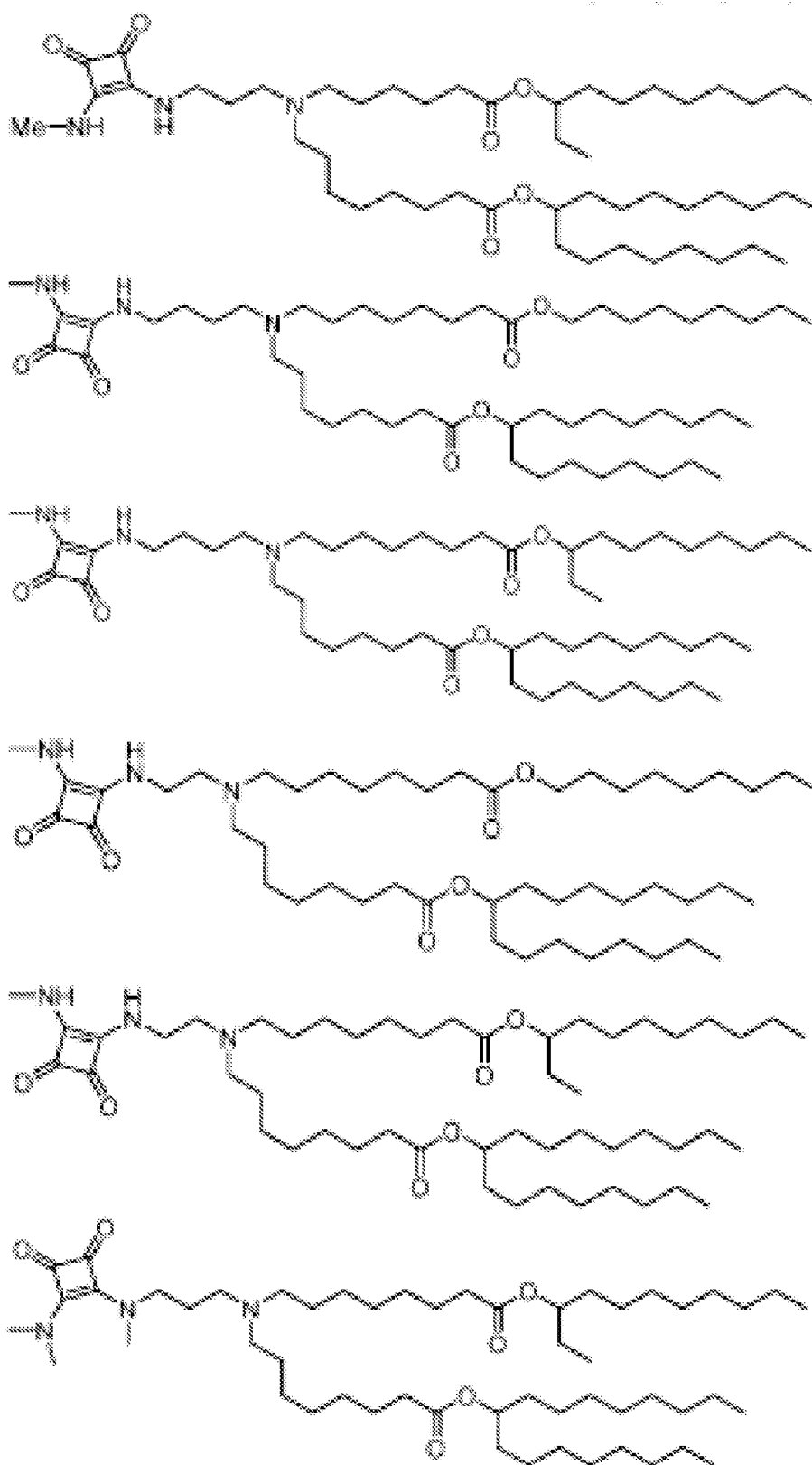


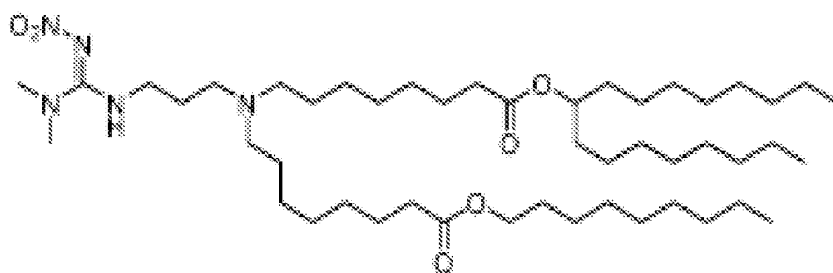
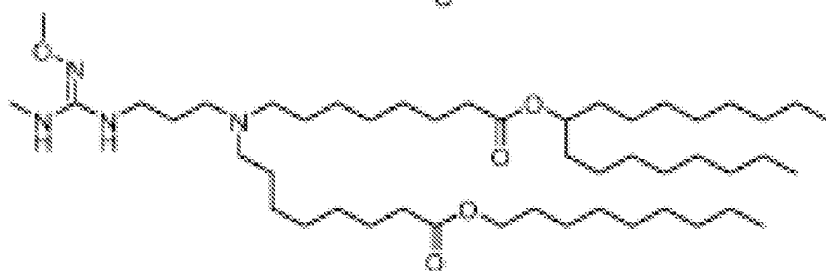
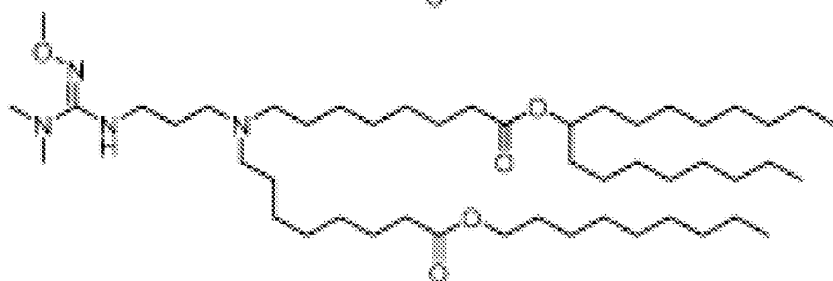
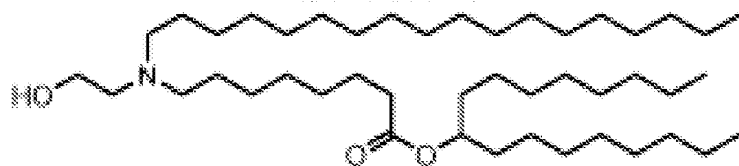
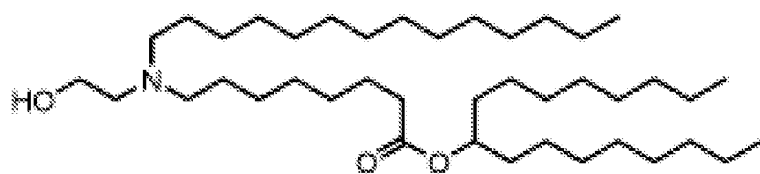
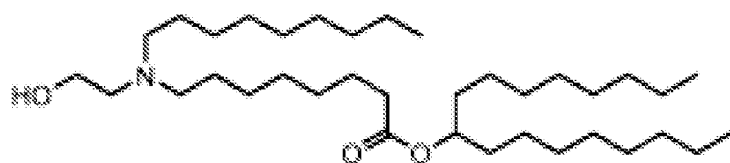


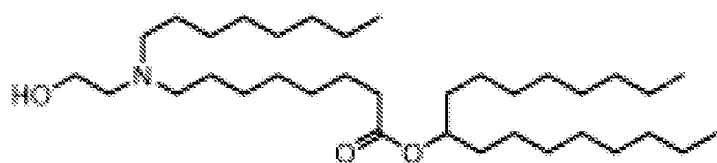
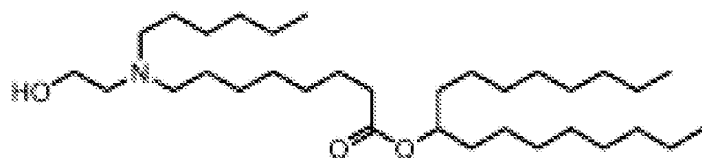
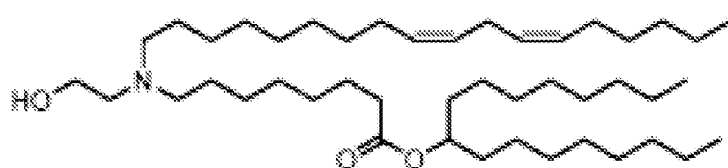
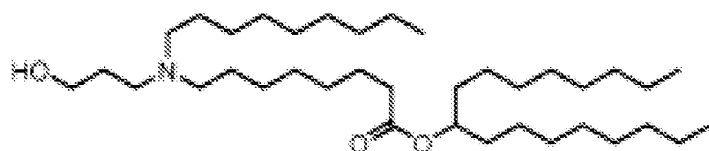
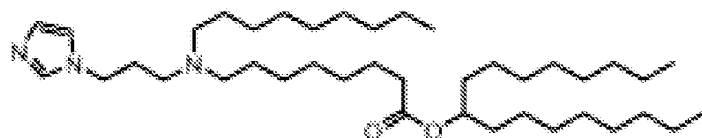
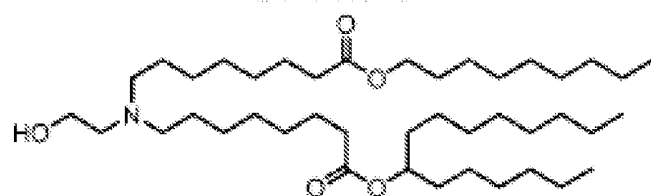
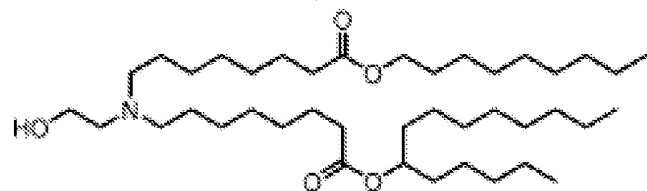
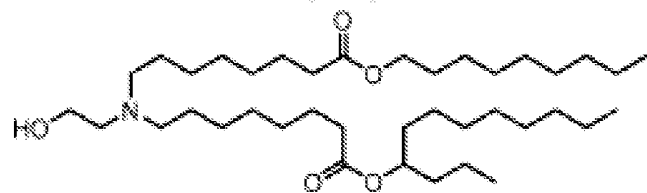


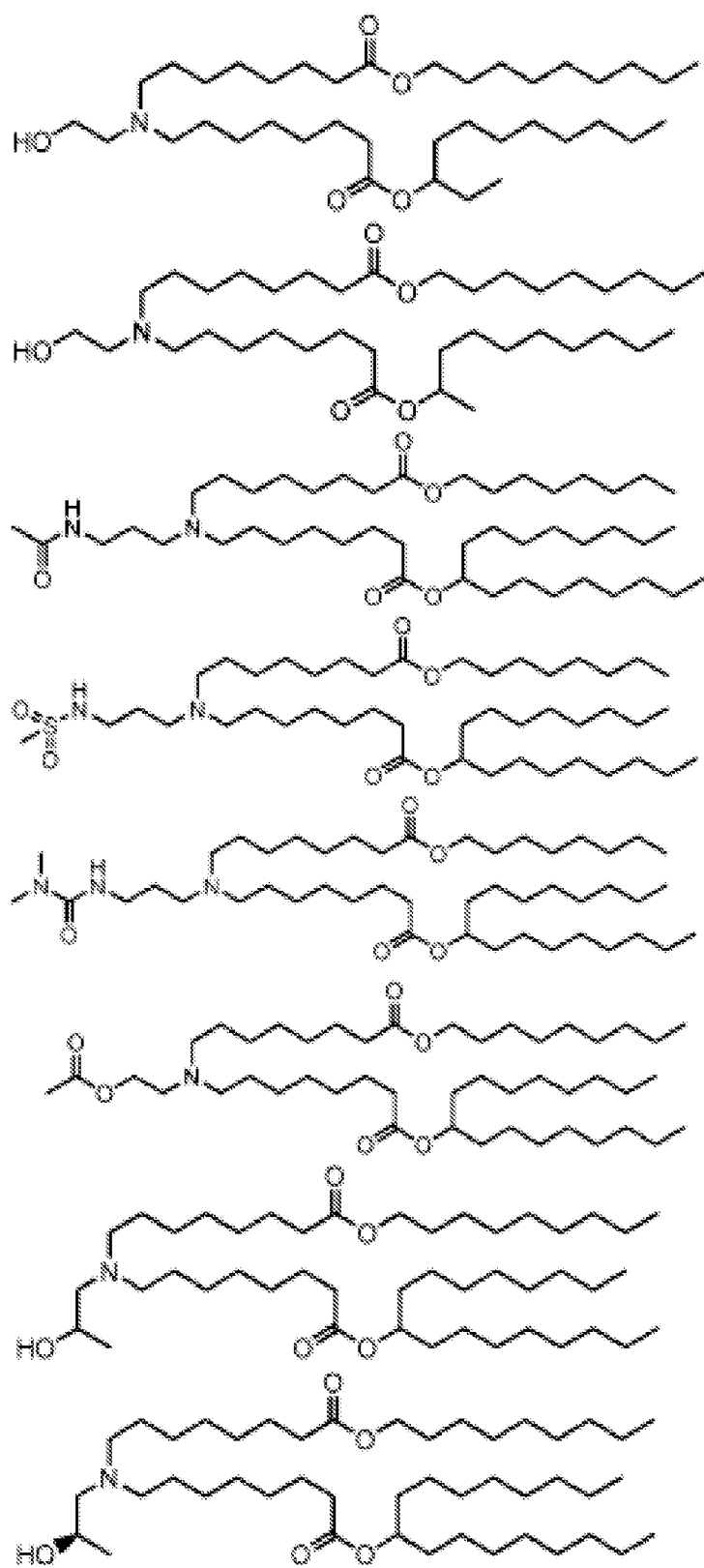


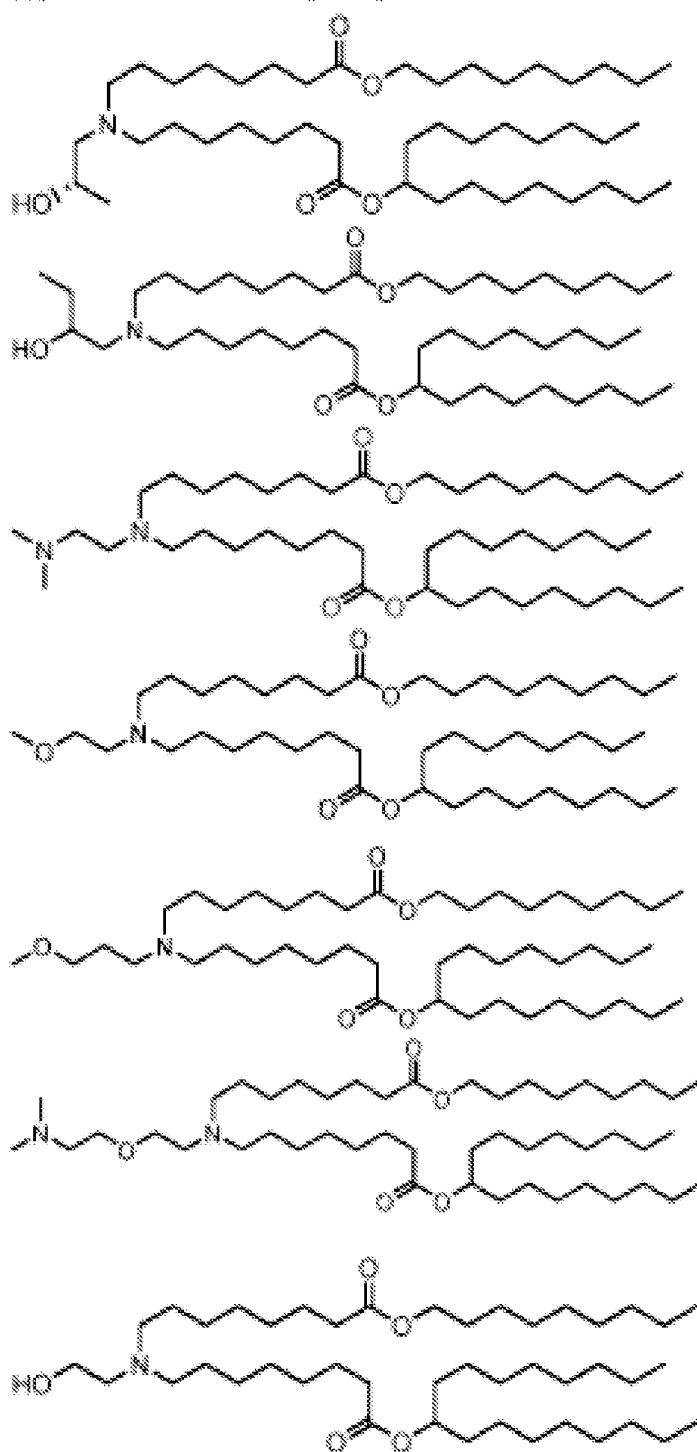


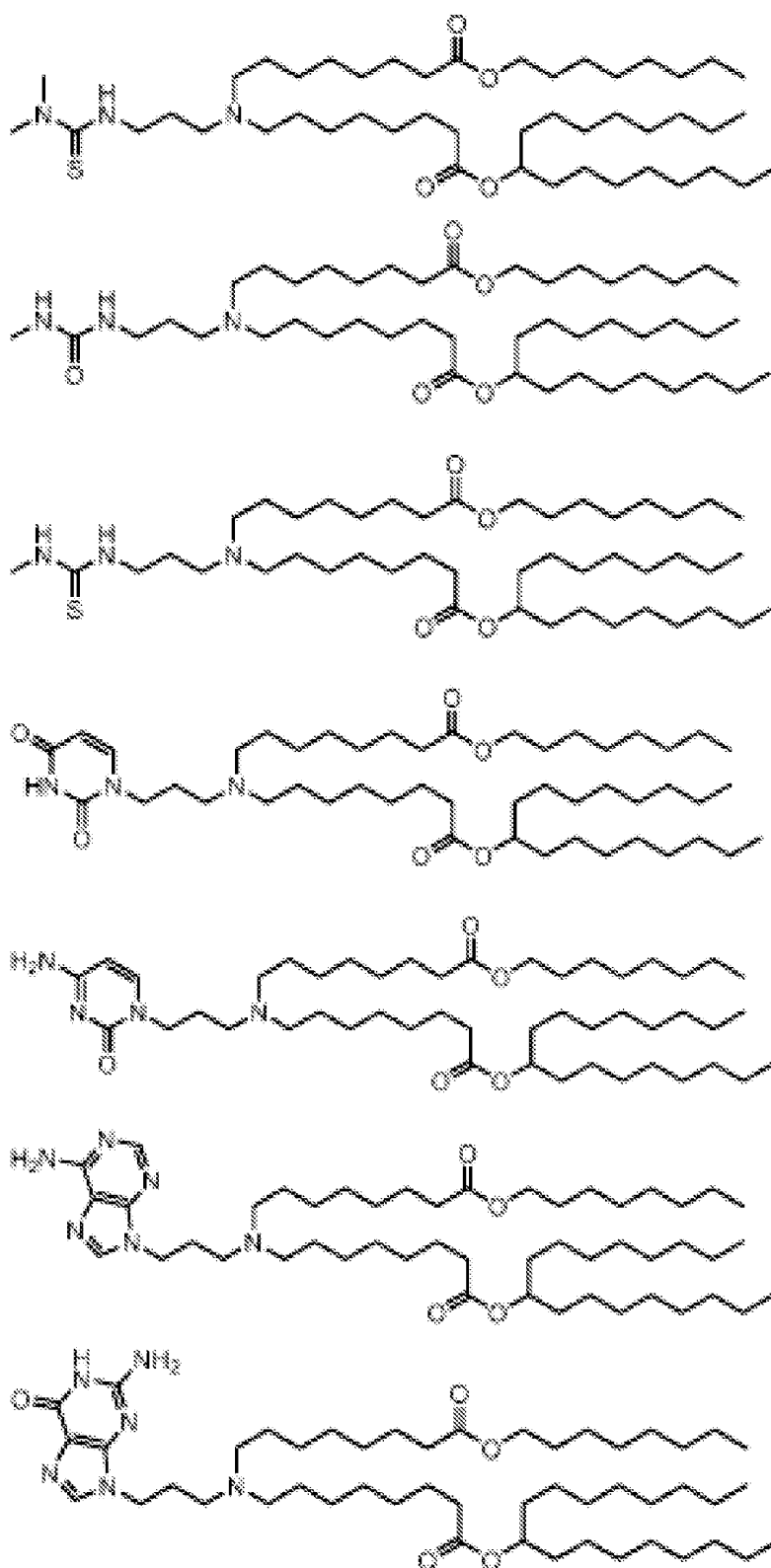


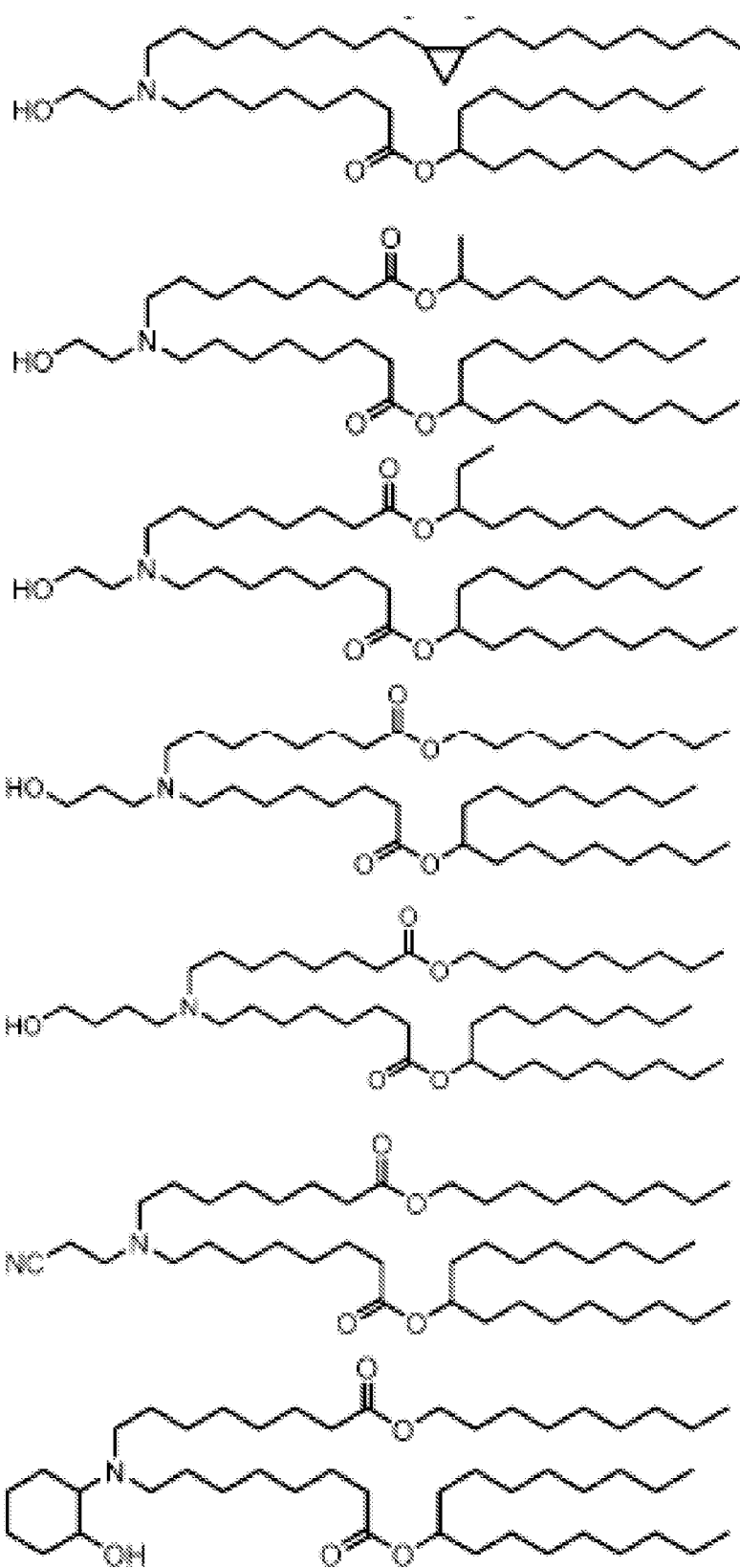


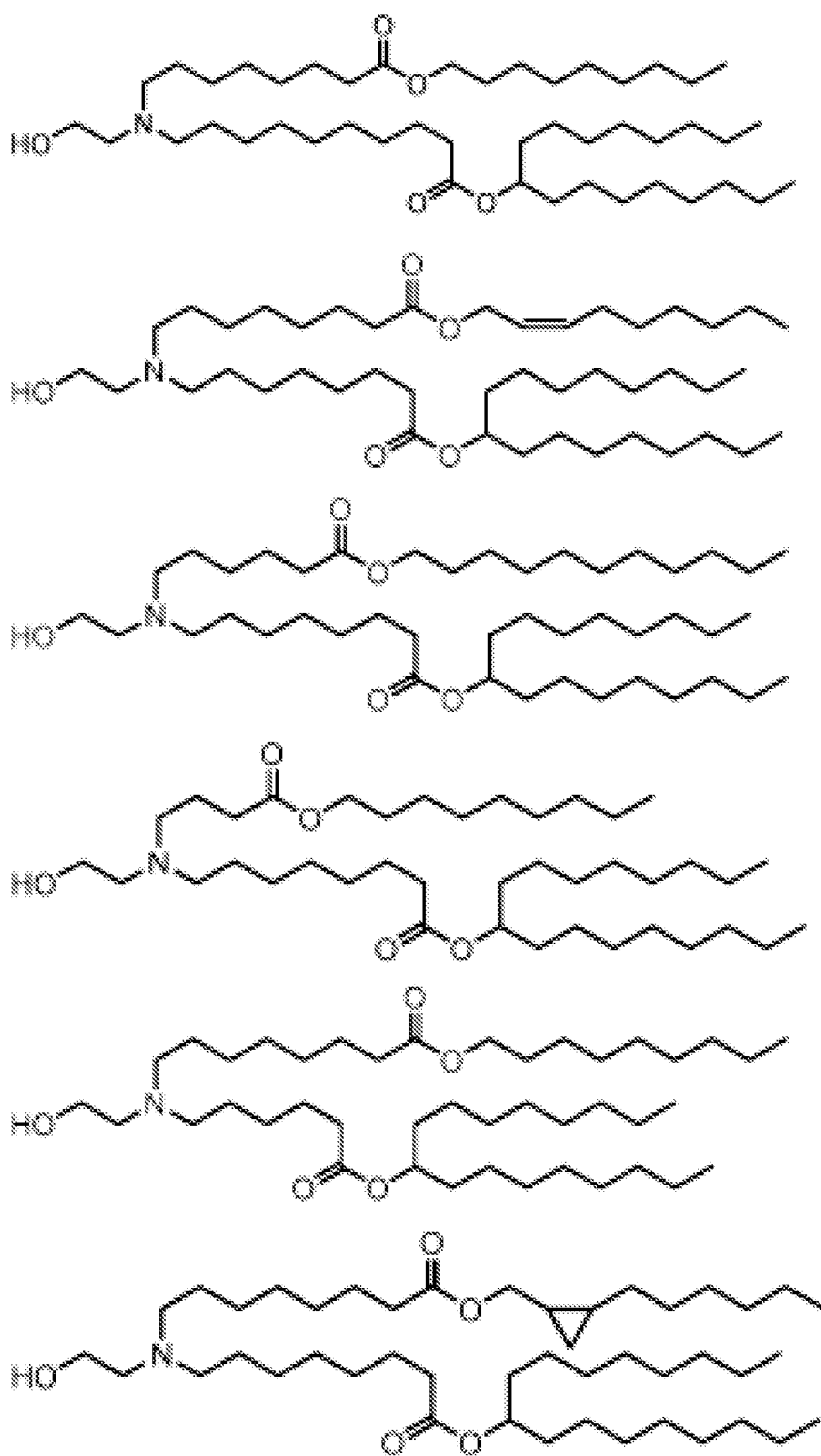


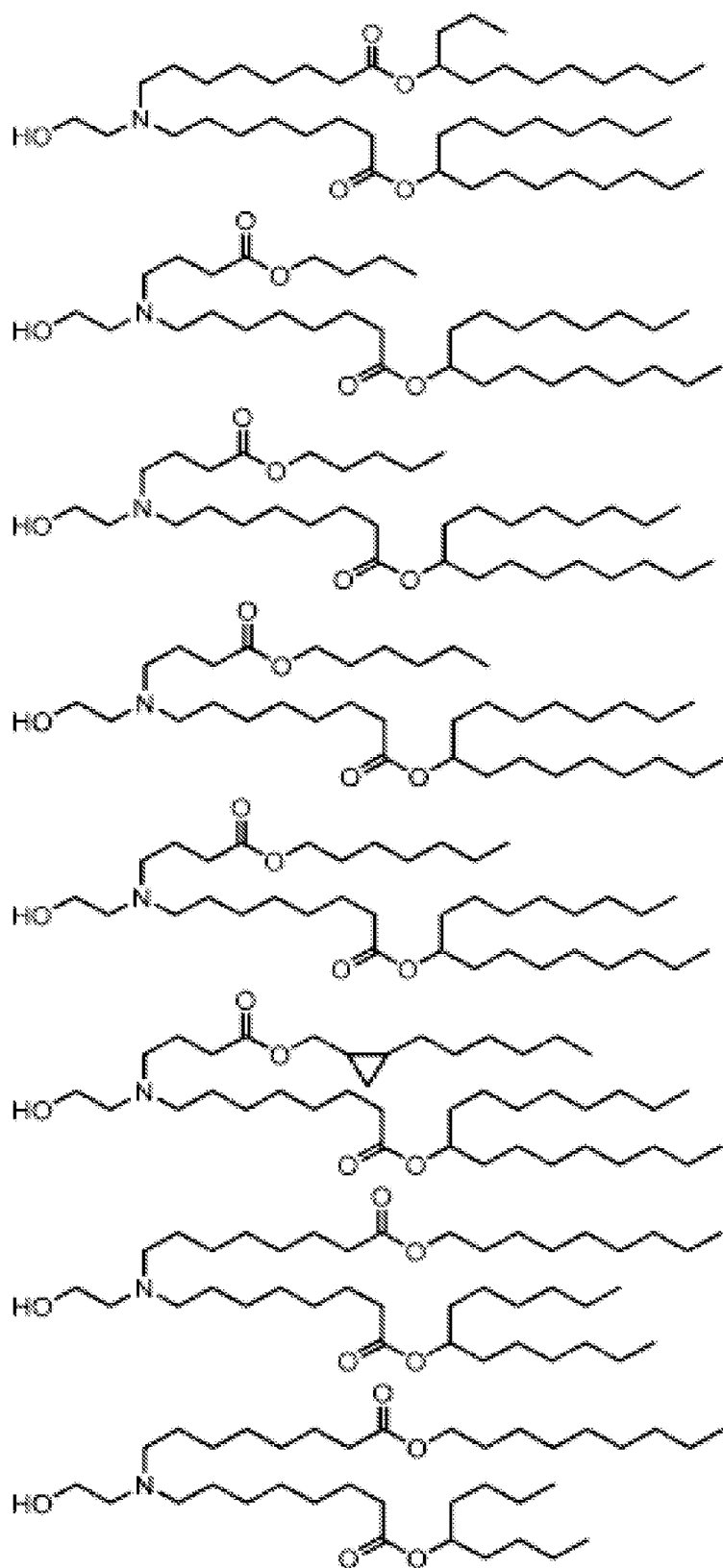


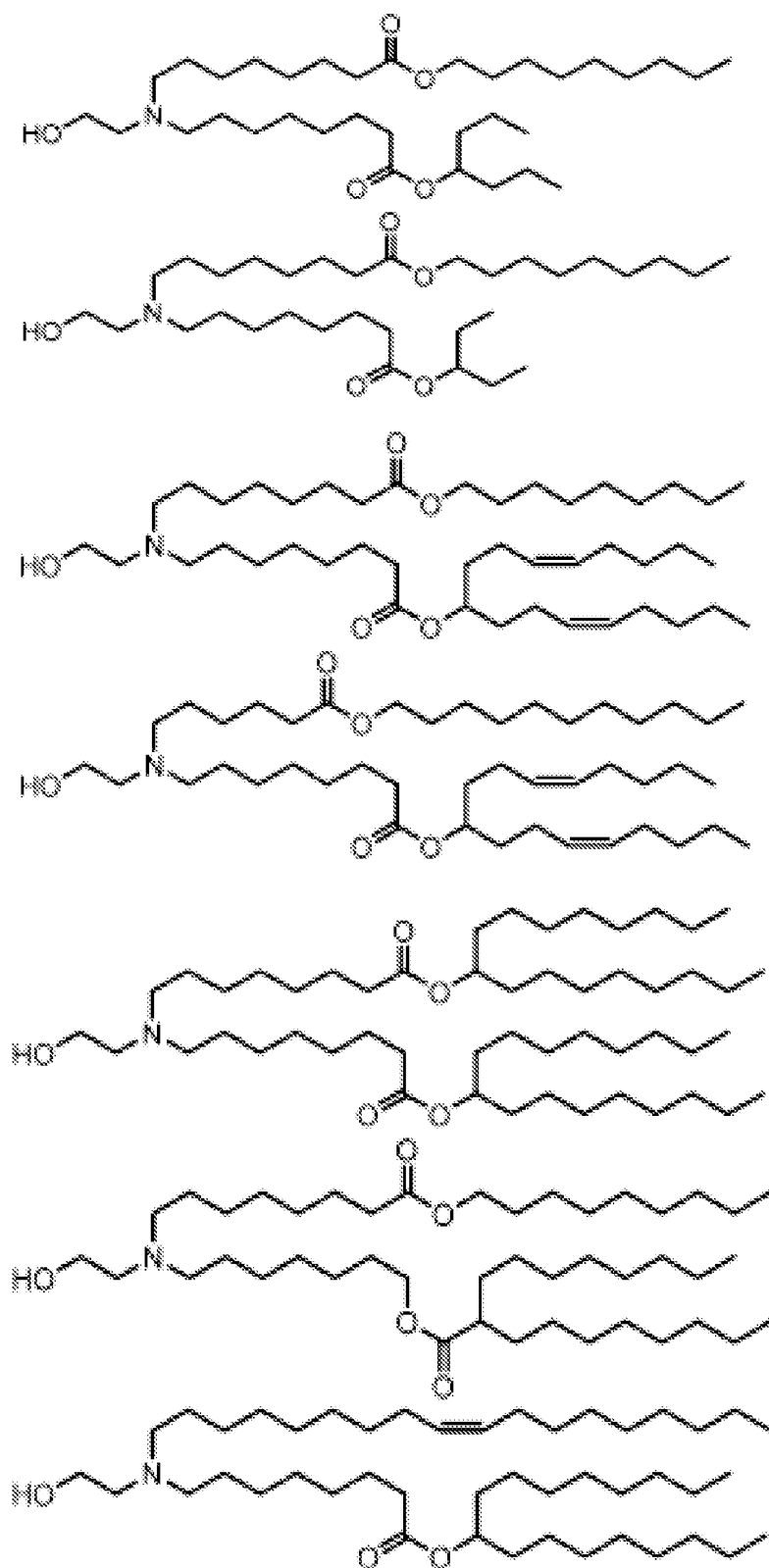


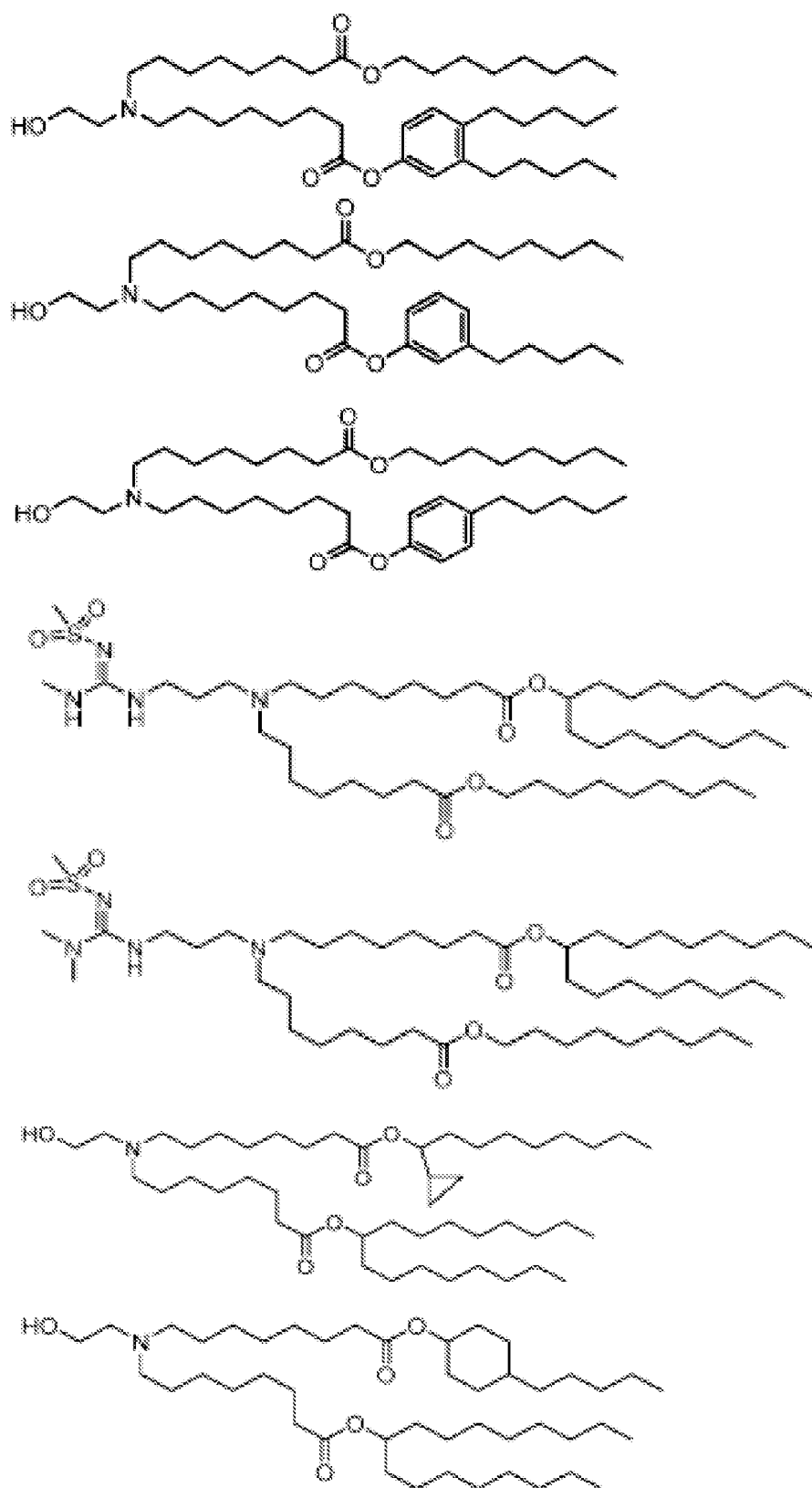


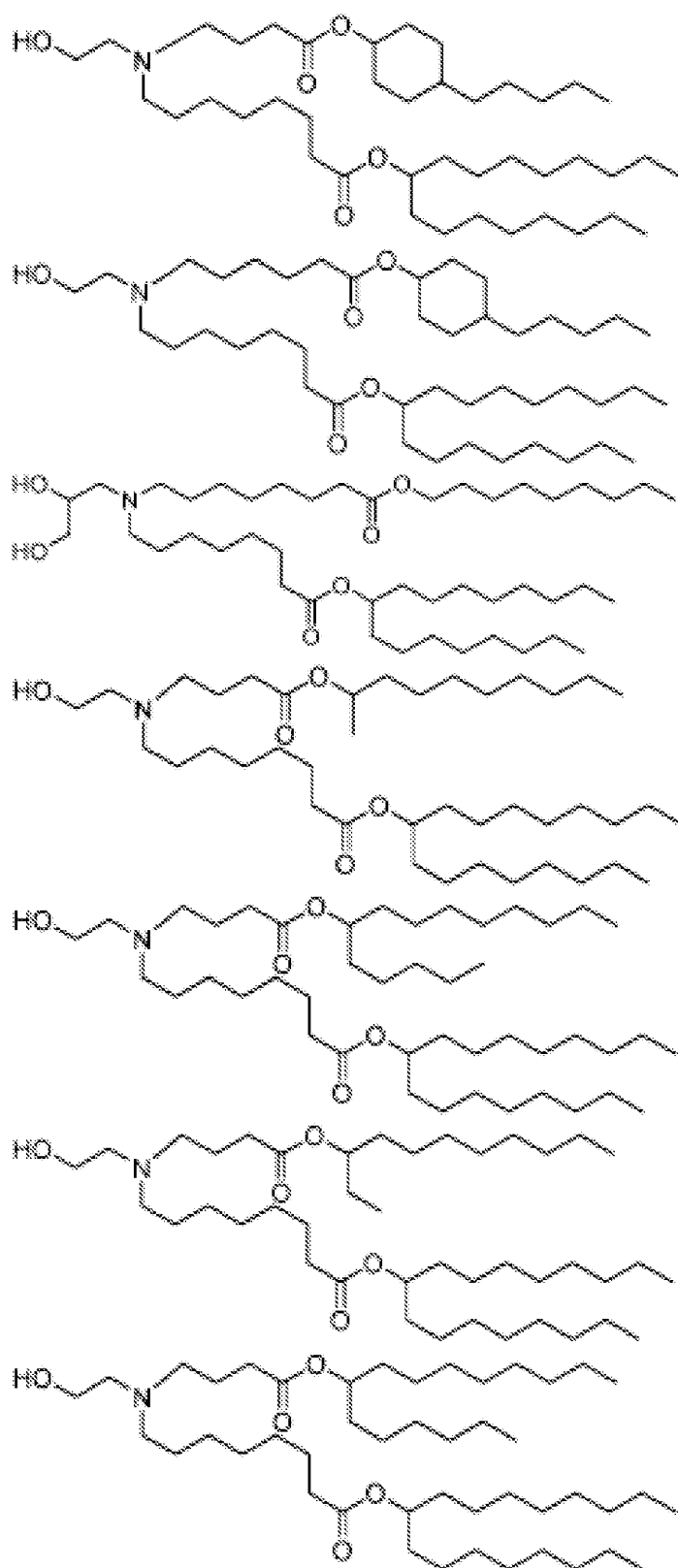


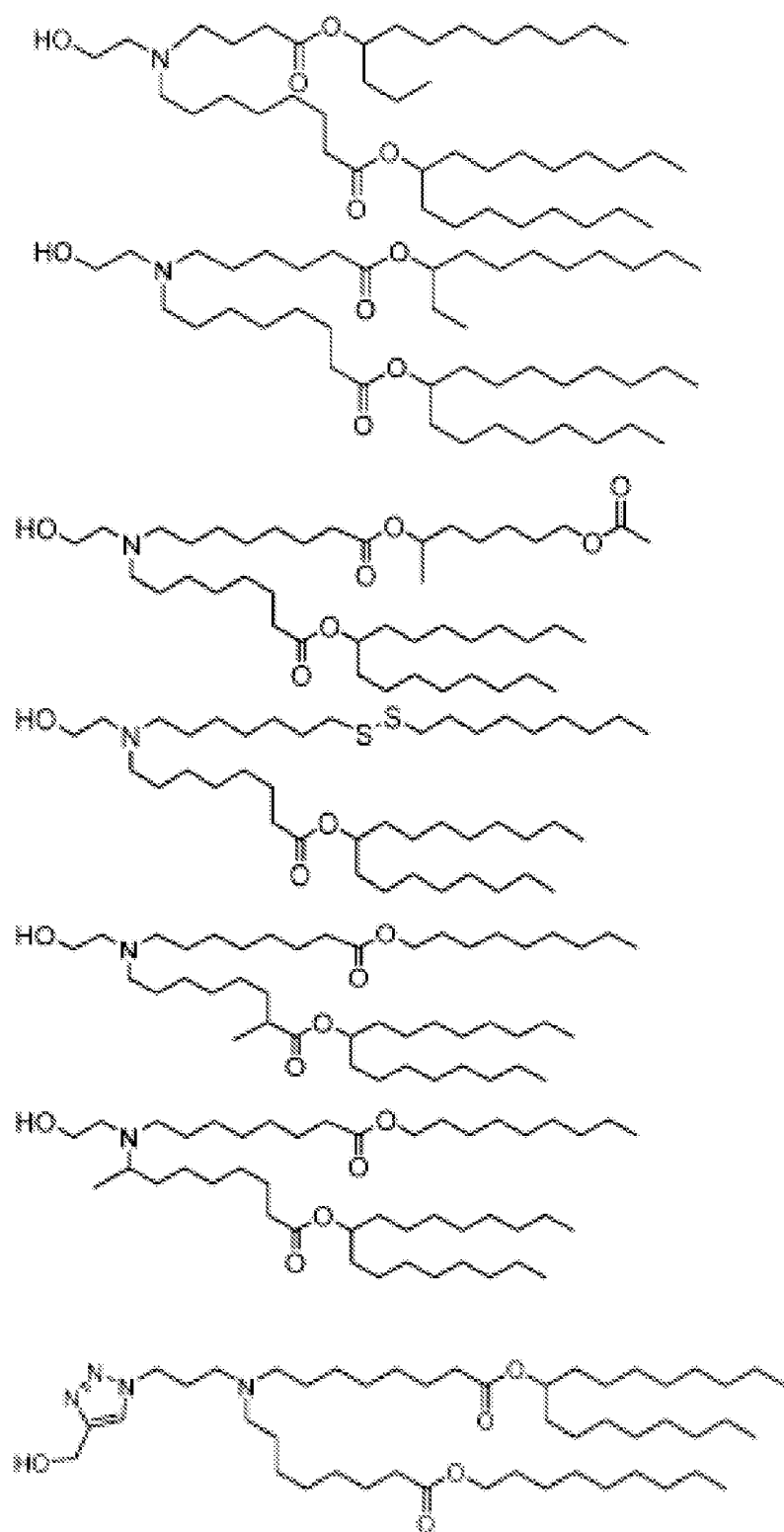


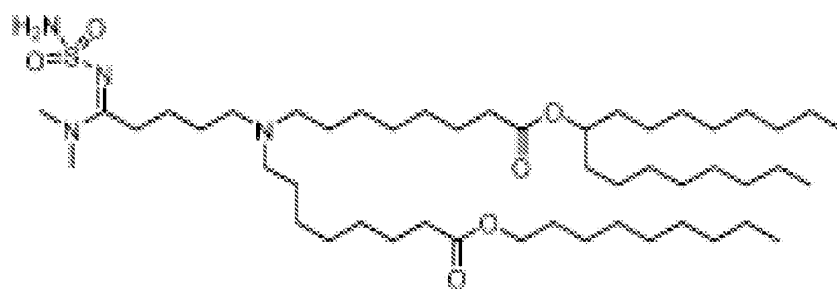
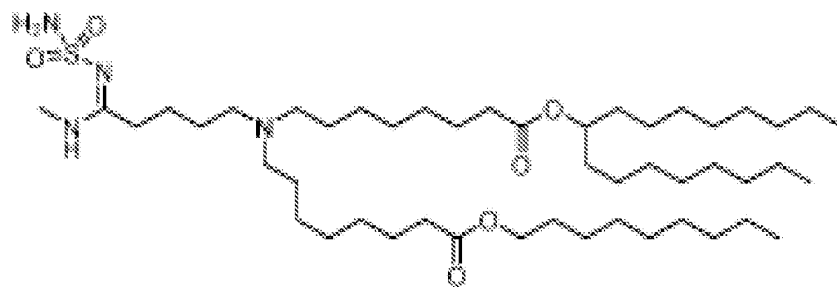
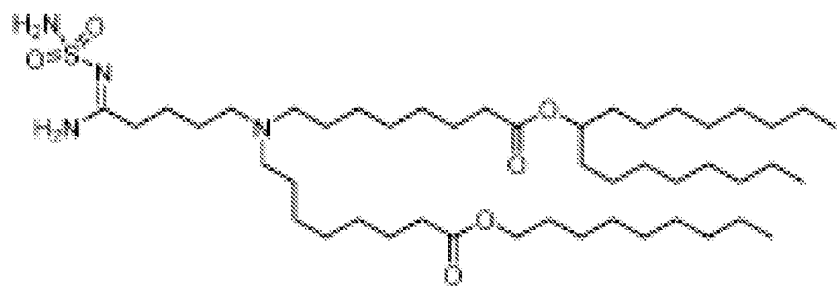
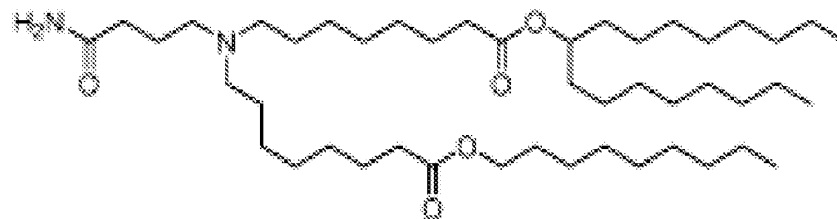
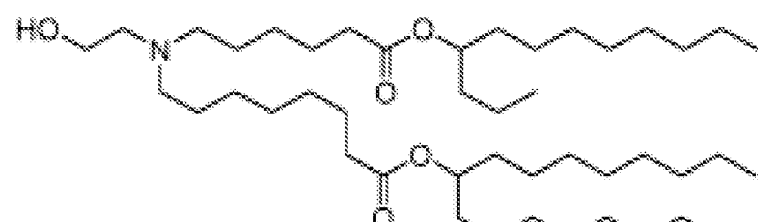
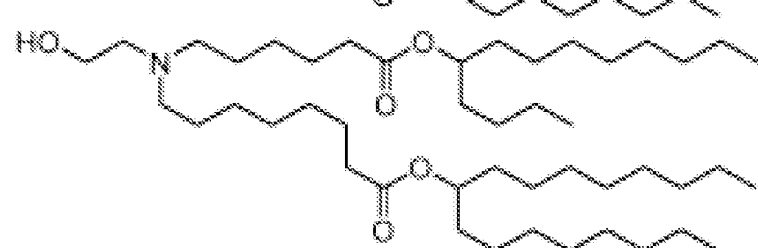


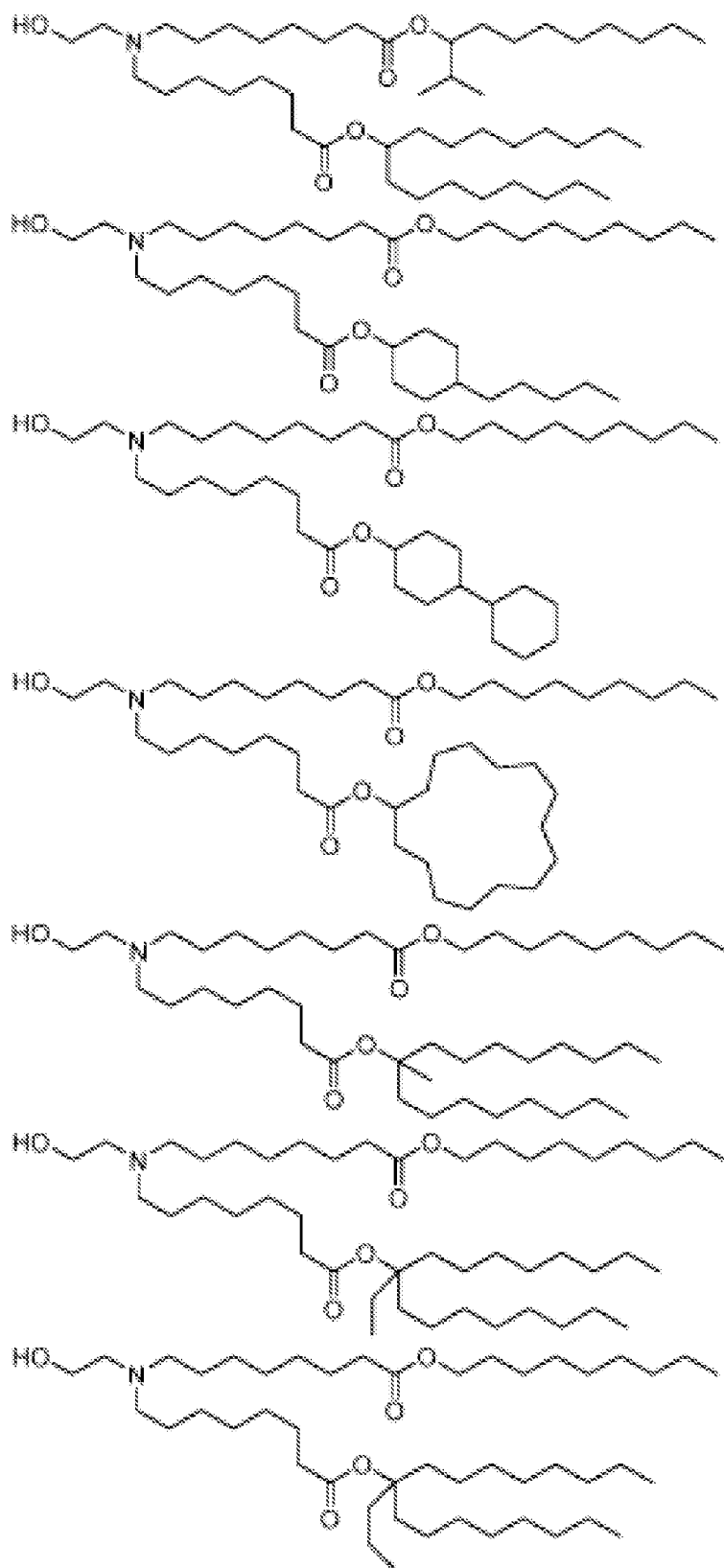


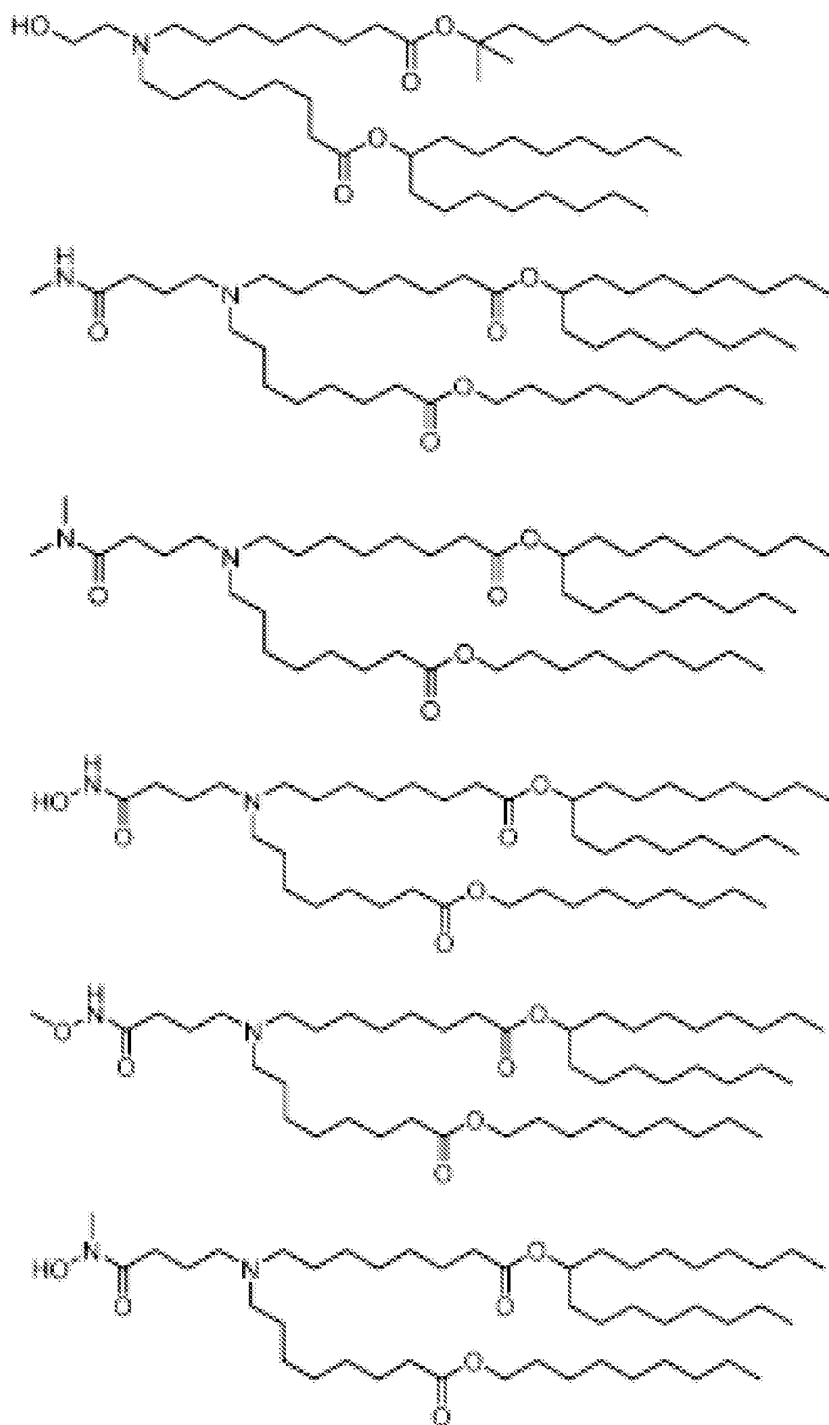


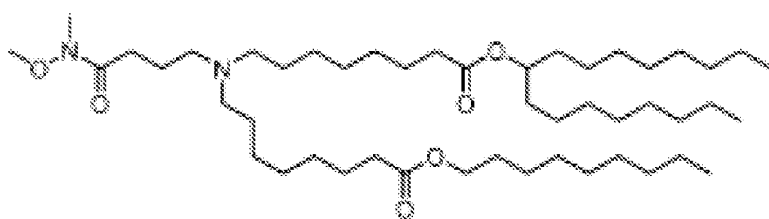
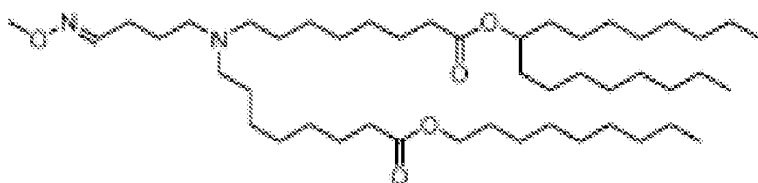
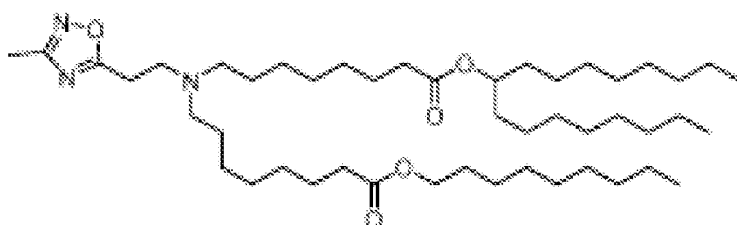
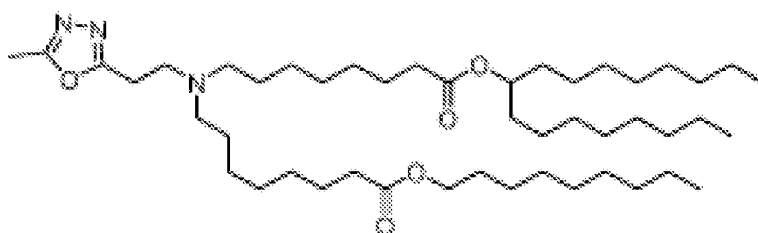
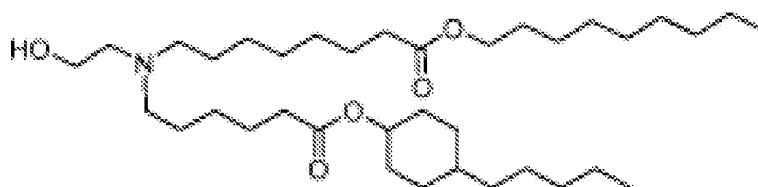
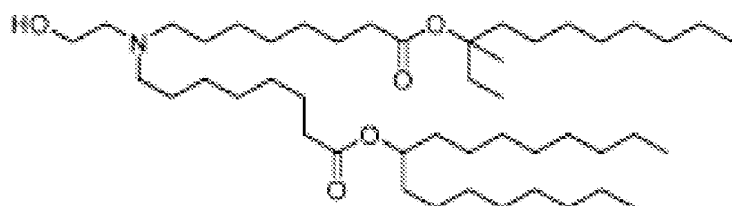
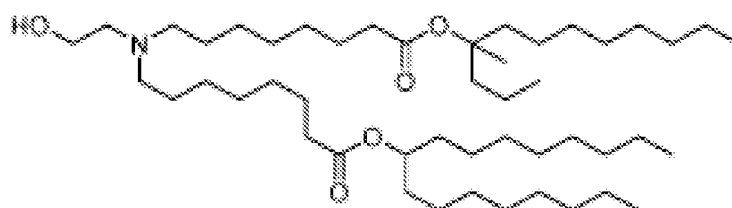


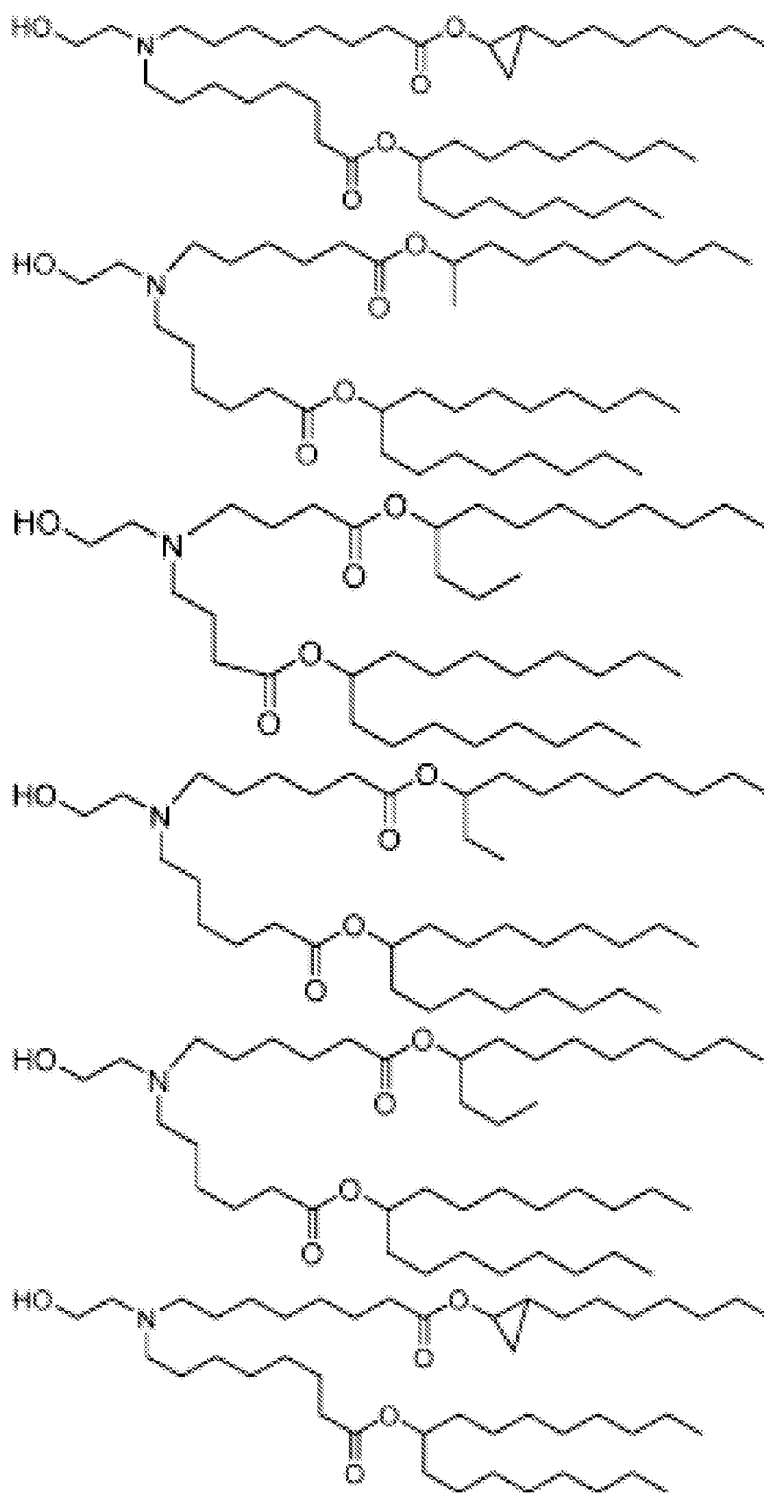


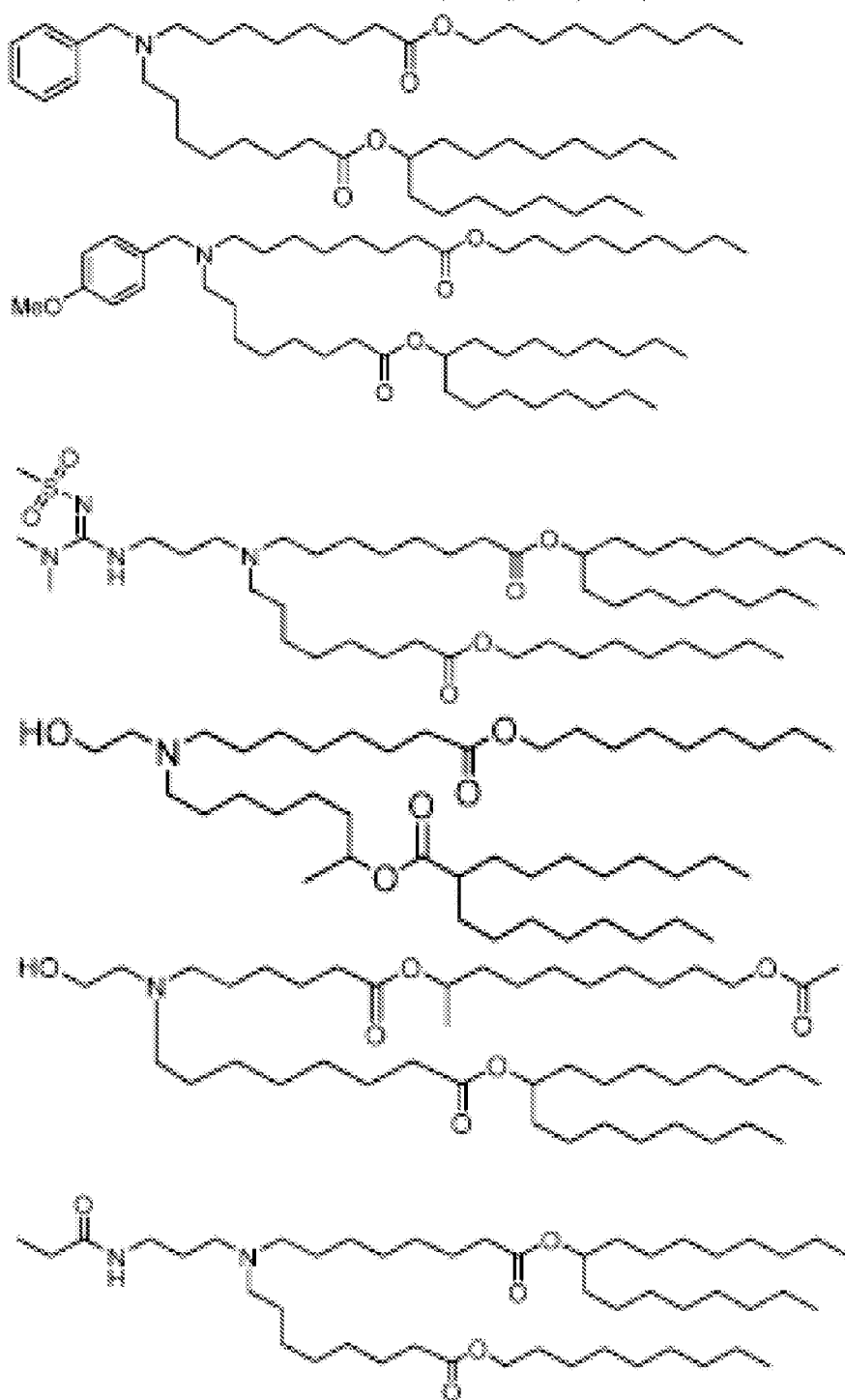


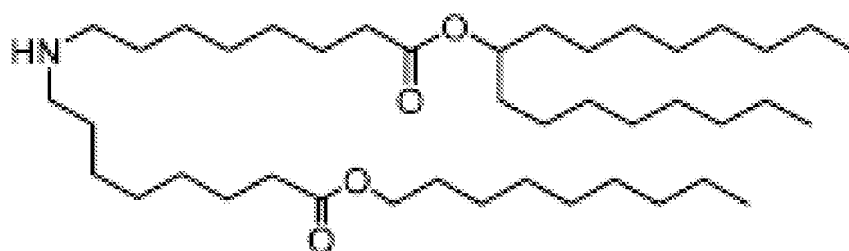
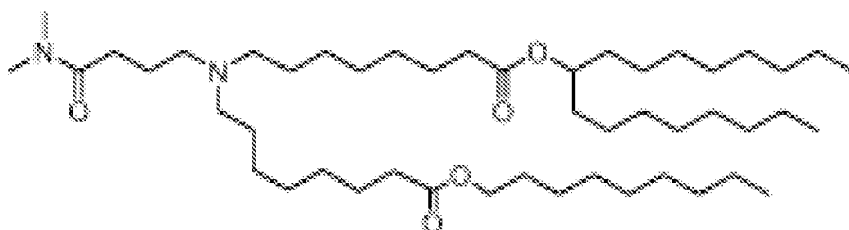
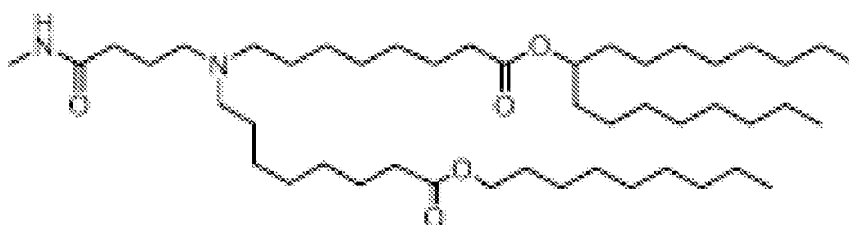
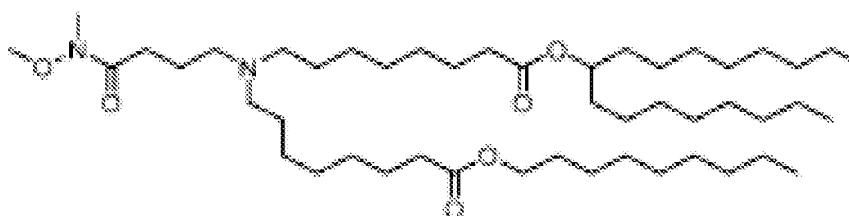
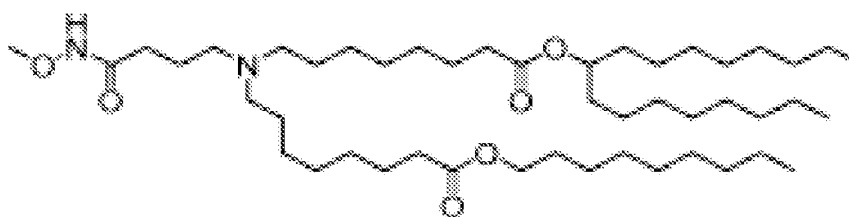
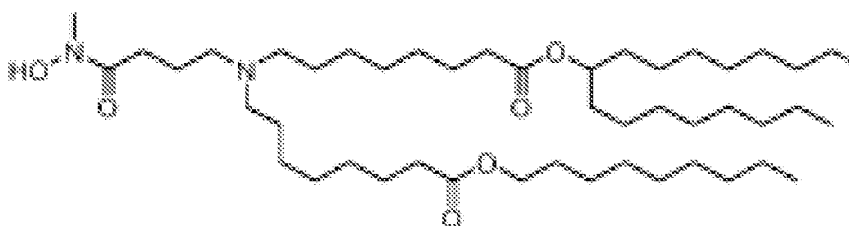


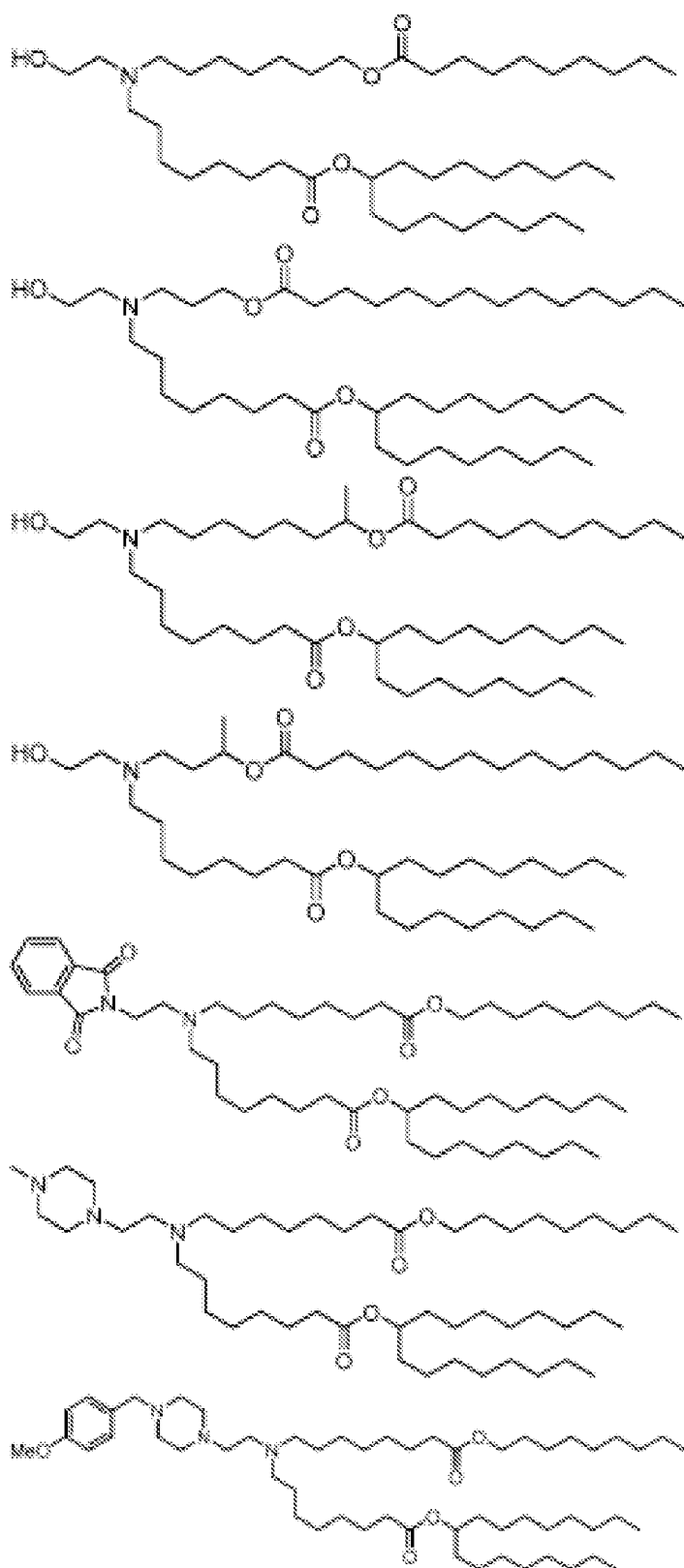


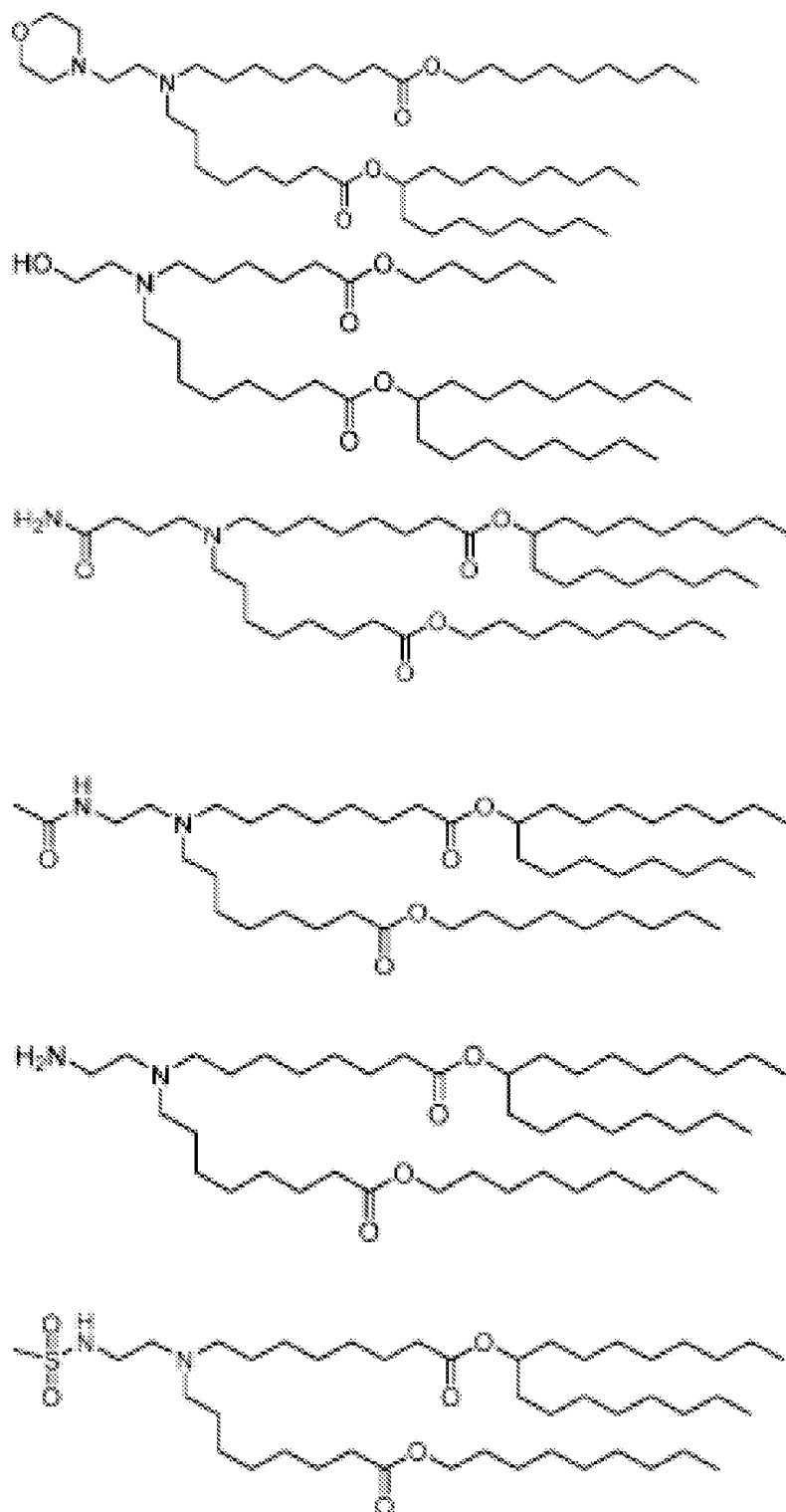


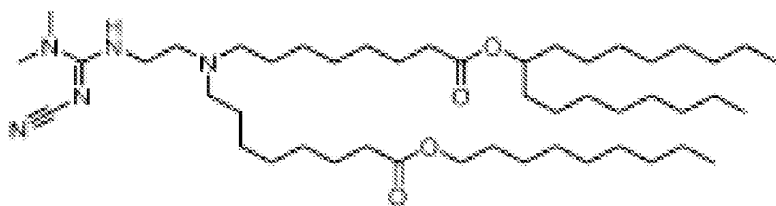
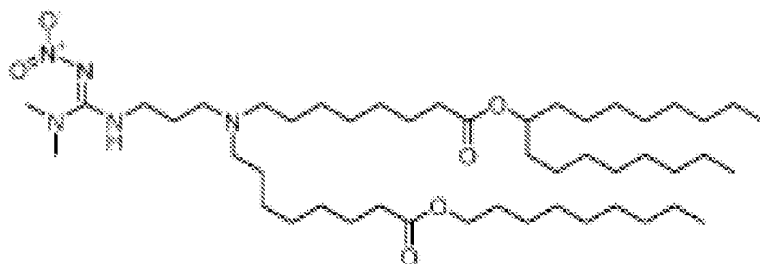
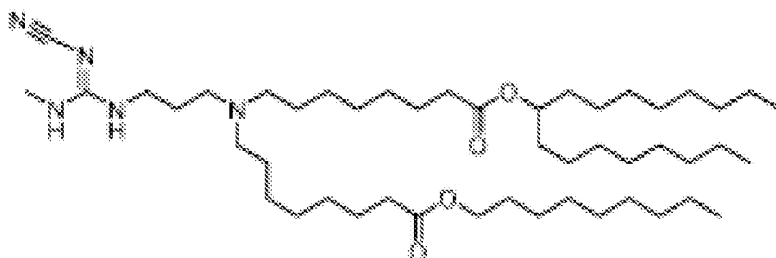
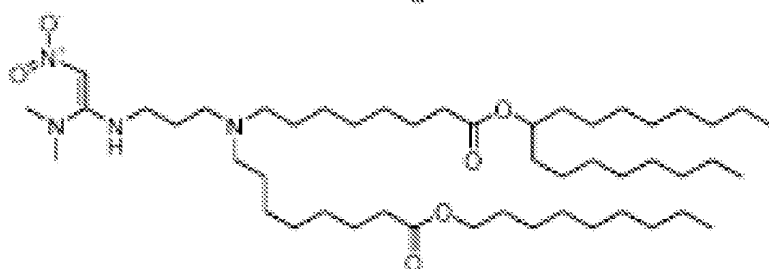
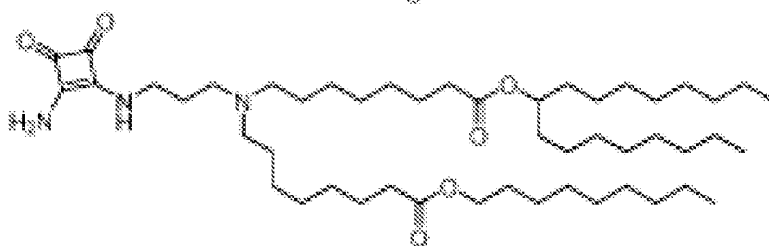


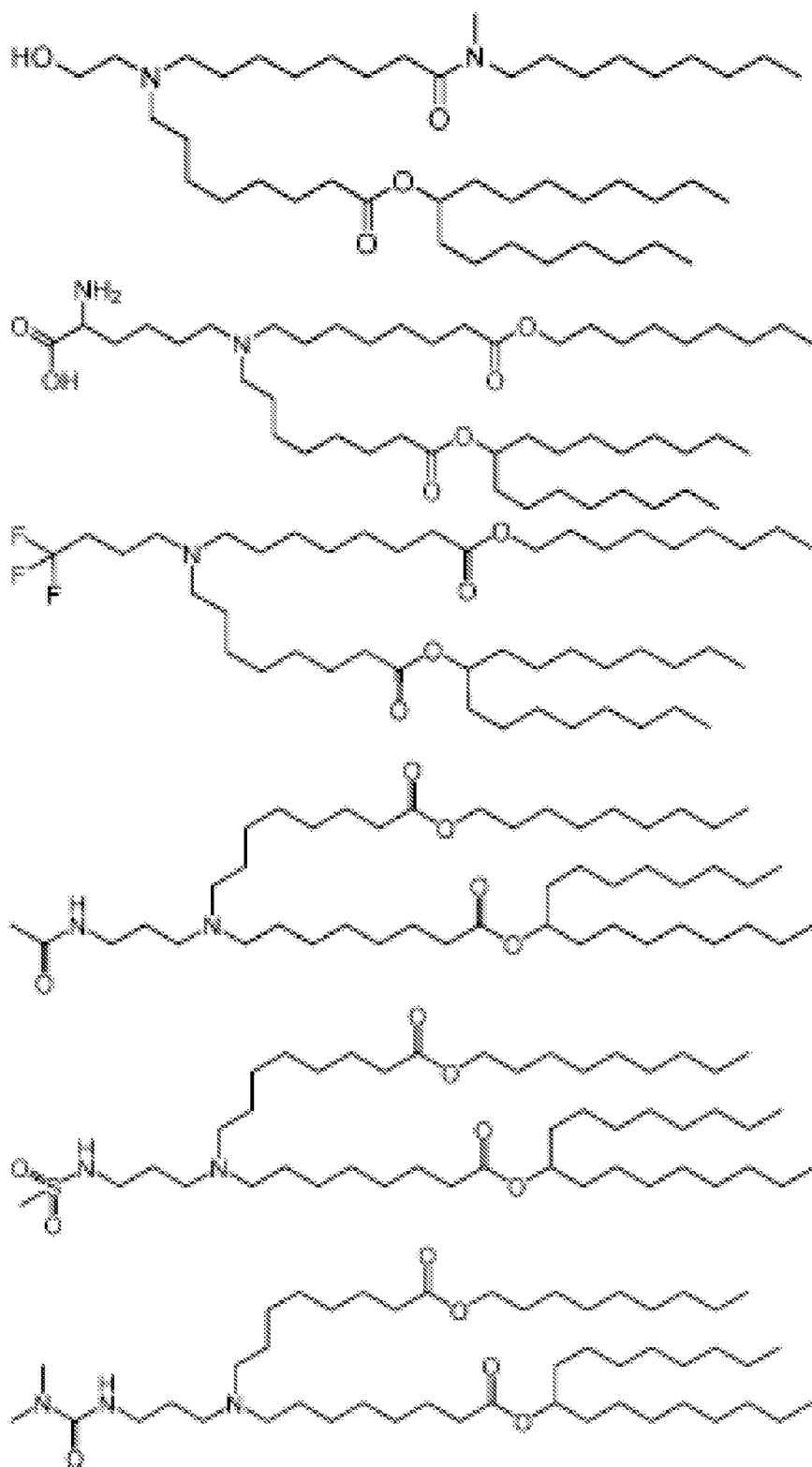


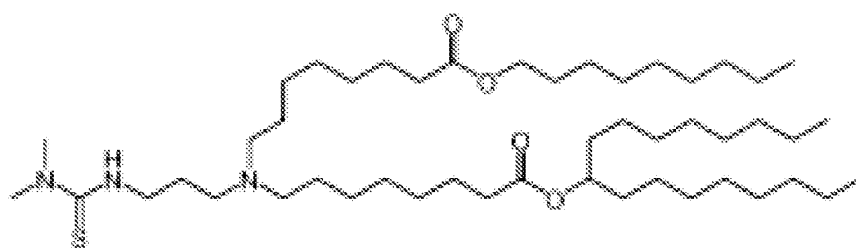
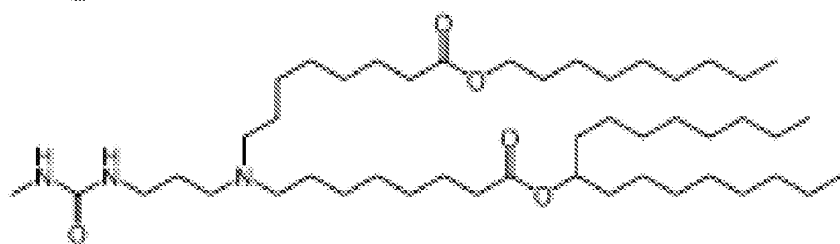
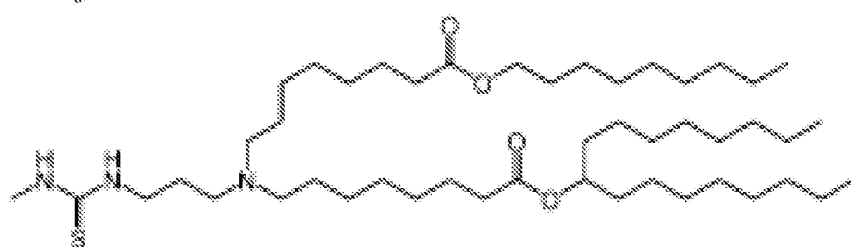
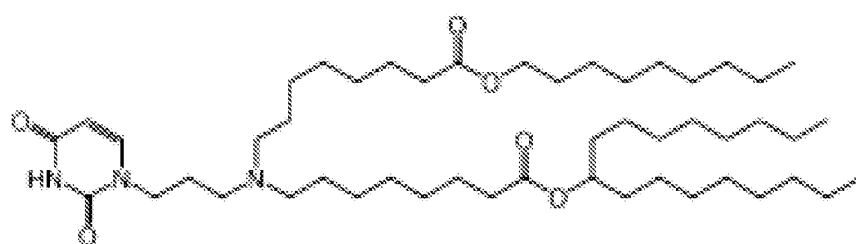
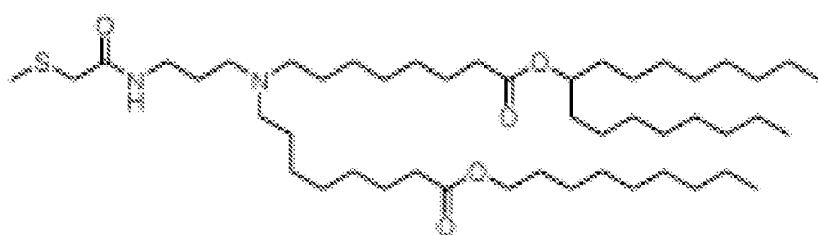
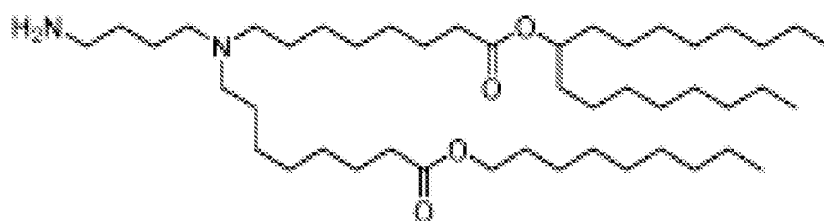


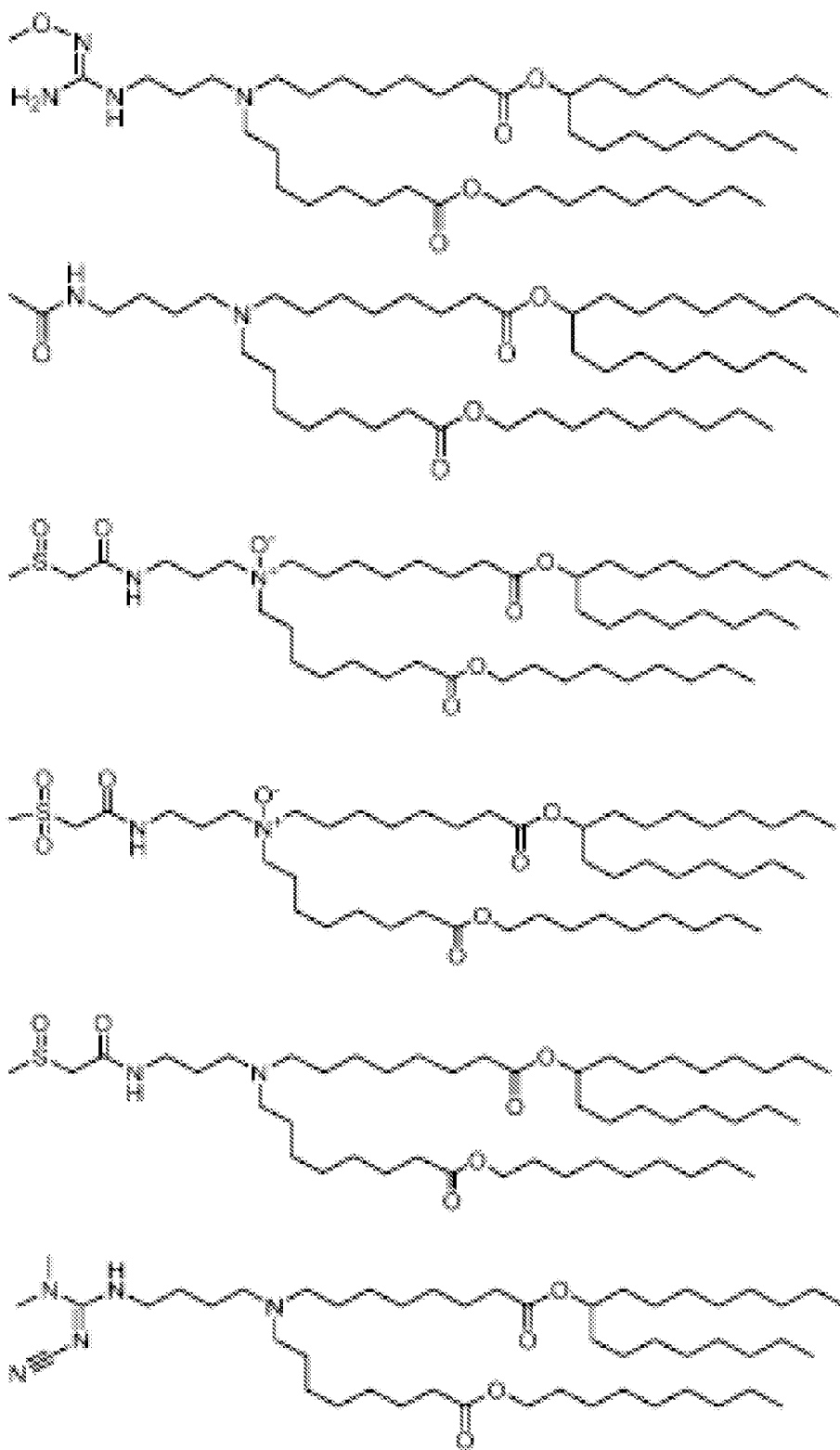


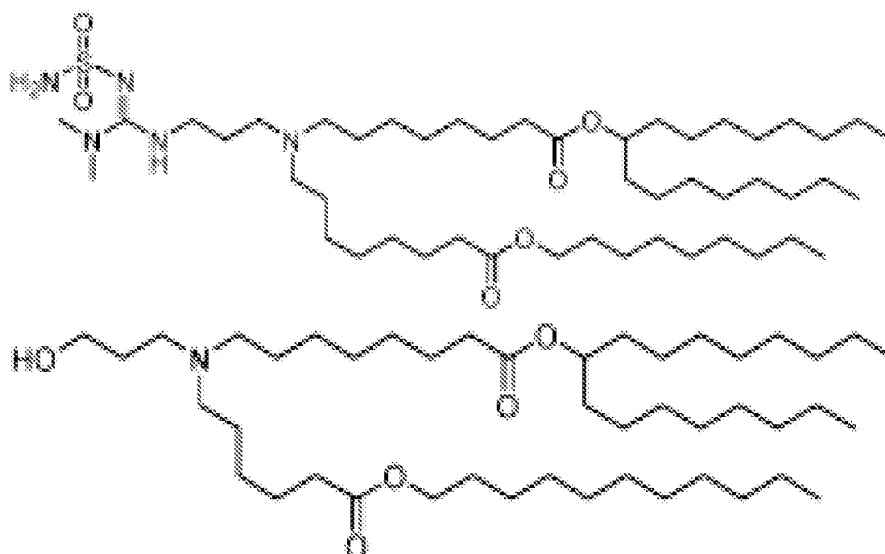










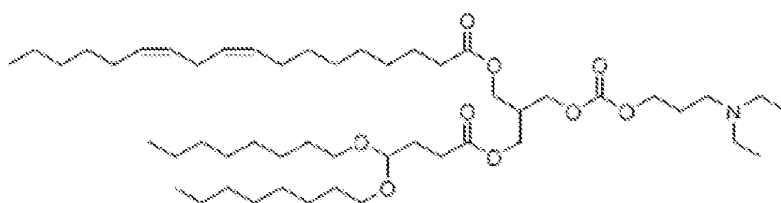


Amine lipids

[0494] In certain embodiments, transfer vehicle compositions for the delivery of circular RNA comprise an amine lipid. In certain embodiments, an ionizable lipid is an amine lipid. In some embodiments, an amine lipid is described in international patent application PCT/US2018/053569.

[0495] In some embodiments, the amine lipid is Lipid E, which is (9Z, 12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-((((3-(diethylamino)propoxy)carbonyl)oxy)methyl)propyl octadeca-9, 12-dienoate.

[0496] Lipid E can be depicted as:



[0497] Lipid E may be synthesized according to WO2015/095340 (*e.g.*, pp. 84-86). In certain embodiments, the amine lipid is an equivalent to Lipid E.

[0498] In certain embodiments, an amine lipid is an analog of Lipid E. In certain embodiments, a Lipid E analog is an acetal analog of Lipid E. In particular transfer vehicle compositions, the acetal analog is a C4-C12 acetal analog. In some embodiments, the acetal analog is a C5-C12 acetal analog. In additional embodiments, the acetal analog is a C5-C10

acetal analog. In further embodiments, the acetal analog is chosen from a C4, C5, C6, C7, C9, C10, C11 and C12 acetal analog.

[0499] Amine lipids and other biodegradable lipids suitable for use in the transfer vehicles, *e.g.*, lipid nanoparticles, described herein are biodegradable *in vivo*. The amine lipids described herein have low toxicity (*e.g.*, are tolerated in animal models without adverse effect in amounts of greater than or equal to 10 mg/kg). In certain embodiments, transfer vehicles composing an amine lipid include those where at least 75% of the amine lipid is cleared from the plasma within 8, 10, 12, 24, or 48 hours, or 3, 4, 5, 6, 7, or 10 days.

[0500] Biodegradable lipids include, for example, the biodegradable lipids of WO2017/173054, WO2015/095340, and WO2014/136086.

[0501] Lipid clearance may be measured by methods known by persons of skill in the art. *See*, for example, Maier, M.A., *et al.* Biodegradable Lipids Enabling Rapidly Eliminated Lipid Nanoparticles for Systemic Delivery of RNAi Therapeutics. *Mol. Ther.* 2013, 21(8), 1570-78.

[0502] Transfer vehicle compositions comprising an amine lipid can lead to an increased clearance rate. In some embodiments, the clearance rate is a lipid clearance rate, for example the rate at which a lipid is cleared from the blood, serum, or plasma. In some embodiments, the clearance rate is an RNA clearance rate, for example the rate at which an circRNA is cleared from the blood, serum, or plasma. In some embodiments, the clearance rate is the rate at which transfer vehicles are cleared from the blood, serum, or plasma. In some embodiments, the clearance rate is the rate at which transfer vehicles are cleared from a tissue, such as liver tissue or spleen tissue. In certain embodiments, a high rate of clearance leads to a safety profile with no substantial adverse effects. The amine lipids and biodegradable lipids may reduce transfer vehicle accumulation in circulation and in tissues. In some embodiments, a reduction in transfer vehicle accumulation in circulation and in tissues leads to a safety profile with no substantial adverse effects.

[0503] Lipids may be ionizable depending upon the pH of the medium they are in. For example, in a slightly acidic medium, the lipid, such as an amine lipid, may be protonated and thus bear a positive charge. Conversely, in a slightly basic medium, such as, for example, blood, where pH is approximately 7.35, the lipid, such as an amine lipid, may not be protonated and thus bear no charge.

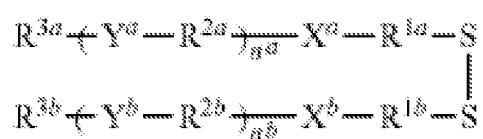
[0504] The ability of a lipid to bear a charge is related to its intrinsic pKa. In some embodiments, the amine lipids of the present disclosure may each, independently, have a pKa in the range of from about 5.1 to about 7.4. In some embodiments, the bioavailable lipids of

the present disclosure may each, independently, have a pKa in the range of from about 5.1 to about 7.4. For example, the amine lipids of the present disclosure may each, independently, have a pKa in the range of from about 5.8 to about 6.5. Lipids with a pKa ranging from about 5.1 to about 7.4 are effective for delivery of cargo *in vivo*, e.g., to the liver. Further, it has been found that lipids with a pKa ranging from about 5.3 to about 6.4 are effective for delivery *in vivo*, e.g., into tumors. See, e.g., WO2014/136086.

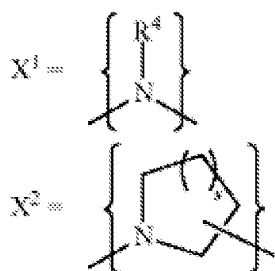
Lipids containing a disulfide bond

[0505] In some embodiments, the ionizable lipid is described in US patent 9,708,628.

[0506] The present invention provides a lipid represented by structure (XXII):



[0507] In structure (XXII), X^a and X^b are each independently X¹ or X² shown below.



[0508] R⁴ in X¹ is an alkyl group having 1-6 carbon atoms, which may be linear, branched or cyclic. The alkyl group preferably has a carbon number of 1-3. Specific examples of the alkyl group having 1-6 carbon atoms include methyl group, ethyl group, propyl group, isopropyl group, n-butyl group, sec-butyl group, isobutyl group, tert-butyl group, pentyl group, isopentyl group, neopentyl group, t-pentyl group, 1,2-dimethylpropyl group, 2-methylbutyl group, 2-methylpentyl group, 3-methylpentyl group, 2,2-dimethylbutyl group, 2,3-dimethylbutyl group, cyclohexyl group and the like. R⁴ is preferably a methyl group, an ethyl group, a propyl group or an isopropyl group, most preferably a methyl group.

[0509] The s in X² is 1 or 2. When s is 1, X² is a pyrrolidinium group, and when s is 2, X² is a piperidinium group. s is preferably 2. While the binding direction of X² is not limited, a nitrogen atom in X² preferably binds to R^{1a} and R^{1b}.

[0510] X^a may be the same as or different from X^b, and X^a is preferably the same group as X^b.

[0511] n^a and n^b are each independently 0 or 1, preferably 1. When n^a is 1, R^{3a} binds to X^a via Y^a and R^{2a} , and when n^a is 0, a structure of $R^{3a}-X^a-R^{1a}-S-$ is taken. Similarly, when n^b is 1, R^{3b} binds to X^b via Y^b and R^{2b} , and when n^b is 0, a structure of $R^{3b}-X^b-R^{1b}-S-$ is taken.

[0512] n^a may be the same as or different from n^b , and n^a is preferably the same as n^b .

[0513] R^{1a} and R^{1b} are each independently an alkylene group having 1-6 carbon atoms, which may be linear or branched, preferably linear. Specific examples of the alkylene group having 1-6 carbon atoms include methylene group, ethylene group, trimethylene group, isopropylene group, tetramethylene group, isobutylene group, pentamethylene group, neopentylene group and the like. R^{1a} and R^{1b} are each preferably a methylene group, an ethylene group, a trimethylene group, an isopropylene group or a tetramethylene group, most preferably an ethylene group.

[0514] R^{1a} may be the same as or different from R^{1b} , and R^{1a} is preferably the same group as R^{1b} .

[0515] R^{2a} and R^{2b} are each independently an alkylene group having 1-6 carbon atoms, which may be linear or branched, preferably linear. Examples of the alkylene group having 1-6 carbon atoms include those recited as the examples of the alkylene group having 1-6 carbon atoms for R^{1a} or R^{1b} . R^{2a} and R^{2b} are each preferably a methylene group, an ethylene group, a trimethylene group, an isopropylene group or a tetramethylene group.

[0516] When X^a and X^b are each X^1 , R^{2a} and R^{2b} are preferably trimethylene groups. When X^a and X^b are each X^2 , R^{2a} and R^{2b} are preferably ethylene groups.

[0517] R^{2a} may be the same as or different from R^{2b} , and R^{2a} is preferably the same group as R^{2b} .

[0518] Y^a and Y^b are each independently an ester bond, an amide bond, a carbamate bond, an ether bond or a urea bond, preferably an ester bond, an amide bond or a carbamate bond, most preferably an ester bond. While the binding direction of Y^a and Y^b is not limited, when Y^a is an ester bond, a structure of $R^{3a}-CO-O-R^{2a}-$ is preferable, and when Y^b is an ester bond, a structure of $R^{3b}-CO-O-R^{2b}-$ is preferable.

[0519] Y^a may be the same as or different from Y^b , and Y^a is preferably the same group as Y^b .

[0520] R^{3a} and R^{3b} are each independently a sterol residue, a liposoluble vitamin residue or an aliphatic hydrocarbon group having 12-22 carbon atoms, preferably a liposoluble vitamin residue or an aliphatic hydrocarbon group having 12-22 carbon atoms, most preferably a liposoluble vitamin residue.

[0521] Examples of the sterol residue include a cholesteryl group (cholesterol residue), a cholestaryl group (cholestanol residue), a stigmasteryl group (stigmasterol residue), a β -sitosteryl group (β -sitosterol residue), a lanosteryl group (lanosterol residue), and an ergosteryl group (ergosterol residue) and the like. The sterol residue is preferably a cholesteryl group or a cholestaryl group.

[0522] As the liposoluble vitamin residue, a residue derived from liposoluble vitamin, as well as a residue derived from a derivative obtained by appropriately converting a hydroxyl group, aldehyde or carboxylic acid, which is a functional group in liposoluble vitamin, to other reactive functional group can be used. As for liposoluble vitamin having a hydroxyl group, for example, the hydroxyl group can be converted to a carboxylic acid by reacting with succinic acid anhydride, glutaric acid anhydride and the like. Examples of the liposoluble vitamin include retinoic acid, retinol, retinal, ergosterol, 7-dehydrocholesterol, calciferol, cholecalciferol, dihydroergocalciferol, dihydrotachysterol, tocopherol, tocotrienol and the like. Preferable examples of the liposoluble vitamin include retinoic acid and tocopherol.

[0523] The aliphatic hydrocarbon group having 12-22 carbon atoms may be linear or branched, preferably linear. The aliphatic hydrocarbon group may be saturated or unsaturated. In the case of an unsaturated aliphatic hydrocarbon group, the aliphatic hydrocarbon group generally contains 1-6, preferably 1-3, more preferably 1-2 unsaturated bonds. While the unsaturated bond includes a carbon-carbon double bond and a carbon-carbon triple bond, it is preferably a carbon-carbon double bond. The aliphatic hydrocarbon group has a carbon number of preferably 12-18, most preferably 13-17. While the aliphatic hydrocarbon group includes an alkyl group, an alkenyl group, an alkynyl group and the like, it is preferably an alkyl group or an alkenyl group. Specific examples of the aliphatic hydrocarbon group having 12-22 carbon atoms include dodecyl group, tridecyl group, tetradecyl group, pentadecyl group, hexadecyl group, heptadecyl group, octadecyl group, nonadecyl group, icosyl group, hencosyl group, docosyl group, dodecenyl group, tridecenyl group, tetradecenyl group, pentadecenyl group, hexadecenyl group, heptadecenyl group, octadecenyl group, nonadecenyl group, icosenyl group, hencosenyl group, docosenyl group, decadienyl group, tridecadienyl group, tetradecadienyl group, pentadecadienyl group, hexadecadienyl group, heptadecadienyl group, octadecadienyl group, nonadecadienyl group, icosadienyl group, hencosadienyl group, docosadienyl group, octadecatrienyl group, icosatrienyl group, icosatetraenyl group, icosapentaenyl group, docosaheptaenyl group, isostearyl group and the like. The aliphatic hydrocarbon group having 12-22 carbon atoms is

preferably tridecyl group, tetradecyl group, heptadecyl group, octadecyl group, heptadecadienyl group or octadecadienyl group, particularly preferably tridecyl group, heptadecyl group or heptadecadienyl group.

[0524] In one embodiment, an aliphatic hydrocarbon group having 12-22 carbon atoms, which is derived from fatty acid, aliphatic alcohol, or aliphatic amine is used. When R^{3a} (or R^{3b}) is derived from fatty acid, Y^a (or Y^b) is an ester bond or an amide bond, and fatty acid-derived carbonyl carbon is included in Y^a (or Y^b). For example, when linoleic acid is used, R^{3a} (or R^{3b}) is a heptadecadienyl group.

[0525] R^{3a} may be the same as or different from R^{3b} , and R^{3a} is preferably the same group as R^{3b} .

[0526] In one embodiment, X^a is the same as X^b , n^a is the same as n^b , R^{1a} is the same as R^{1b} , R^{2a} is the same as R^{2b} , R^{3a} is the same as R^{3b} , and Y^a is the same as Y^b .

[0527] In one embodiment,

X^a and X^b are each independently X^1 ,

R^4 is an alkyl group having 1-3 carbon atoms, n^a and n^b are each 1,

R^{1a} and R^{1b} are each independently an alkylene group having 1-6 carbon atoms,

R^{2a} and R^{2b} are each independently an alkylene group having 1-6 carbon atoms,

Y^a and Y^b are each an ester bond or an amide bond, and

R^{3a} and R^{3b} are each independently an aliphatic hydrocarbon group having 12-22 carbon atoms.

[0528] In one embodiment,

X^a and X^b are each X^1 ,

R^4 is an alkyl group having 1-3 carbon atoms, n^a and n^b are each 1,

R^{1a} and R^{1b} are each an alkylene group having 1-6 carbon atoms,

R^{2a} and R^{2b} are each an alkylene group having 1-6 carbon atoms,

Y^a and Y^b are each an ester bond or an amide bond,

R^{3a} and R^{3b} are each an aliphatic hydrocarbon group having 12-22 carbon atoms,

X^a is the same as X^b ,

R^{1a} is the same as R^{1b} ,

R^{2a} is the same as R^{2b} , and

R^{3a} is the same as R^{3b} .

[0529] In one embodiment,

X^a and X^b are each X^1 ,

R^4 is a methyl group, n^a and n^b are each 1,

R^{1a} and R^{1b} are each an ethylene group,
 R^{2a} and R^{2b} are each a trimethylene group,
 Y^a and Y^b are each —CO—O— , and
 R^{3a} and R^{3b} are each independently an alkyl group or alkenyl group having 13-17 carbon atoms.

[0530] In one embodiment,

X^a and X^b are each X^1 ,
 R^4 is a methyl group, n^a and n^b are each 1,
 R^{1a} and R^{1b} are each an ethylene group,
 R^{2a} and R^{2b} are each a trimethylene group,
 Y^a and Y^b are each —CO—O— ,
 R^{3a} and R^{3b} are each an alkyl group or alkenyl group having 13-17 carbon atoms, and
 R^{3a} is the same as R^{3b} .

[0531] In one embodiment,

X^a and X^b are each independently X^1 ,
 R^4 is an alkyl group having 1-3 carbon atoms, n^a and n^b are each 1,
 R^{1a} and R^{1b} are each independently an alkylene group having 1-6 carbon atoms,
 R^{2a} and R^{2b} are each independently an alkylene group having 1-6 carbon atoms,
 Y^a and Y^b are each an ester bond or an amide bond, and
 R^{3a} and R^{3b} are each independently a liposoluble vitamin residue (e.g., retinoic acid residue, tocopherol residue).

[0532] In one embodiment,

X^a and X^b are each X^1 ,
 R^4 is an alkyl group having 1-3 carbon atoms, n^a and n^b are each 1,
 R^{1a} and R^{1b} are each an alkylene group having 1-6 carbon atoms,
 R^{2a} and R^{2b} are each an alkylene group having 1-6 carbon atoms,
 Y^a and Y^b are each an ester bond or an amide bond,
 R^{3a} and R^{3b} are each a liposoluble vitamin residue (e.g., retinoic acid residue, tocopherol residue),

X^a is the same as X^b ,
 R^{1a} is the same as R^{1b} ,
 R^{2a} is the same as R^{2b} , and
 R^{3a} is the same as R^{3b} .

[0533] In one embodiment,

X^a and X^b are each X^1 ,
 R^4 is a methyl group, n^a and n^b are each 1,
 R^{1a} and R^{1b} are each an ethylene group,
 R^{2a} and R^{2b} are each a trimethylene group,
 Y^a and Y^b are each ---CO---O--- , and
 R^{3a} and R^{3b} are each independently a liposoluble vitamin residue (e.g., retinoic acid residue, tocopherol residue).

[0534] In one embodiment,

X^a and X^b are each X^1 ,
 R^4 is a methyl group, n^a and n^b are each 1,
 R^{1a} and R^{1b} are each an ethylene group,
 R^{2a} and R^{2b} are each a trimethylene group,
 Y^a and Y^b are each ---CO---O--- ,
 R^{3a} and R^{3b} are each a liposoluble vitamin residue (e.g., retinoic acid residue, tocopherol residue), and
 R^{3a} is the same as R^{3b} .

[0535] In one embodiment,

X^a and X^b are each independently X^2 ,
 t is 2,
 R^{1a} and R^{1b} are each independently an alkylene group having 1-6 carbon atoms,
 R^{2a} and R^{2b} are each independently an alkylene group having 1-6 carbon atoms,
 Y^a and Y^b are each an ester bond, and
 R^{3a} and R^{3b} are each independently a liposoluble vitamin residue (e.g., retinoic acid residue, tocopherol residue) or an aliphatic hydrocarbon group having 12-22 carbon atoms (e.g., alkyl group having 12-22 carbon atoms).

[0536] In one embodiment,

X^a and X^b are each independently X^2 ,
 t is 2,
 R^{1a} and R^{1b} are each independently an alkylene group having 1-6 carbon atoms,
 R^{2a} and R^{2b} are each independently an alkylene group having 1-6 carbon atoms,
 Y^a and Y^b are each an ester bond,
 R^{3a} and R^{3b} are each independently a liposoluble vitamin residue (e.g., retinoic acid residue, tocopherol residue) or an aliphatic hydrocarbon group having 12-22 carbon atoms (e.g., alkyl group having 12-22 carbon atoms),

X^a is the same as X^b ,

R^{1a} is the same as R^{1b} ,

R^{2a} is the same as R^{2b} , and

R^{3a} is the same as R^{3b} .

[0537] In one embodiment,

X^a and X^b are each independently X^2 ,

t is 2,

R^{1a} and R^{1b} are each an ethylene group,

R^{2a} and R^{2b} are each independently an alkylene group having 1-6 carbon atoms,

Y^a and Y^b are each an ester bond,

R^{3a} and R^{3b} are each independently a liposoluble vitamin residue (e.g., retinoic acid residue, tocopherol residue) or an aliphatic hydrocarbon group having 12-22 carbon atoms (e.g., alkyl group having 12-22 carbon atoms),

X^a is the same as X^b ,

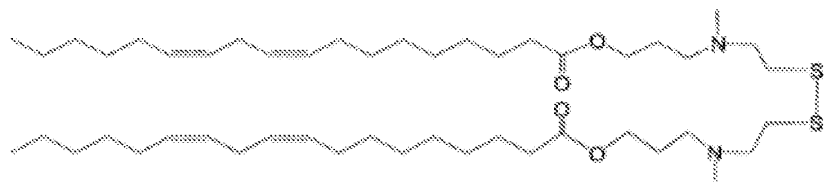
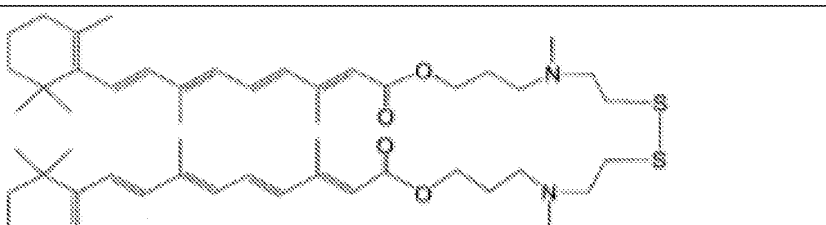
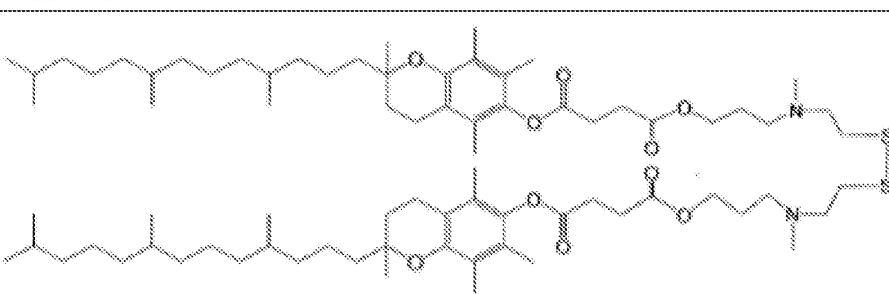
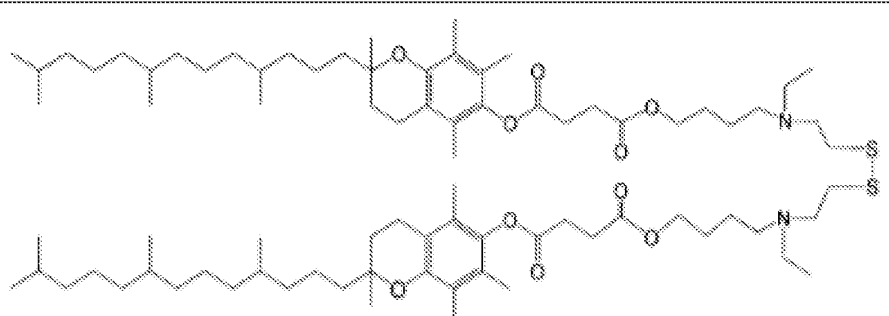
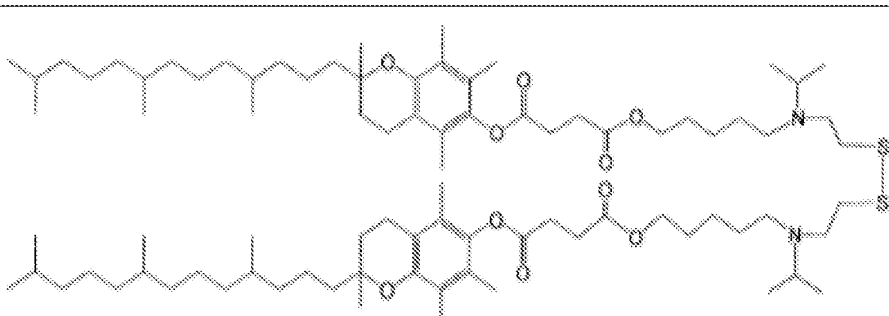
R^{2a} is the same as R^{2b} , and

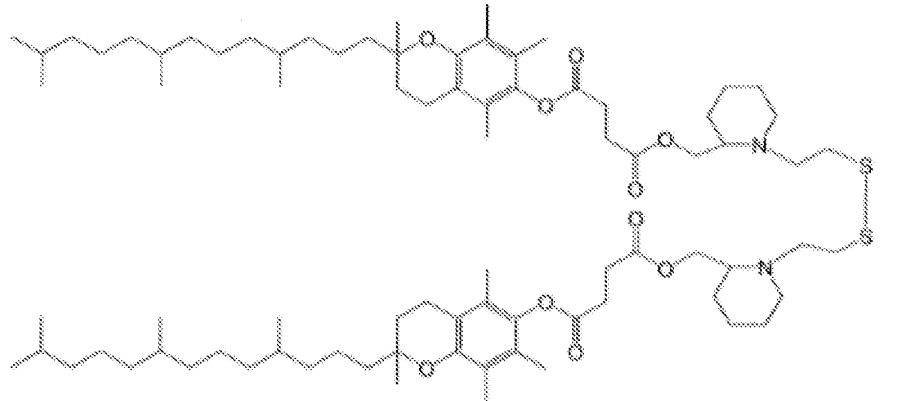
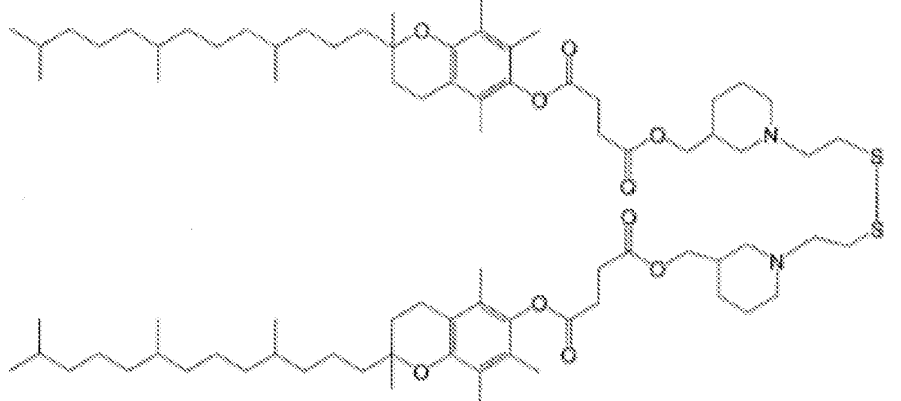
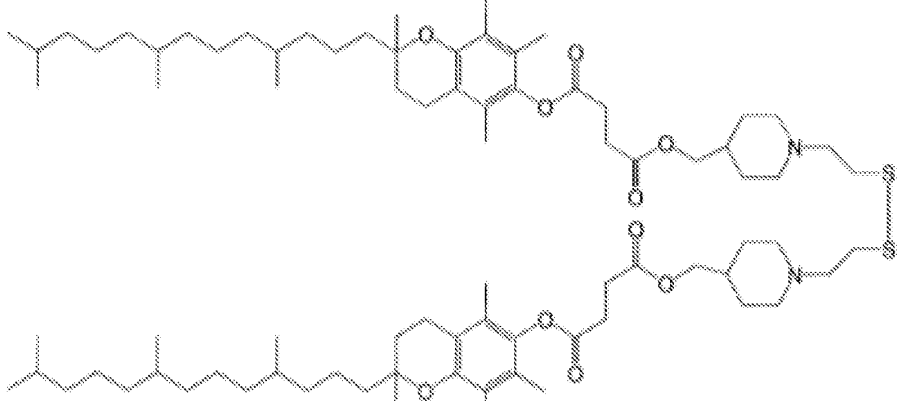
R^{3a} is the same as R^{3b} .

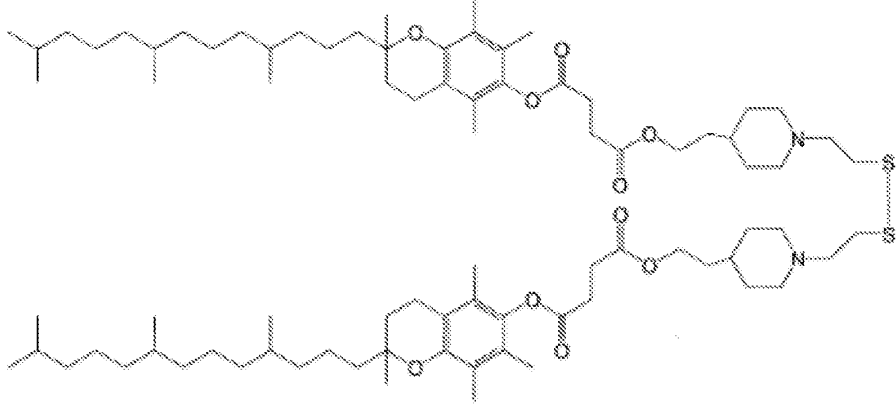
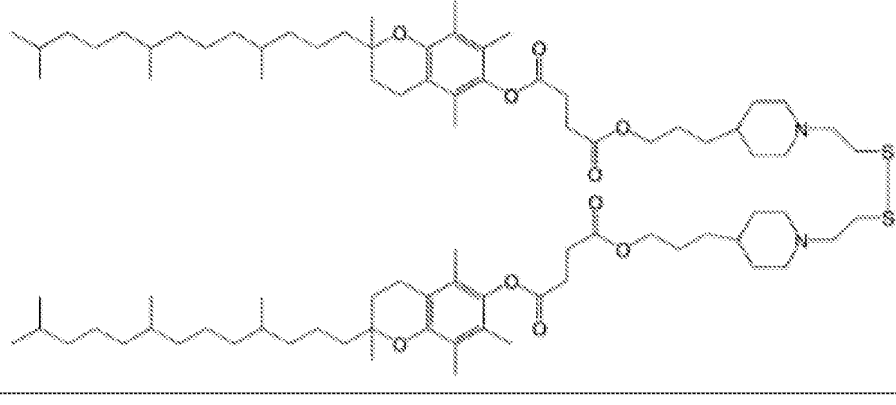
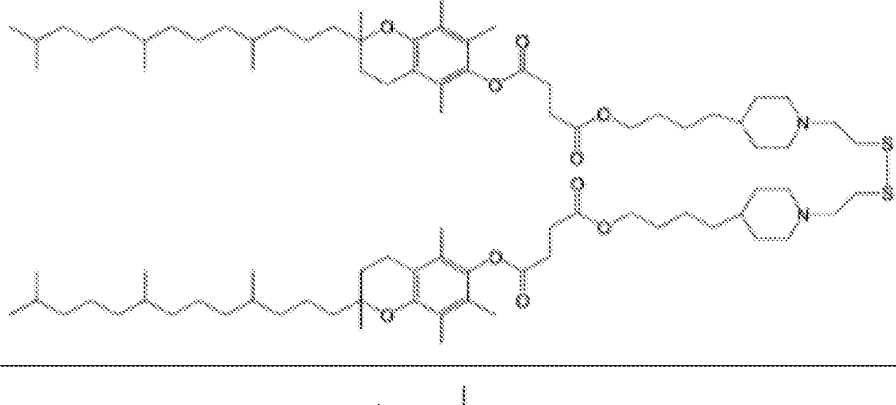
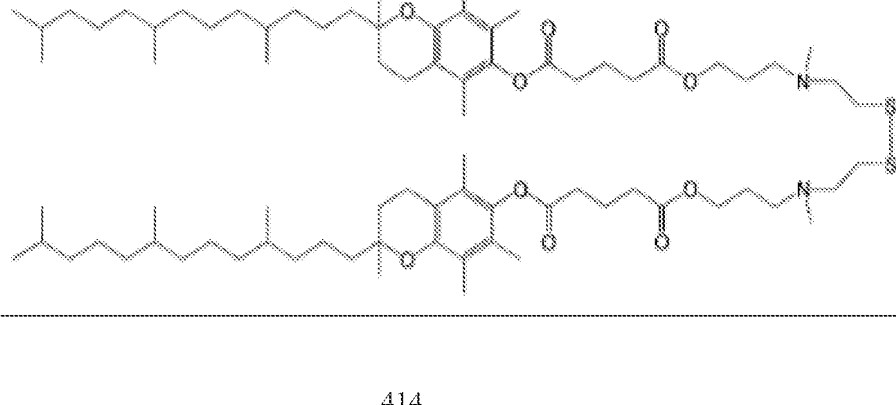
[0538] In some embodiments, an ionizable lipid has one of the structures set forth in Table 15b below.

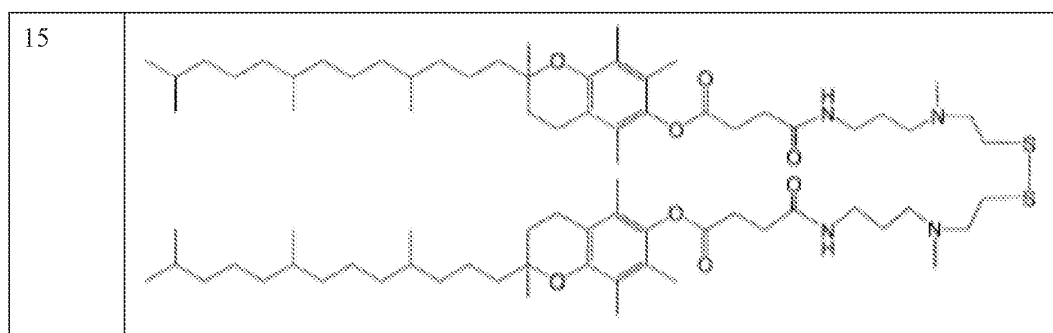
Table 15b

Number	Structure
1	
2	

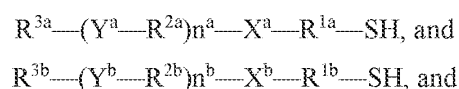
3	
4	
5	
6	
7	

8	
9	
10	

11	
12	
13	
14	



[0539] A lipid of the present invention may have an —S—S— (disulfide) bond. The production method for such a compound includes, for example, a method including producing



subjecting them to oxidation (coupling) to give a compound containing —S—S—, a method including sequentially bonding necessary parts to a compound containing an —S—S— bond to finally obtain the compound of the present invention and the like. Preferred is the latter method.

[0540] A specific example of the latter method is shown below, which is not to be construed as limiting.

[0541] Examples of the starting compound include —S—S— bond-containing two terminal carboxylic acid, two terminal carboxylate, two terminal amine, two terminal isocyanate, two terminal alcohol, two terminal alcohol having a leaving group such as MsO (mesylate group) and the like, a two terminal carbonate having a leaving group such as pNP (p-nitrophenylcarbonate group) and the like.

[0542] For example, when a compound containing X^1 or X^2 for X^a and X^b is produced, two terminal functional groups of compound (1) containing an —S—S— bond are reacted with an —NH— group in compound (2) having the —NH— group and one functional group at the terminal, the functional group at the terminal in compound (2) which did not contribute to the reaction is reacted with a functional group in compound (3) containing R^3 , whereby the compound of the present invention containing an —S—S— bond, R^{1a} and R^{1b} , X^a and X^b , R^{2a} and R^{2b} , Y^a and Y^b , and R^{3a} and R^{3b} can be obtained.

[0543] In the reaction of compound (1) and compound (2), an alkali catalyst such as potassium carbonate, sodium carbonate, potassium t-butoxide and the like may be used as a catalyst, or the reaction may be performed without a catalyst. Preferably, potassium carbonate

or sodium carbonate is used as a catalyst.

[0544] The amount of catalyst is 0.1-100 molar equivalents, preferably, 0.1-20 molar equivalents, more preferably 0.1-5 molar equivalents, relative to compound (1). The amount of compound (2) to be charged is 1-50 molar equivalents, preferably 1-10 molar equivalents, relative to compound (1).

[0545] The solvent to be used for the reaction of compound (1) and compound (2) is not particularly limited as long as it is a solvent or aqueous solution that does not inhibit the reaction. For example, ethyl acetate, dichloromethane, chloroform, benzene, toluene and the like can be mentioned. Among these, toluene and chloroform are preferable.

[0546] The reaction temperature is -20 to 200°C ., preferably 0 to 80°C ., more preferably 20 to 50°C ., and the reaction time is 1-48 hr, preferably 2-24 hr.

[0547] When the reaction product of compound (1) and compound (2) is reacted with compound (3), an alkali catalyst such as potassium carbonate, sodium carbonate, potassium *t*-butoxide and the like, or an acid catalyst such as PTS (*p*-toluenesulfonic acid), MSA (methanesulfonic acid) and the like may be used, like the catalyst used for the reaction of compound (1) and compound (2), or the reaction may be performed without a catalyst.

[0548] In addition, the reaction product of compound (1) and compound (2) may be directly reacted with compound (3) by using a condensing agent such as DCC (dicyclohexylcarbodiimide), DIC (diisopropylcarbodiimide), EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride) and the like. Alternatively, compound (3) may be treated with a condensing agent to be once converted to an anhydride and the like, after which it is reacted with the reaction product of compound (1) and compound (2).

[0549] The amount of compound (3) to be charged is 1-50 molar equivalents, preferably 1-10 molar equivalents, relative to the reaction product of compound (1) and compound (2).

[0550] The catalyst to be used is appropriately selected according to the functional groups to be reacted.

[0551] The amount of catalyst is 0.05-100 molar equivalents, preferably 0.1-20 molar equivalents, more preferably 0.2-5 molar equivalent, relative to compound (1).

[0552] The solvent to be used for the reaction of the reaction product of compound (1) and compound (2) with compound (3) is not particularly limited as long as it is a solvent or aqueous solution that does not inhibit the reaction. For example, ethyl acetate, dichloromethane, chloroform, benzene, toluene and the like can be mentioned. Among these, toluene and chloroform are preferable.

[0553] The reaction temperature is 0 to 200°C ., preferably 0 to 120°C ., more preferably

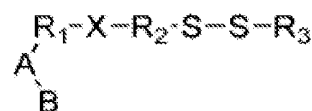
20 to 50° C., and the reaction time is 1 hr-48 hr, preferably 2-24 hr.

[0554] The reaction product obtained by the above-mentioned reaction can be appropriately purified by a general purification method, for example, washing with water, silica gel column chromatography, crystallization, recrystallization, liquid-liquid extraction, reprecipitation, ion exchange column chromatography and the like.

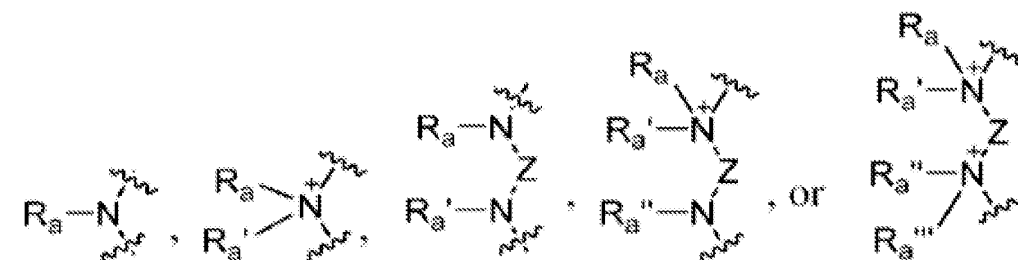
Structure XXIII lipids

[0555] In some embodiments, an ionizable lipid is described in US patent 9,765,022.

[0556] The present invention provides a compound represented by structure (XXIII):



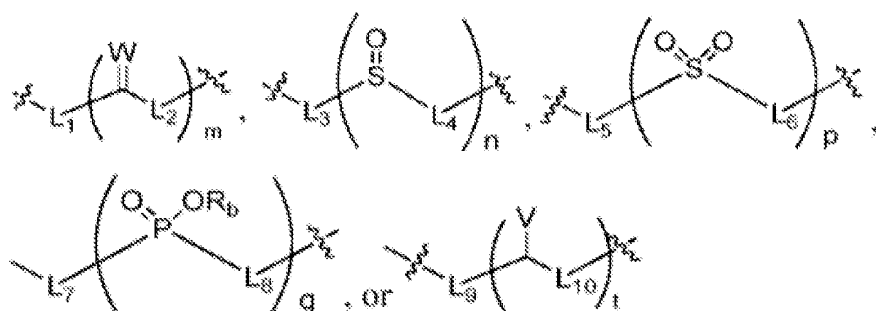
[0557] In structure XXIII, a hydrophilic and optionally positively charged head is



in which each of R_a , R_a' , R_a'' , and R_a''' , independently, is H, a C₁-C₂₀ monovalent aliphatic radical, a C₁-C₂₀ monovalent heteroaliphatic radical, a monovalent aryl radical, or a monovalent heteroaryl radical, and Z is a C₁-C₂₀ bivalent aliphatic radical, a C₁-C₂₀ bivalent heteroaliphatic radical, a bivalent aryl radical, or a bivalent heteroaryl radical; B is a C₁-C₂₄ monovalent aliphatic radical, a C₁-C₂₄ monovalent heteroaliphatic radical, a monovalent aryl

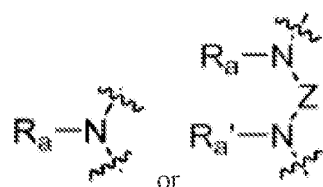
radical, a monovalent heteroaryl radical, or $\begin{array}{c} \text{Z} \\ | \\ -R_4-Y-R_5-S-S-R_6 \end{array}$ each of R_1 and R_4 , independently, is a bond, a C₁-C₁₀ bivalent aliphatic radical, a C₁-C₁₀ bivalent heteroaliphatic radical, a bivalent aryl radical, or a bivalent heteroaryl radical; each of R_2 and R_5 , independently, is a bond, a C₁-C₂₀ bivalent aliphatic radical, a C₁-C₂₀ bivalent heteroaliphatic radical, a bivalent aryl radical, or a bivalent heteroaryl radical; each of R_3 and R_6 , independently, is a C₁-C₂₀ monovalent aliphatic radical, a C₁-C₂₀ monovalent heteroaliphatic radical, a monovalent aryl radical, or a monovalent heteroaryl radical; each of

$\begin{array}{c} \text{Z} \\ | \\ -R_2-S-S-R_3 \end{array}$ a hydrophobic tail, and $\begin{array}{c} \text{Z} \\ | \\ -R_5-S-S-R_6 \end{array}$ also a hydrophobic tail, has 8 to 24 carbon atoms; and each of X, a linker, and Y, also a linker, independently, is



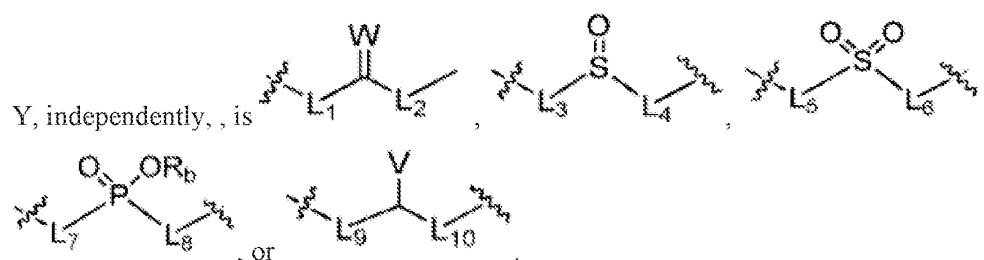
in which each of m , n , p , q , and t , independently, is 1-6; W is O, S, or NR_c ; each of L_1 , L_3 , L_5 , L_7 , and L_9 , directly linked to R_1 , R_2 , R_4 , or R_5 , independently, is a bond, O, S, or NR_d ; each of L_2 , L_4 , L_6 , L_8 , and L_{10} , independently, is a bond, O, S, or NR_e ; V is OR_f , SR_g , or NR_hR_i ; and each of R_b , R_c , R_d , R_e , R_f , R_g , R_h , and R_i , independently, is H, OH, C_{1-10} oxyaliphatic radical, C_{1-10} monovalent aliphatic radical, C_{1-10} monovalent heteroaliphatic radical, a monovalent aryl radical, or a monovalent heteroaryl radical.

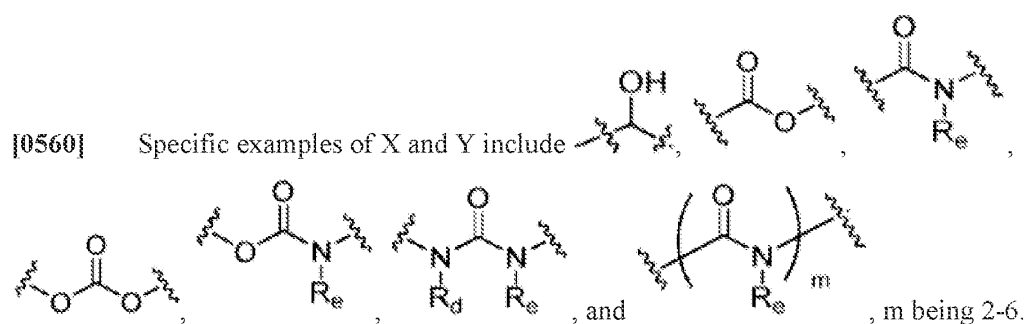
[0558] A subset of the above-described lipid-like compounds include those in which A is



, each of R_a and R_a' , independently, being a C_{1-10} monovalent aliphatic radical, a C_{1-10} monovalent heteroaliphatic radical, a monovalent aryl radical, or a monovalent heteroaryl radical; and Z being a C_{1-10} bivalent aliphatic radical, a C_{1-10} bivalent heteroaliphatic radical, a bivalent aryl radical, or a bivalent heteroaryl radical.

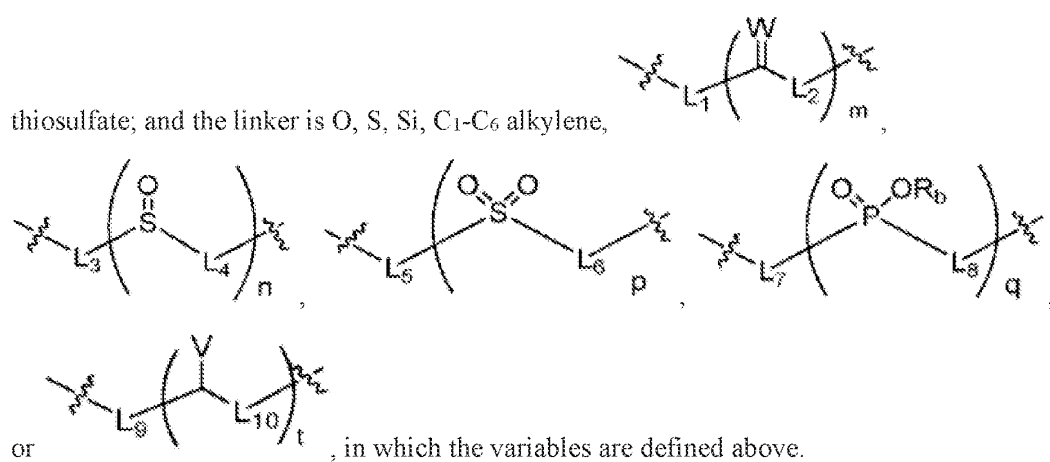
[0559] Some lipid-like compounds of this invention feature each of R_1 and R_4 , independently, being C_{1-6} (e.g., C_{1-4}) bivalent aliphatic radical or a C_{1-6} (e.g., C_{1-4}) bivalent heteroaliphatic radical, the total carbon number for R_2 and R_3 being 12-20 (e.g., 14-18), the total carbon number of R_5 and R_6 also being 12-20 (e.g., 14-18), and each of X and



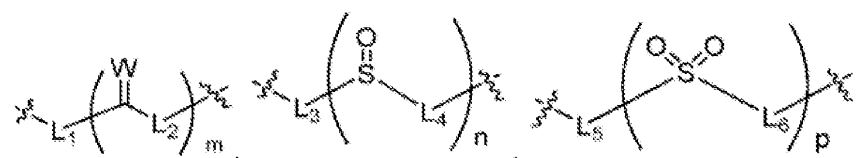


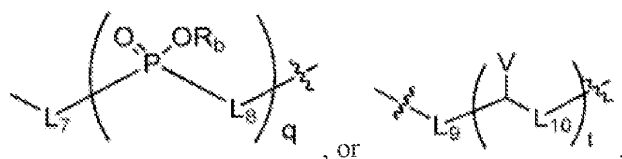
[0561] Still within the scope of this invention is a pharmaceutical composition containing a nanocomplex that is formed of a protein and a bio-reducible compound. In this pharmaceutical composition, the nanocomplex has a particle size of 50 to 500 nm; the bio-reducible compound contains a disulfide hydrophobic moiety, a hydrophilic moiety, and a linker joining the disulfide hydrophobic moiety and the hydrophilic moiety; and the protein binds to the bio-reducible compound via a non-covalent interaction, a covalent bond, or both.

[0562] In certain embodiments, the disulfide hydrophobic moiety is a heteroaliphatic radical containing one or more —S—S— groups and 8 to 24 carbon atoms; the hydrophilic moiety is an aliphatic or heteroaliphatic radical containing one or more hydrophilic groups and 1-20 carbon atoms, each of the hydrophilic groups being amino, alkylamino, dialkylamino, trialkylamino, tetraalkylammonium, hydroxyamino, hydroxyl, carboxyl, carboxylate, carbamate, carbamide, carbonate, phosphate, phosphite, sulfate, sulfite, or



[0563] Specific examples of X and Y include O, S, Si, C₁-C₆ alkylene,

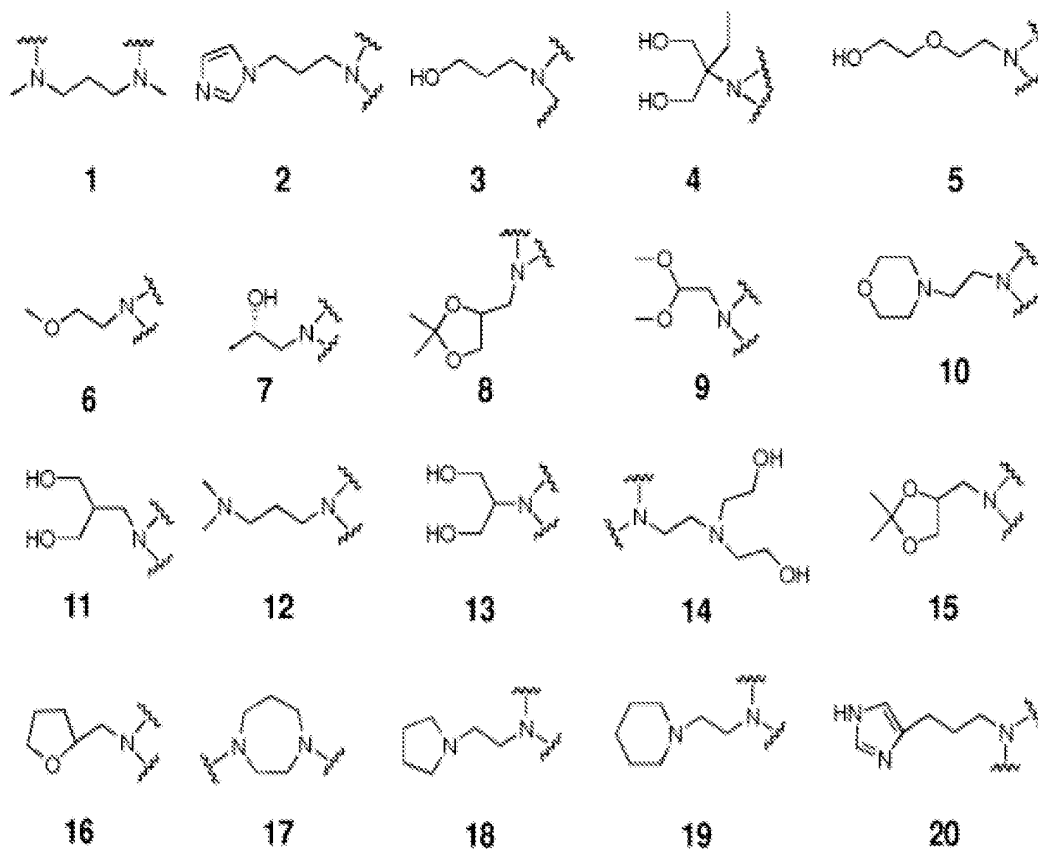


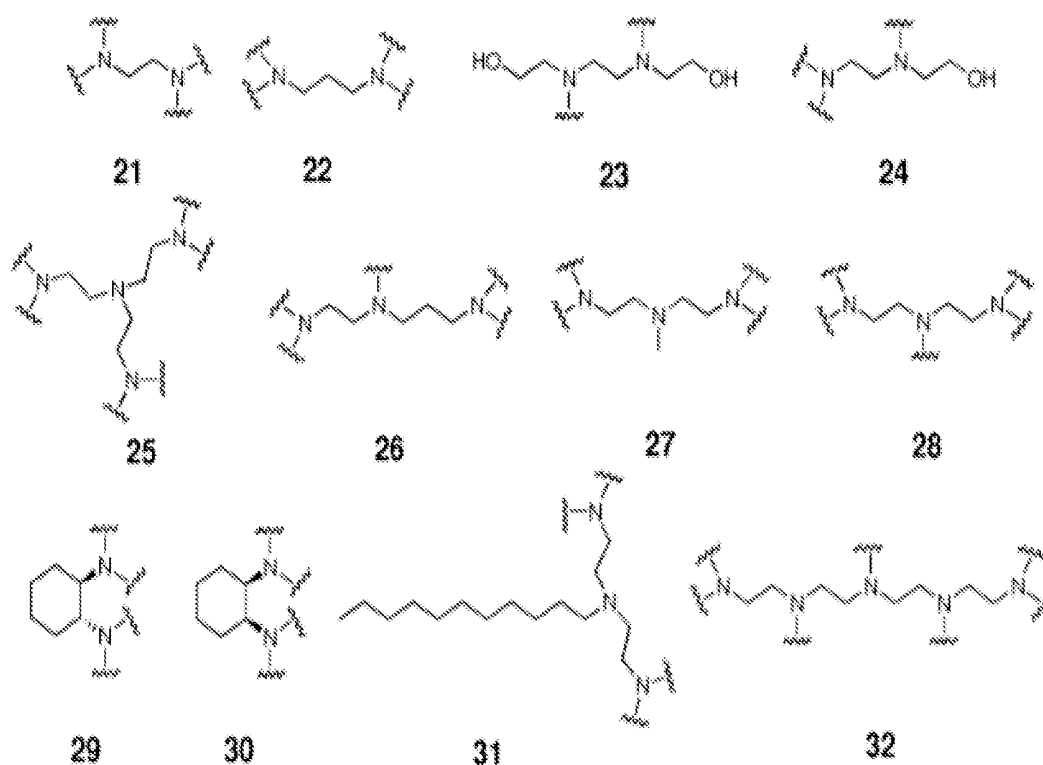


[0564] In some embodiments, a lipid-like compound of this invention, as shown in structure XXIII above, includes (i) a hydrophilic head, A; (ii) a hydrophobic tail, R₂-S-S-R₃; and (iii) a linker, X. Optionally, these compounds contain a second hydrophobic tail, R₅-S-S-R₆ and a second linker, Y.

[0565] The hydrophilic head of structure XXIII contains one or more hydrophilic functional groups, e.g., hydroxyl, carbonyl, carboxyl, amino, sulfhydryl, phosphate, amide, ester, ether, carbamate, carbonate, carbamide, and phosphodiester. These groups can form hydrogen bonds and are optionally positively or negatively charged.

[0566] Examples of the hydrophilic head include:

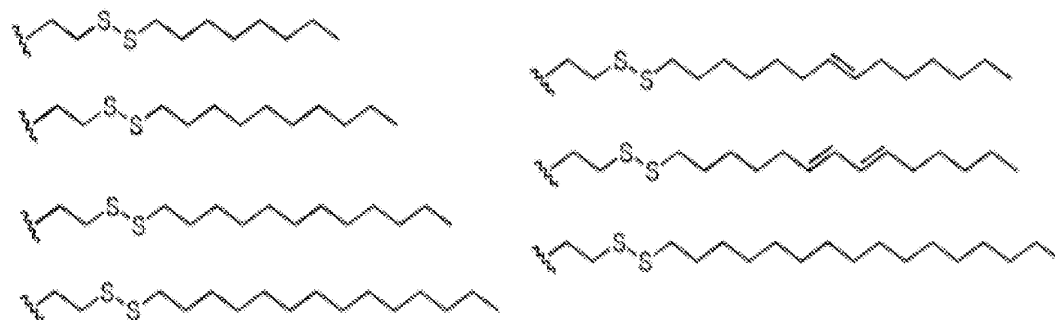




[0567] Other examples include those described in Akinc et al., *Nature Biotechnology*, 26, 561-69 (2008) and Mahon et al., US Patent Application Publication 2011/0293703.

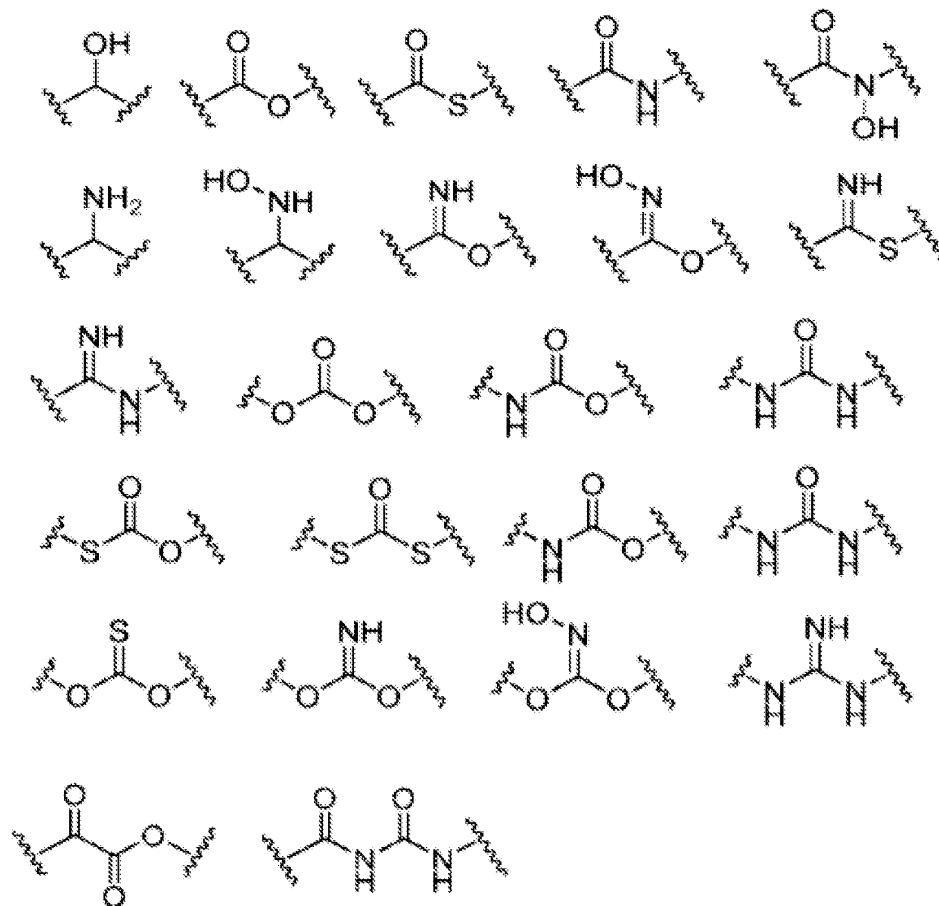
[0568] The hydrophobic tail of structure XXIII is a saturated or unsaturated, linear or branched, acyclic or cyclic, aromatic or nonaromatic hydrocarbon moiety containing a disulfide bond and 8-24 carbon atoms. One or more of the carbon atoms can be replaced with a heteroatom, such as N, O, P, B, S, Si, Sb, Al, Sn, As, Se, and Ge. The tail is optionally substituted with one or more groups described above. The lipid-like compounds containing this disulfide bond can be bioreducible.

[0569] Examples include:

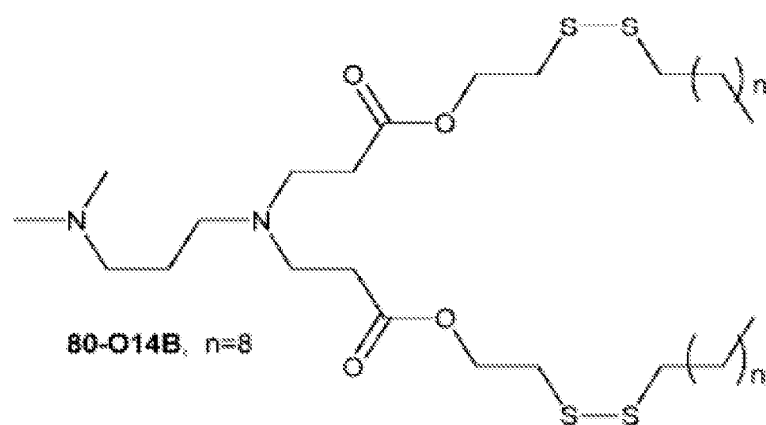
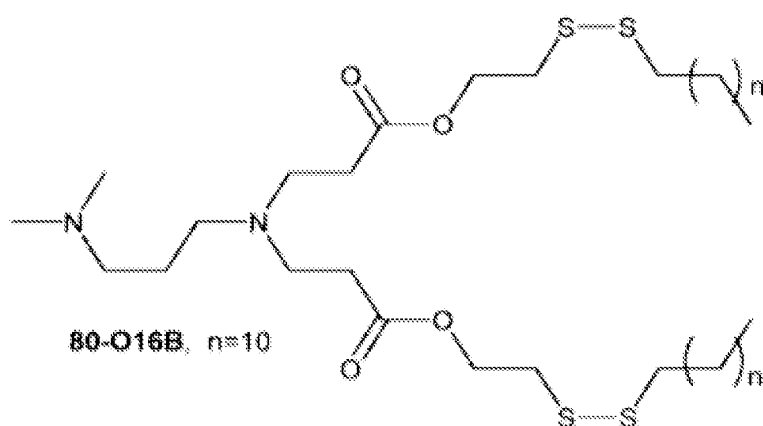
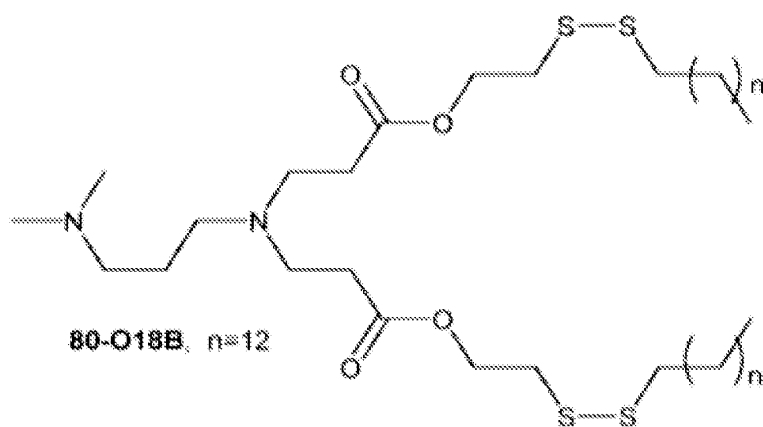


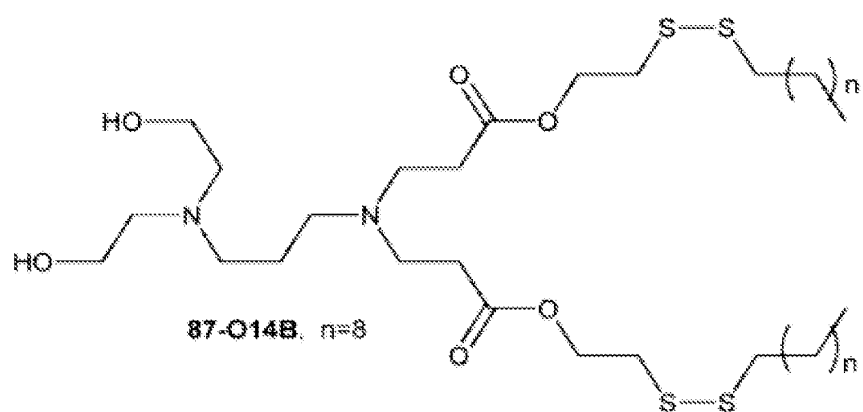
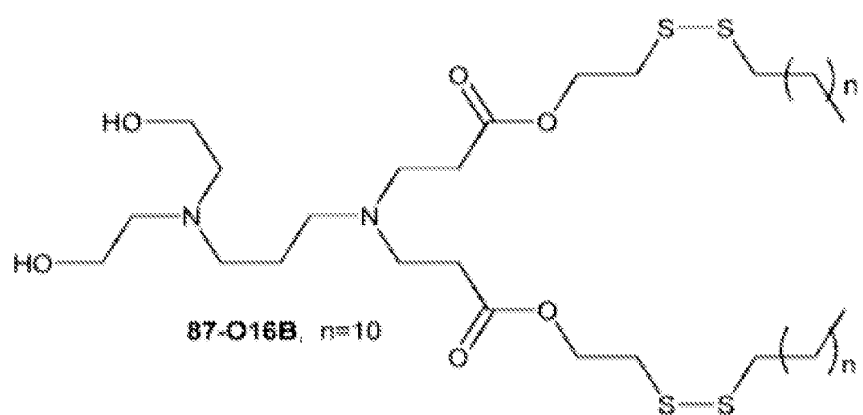
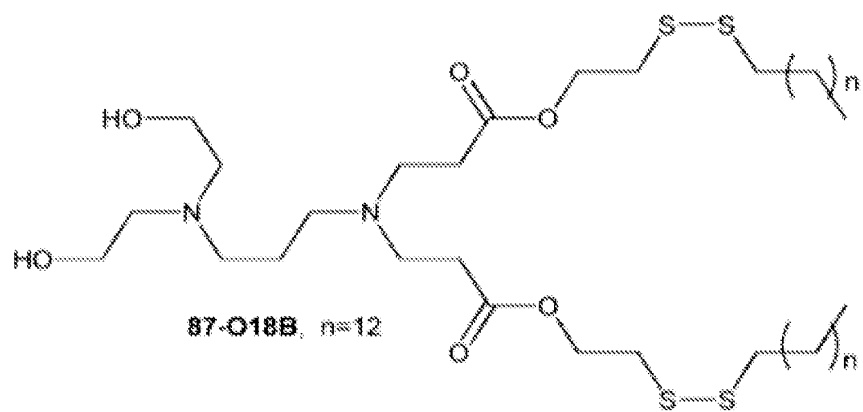
[0570] A linker of structure XXIII links the hydrophilic head and the hydrophobic tail.

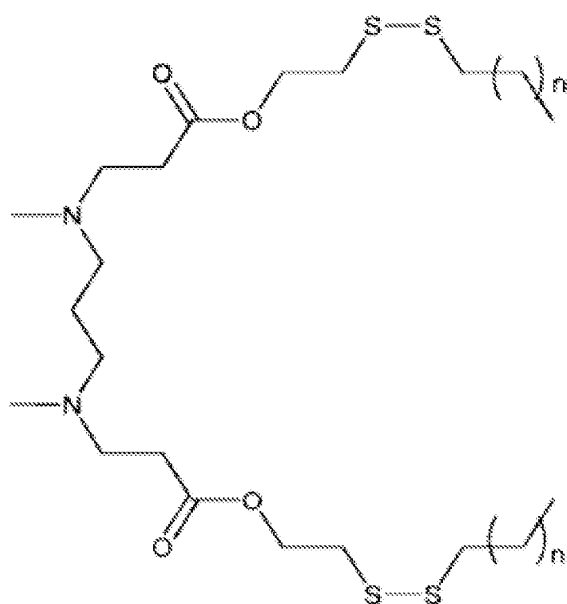
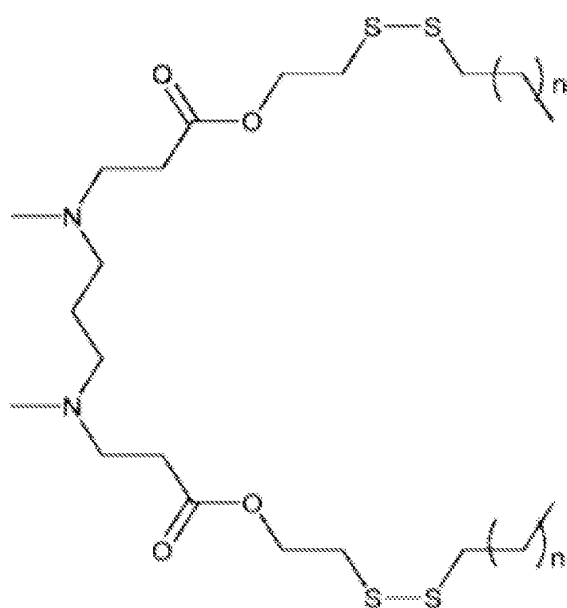
The linker can be any chemical group that is hydrophilic or hydrophobic, polar or non-polar, e.g., O, S, Si, amino, alkylene, ester, amide, carbamate, carbamide, carbonate, phosphate, phosphite, sulfate, sulfite, and thiosulfate. Examples include:



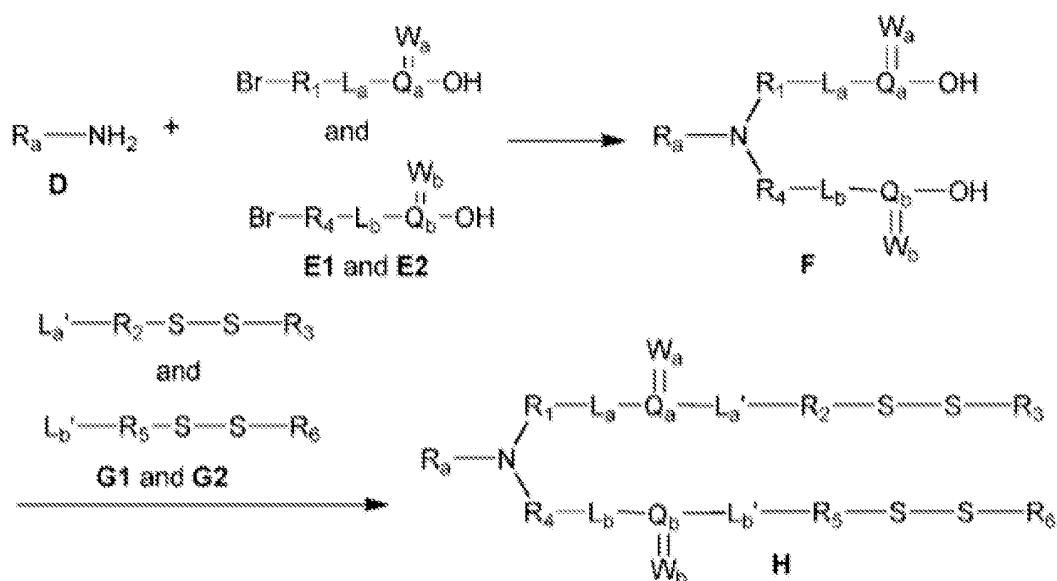
[0571] Shown below are exemplary lipid-like compounds of this invention:





**1-O16B, n=10****1-O18B, n=12**

[0572] The lipid-like compounds of structure XXIII can be prepared by methods well known the art. See Wang et al., ACS Synthetic Biology, 1, 403-07 (2012); Manoharan, et al., International Patent Application Publication WO 2008/042973; and Zugates et al., US Patent 8,071,082. The route shown below exemplifies synthesis of these lipid-like compounds:



[0573] Each of L_a , L_a' , L , and L' can be one of L_1 - L_{10} ; each of W_a and W_b , independently, is W or V; and R_a and R_1 - R_6 are defined above, as well as L_1 - L_{10} , W, and V.

[0574] In this exemplary synthetic route, an amine compound, i.e., compound D, reacts with bromides E1 and E2 to form compound F, which is then coupled with both G1 and G2 to afford the final product, i.e., compound H. One or both of the double bonds in this compound (shown above) can be reduced to one or two single bonds to obtain different lipid-like compounds of structure XXIII.

[0575] Other lipid-like compounds of this invention can be prepared using other suitable starting materials through the above-described synthetic route and others known in the art. The method set forth above can include an additional step(s) to add or remove suitable protecting groups in order to ultimately allow synthesis of the lipid-like compounds. In addition, various synthetic steps can be performed in an alternate sequence or order to give the desired material. Synthetic chemistry transformations and protecting group methodologies (protection and deprotection) useful in synthesizing applicable lipid-like compounds are known in the art, including, for example, R. Larock, *Comprehensive Organic Transformations* (2nd Ed., VCH Publishers 1999); P. G. M. Wuts and T. W. Greene, *Greene's Protective Groups in Organic Synthesis* (4th Ed., John Wiley and Sons 2007); L. Fieser and M. Fieser, *Fieser and Fieser's Reagents for Organic Synthesis* (John Wiley and Sons 1994); and L. Paquette, ed., *Encyclopedia of Reagents for Organic Synthesis* (2nd ed., John Wiley and Sons 2009) and subsequent editions thereof. Certain lipid-like compounds may contain a non-aromatic double bond and one or more asymmetric centers. Thus, they can occur as

racemates and racemic mixtures, single enantiomers, individual diastereomers, diastereomeric mixtures, and cis- or trans- isomeric forms. All such isomeric forms are contemplated.

[0576] As mentioned above, these lipid-like compounds are useful for delivery of pharmaceutical agents. They can be preliminarily screened for their efficacy in delivering pharmaceutical agents by an in vitro assay and then confirmed by animal experiments and clinic trials. Other methods will also be apparent to those of ordinary skill in the art.

[0577] Not to be bound by any theory, the lipid-like compounds of structure XXIII facilitate delivery of pharmaceutical agents by forming complexes, e.g., nanocomplexes and microparticles. The hydrophilic head of such a lipid-like compound, positively or negatively charged, binds to a moiety of a pharmaceutical agent that is oppositely charged and its hydrophobic moiety binds to a hydrophobic moiety of the pharmaceutical agent. Either binding can be covalent or non-covalent.

[0578] The above described complexes can be prepared using procedures described in publications such as Wang et al., ACS Synthetic Biology, 1, 403-07 (2012). Generally, they are obtained by incubating a lipid-like compound and a pharmaceutical agent in a buffer such as a sodium acetate buffer or a phosphate buffered saline ("PBS").

Hydrophilic groups

[0579] In certain embodiments, the selected hydrophilic functional group or moiety may alter or otherwise impart properties to the compound or to the transfer vehicle of which such compound is a component (e.g., by improving the transfection efficiencies of a lipid nanoparticle of which the compound is a component). For example, the incorporation of guanidinium as a hydrophilic head-group in the compounds disclosed herein may promote the fusogenicity of such compounds (or of the transfer vehicle of which such compounds are a component) with the cell membrane of one or more target cells, thereby enhancing, for example, the transfection efficiencies of such compounds. It has been hypothesized that the nitrogen from the hydrophilic guanidinium moiety forms a six-membered ring transition state which grants stability to the interaction and thus allows for cellular uptake of encapsulated materials. (Wender, *et al.*, Adv. Drug Del. Rev. (2008) 60: 452-472.) Similarly, the incorporation of one or more amino groups or moieties into the disclosed compounds (e.g., as a head-group) may further promote disruption of the endosomal/lysosomal membrane of the target cell by exploiting the fusogenicity of such amino groups. This is based not only on the pKa of the amino group of the composition, but also on the ability of the amino group to

undergo a hexagonal phase transition and fuse with the target cell surface, *i.e.* the vesicle membrane. (Koltover, *et al.* Science (1998) 281: 78-81.) The result is believed to promote the disruption of the vesicle membrane and release of the lipid nanoparticle contents into the target cell.

[0580] Similarly, in certain embodiments the incorporation of, for example, imidazole as a hydrophilic head-group in the compounds disclosed herein may serve to promote endosomal or lysosomal release of, for example, contents that are encapsulated in a transfer vehicle (*e.g.*, lipid nanoparticle) of the invention. Such enhanced release may be achieved by one or both of a proton-sponge mediated disruption mechanism and/or an enhanced fusogenicity mechanism. The proton-sponge mechanism is based on the ability of a compound, and in particular a functional moiety or group of the compound, to buffer the acidification of the endosome. This may be manipulated or otherwise controlled by the pKa of the compound or of one or more of the functional groups comprising such compound (*e.g.*, imidazole). Accordingly, in certain embodiments the fusogenicity of, for example, the imidazole-based compounds disclosed herein (*e.g.*, HGT4001 and HGT4004) are related to the endosomal disruption properties, which are facilitated by such imidazole groups, which have a lower pKa relative to other traditional ionizable lipids. Such endosomal disruption properties in turn promote osmotic swelling and the disruption of the liposomal membrane, followed by the transfection or intracellular release of the polynucleotide materials loaded or encapsulated therein into the target cell. This phenomenon can be applicable to a variety of compounds with desirable pKa profiles in addition to an imidazole moiety. Such embodiments also include multi-nitrogen based functionalities such as polyamines, polypeptide (histidine), and nitrogen-based dendritic structures.

[0581] Exemplary ionizable and/or cationic lipids are described in International PCT patent publications WO2015/095340, WO2015/199952, WO2018/011633, WO2017/049245, WO2015/061467, WO2012/040184, WO2012/000104, WO2015/074085, WO2016/081029, WO2017/004 143, WO2017/075531, WO2017/117528, WO2011/022460, WO2013/148541, WO2013/116126, WO2011/153120, WO2012/044638, WO2012/054365, WO2011/090965, WO2013/016058, WO2012/162210, WO2008/042973, WO2010/129709, WO2010/144740, WO20 12/099755, WO2013/049328, WO2013/086322, WO2013/086373, WO2011/071860, WO2009/132131, WO2010/048536, WO2010/088537, WO2010/054401, WO2010/054406, WO2010/054405, WO2010/054384, WO2012/016184, WO2009/086558, WO2010/042877, WO2011/000106, WO2011/000107, WO2005/120152, WO2011/141705, WO2013/126803, WO2006/007712, WO2011/038160, WO2005/121348, WO2011/066651, WO2009/127060,

WO2011/141704, WO2006/069782, WO2012/031043, WO2013/006825, WO2013/033563, WO2013/089151, WO2017/099823, WO2015/095346, and WO2013/086354, and US patent publications US2016/0311759, US2015/0376115, US2016/0151284, US2017/0210697, US2015/0140070, US2013/0178541, US2013/0303587, US2015/0141678, US2015/0239926, US2016/0376224, US2017/0119904, US2012/0149894, US2015/0057373, US2013/0090372, US2013/0274523, US2013/0274504, US2013/0274504, US2009/0023673, US2012/0128760, US2010/0324120, US2014/0200257, US2015/0203446, US2018/0005363, US2014/0308304, US2013/0338210, US2012/0101148, US2012/0027796, US2012/0058144, US2013/0323269, US2011/0117125, US2011/0256175, US2012/0202871, US2011/0076335, US2006/0083780, US2013/0123338, US2015/0064242, US2006/0051405, US2013/0065939, US2006/0008910, US2003/0022649, US2010/0130588, US2013/0116307, US2010/0062967, US2013/0202684, US2014/0141070, US2014/0255472, US2014/0039032, US2018/0028664, US2016/0317458, and US2013/0195920, the contents of all of which are incorporated herein by reference in their entirety. International patent application WO 2019/131770 is also incorporated herein by reference in its entirety.

8. PEG lipids

[0582] The use and inclusion of polyethylene glycol (PEG)-modified phospholipids and derivatized lipids such as derivatized ceramides (PEG-CER), including N-Octanoyl-Sphingosine-1-[Succinyl(Methoxy Polyethylene Glycol)-2000] (C8 PEG-2000 ceramide) in the liposomal and pharmaceutical compositions described herein is contemplated, preferably in combination with one or more of the compounds and lipids disclosed herein. Contemplated PEG-modified lipids include, but are not limited to, a polyethylene glycol chain of up to 5 kDa in length covalently attached to a lipid with alkyl chain(s) of C6-C20 length. In some embodiments, the PEG-modified lipid employed in the compositions and methods of the invention is 1,2-dimyristoyl-sn-glycerol, methoxypolyethylene Glycol (2000 MW PEG) “DMG-PEG2000.” The addition of PEG-modified lipids to the lipid delivery vehicle may prevent complex aggregation and may also provide a means for increasing circulation lifetime and increasing the delivery of the lipid-polynucleotide composition to the target tissues, (Klibanov *et al.* (1990) FEBS Letters, 268 (1): 235-237), or they may be selected to rapidly exchange out of the formulation *in vivo* (see U.S. Pat. No. 5,885,613). Particularly useful exchangeable lipids are PEG-ceramides having shorter acyl chains (*e.g.*, C14 or C18). The PEG-modified phospholipid and derivatized lipids of the present invention may comprise

a molar ratio from about 0% to about 20%, about 0.5% to about 20%, about 1% to about 15%, about 4% to about 10%, or about 2% of the total lipid present in a liposomal lipid nanoparticle.

[0583] In an embodiment, a PEG-modified lipid is described in International Pat. Appl. No. PCT/US2019/015913, which is incorporated herein by reference in their entirety. In an embodiment, a transfer vehicle comprises one or more PEG-modified lipids.

[0584] Non-limiting examples of PEG-modified lipids include PEG-modified phosphatidylethanolamines and phosphatidic acids, PEG-ceramide conjugates (*e.g.*, PEG-CerC14 or PEG-CerC20), PEG-modified dialkylamines and PEG-modified 1,2-diacyloxypropan-3-amines. In some further embodiments, a PEG-modified lipid may be, *e.g.*, PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, or a PEG-DSPE.

[0585] In some still further embodiments, the PEG-modified lipid includes, but is not limited to 1,2-dimyristoyl-sn-glycerol methoxypolyethylene glycol (PEG-DMG), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[amino(polyethylene glycol)] (PEG-DSPE), PEG-disteryl glycerol (PEG-DSG), PEG-dipalmitoleyl, PEG-dioleoyl, PEG-distearyl, PEG-diacylglyceramide (PEG-DAG), PEG-dipalmitoyl phosphatidylethanolamine (PEG-DPPE), or PEG-1,2-dimyristyloxypropyl-3-amine (PEG-c-DMA).

[0586] In various embodiments, a PEG-modified lipid may also be referred to as “PEGylated lipid” or “PEG-lipid.”

[0587] In one embodiment, the PEG-lipid is selected from the group consisting of a PEG-modified phosphatidylethanolamine, a PEG-modified phosphatidic acid, a PEG-modified ceramide, a PEG-modified dialkylamine, a PEG-modified diacylglycerol, a PEG-modified dialkylglycerol, and mixtures thereof.

[0588] In some embodiments, the lipid moiety of the PEG-lipids includes those having lengths of from about C₁₄ to about C₂₂, such as from about C₁₄ to about C₁₆. In some embodiments, a PEG moiety, for example a mPEG-NH₂, has a size of about 1000, about 2000, about 5000, about 10,000, about 15,000 or about 20,000 daltons. In one embodiment, the PEG-lipid is PEG2k-DMG.

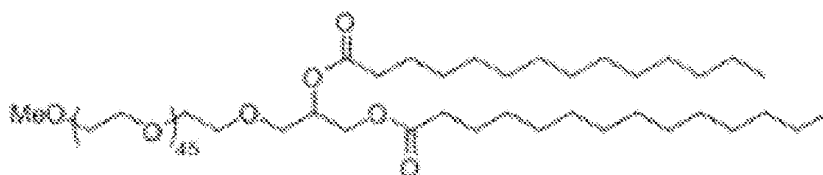
[0589] In one embodiment, the lipid nanoparticles described herein can comprise a lipid modified with a non-diffusible PEG. Non-limiting examples of non-diffusible PEGs include PEG-DSG and PEG-DSPE.

[0590] PEG-lipids are known in the art, such as those described in U.S. Pat. No. 8,158,601 and International Pat. Publ. No. WO2015/130584 A2, which are incorporated herein by reference in their entirety..

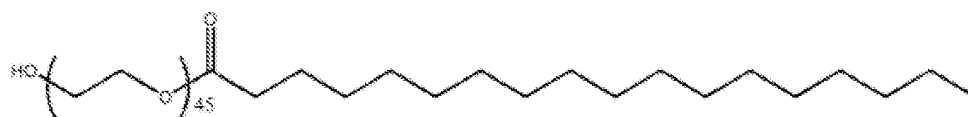
[0591] In various embodiments, lipids (e.g., PEG-lipids), described herein may be synthesized as described International Pat. Publ. No. PCT/US2016/000129, which is incorporated by reference in its entirety.

[0592] The lipid component of a lipid nanoparticle composition may include one or more molecules comprising polyethylene glycol, such as PEG or PEG-modified lipids. Such species may be alternately referred to as PEGylated lipids. A PEG lipid is a lipid modified with polyethylene glycol. A PEG lipid may be selected from the non-limiting group including PEG-modified phosphatidylethanolamines, PEG-modified phosphatidic acids, PEG-modified ceramides, PEG-modified dialkylamines, PEG-modified diacylglycerols, PEG-modified dialkylglycerols, and mixtures thereof. For example, a PEG lipid may be PEG-c-DMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, or a PEG-DSPE lipid.

[0593] In some embodiments the PEG-modified lipids are a modified form of PEG-DMG. PEG-DMG has the following structure:

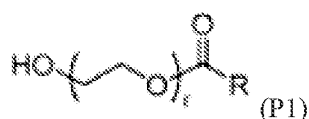


[0594] In some embodiments the PEG-modified lipids are a modified form of PEG-C18, or PEG-1. PEG-1 has the following structure



[0595] In one embodiment, PEG lipids useful in the present invention can be PEGylated lipids described in International Publication No. WO2012099755, the contents of which is herein incorporated by reference in its entirety. Any of these exemplary PEG lipids described herein may be modified to comprise a hydroxyl group on the PEG chain. In certain embodiments, the PEG lipid is a PEG-OH lipid. In certain embodiments, the PEG-OH lipid includes one or more hydroxyl groups on the PEG chain. In certain embodiments, a PEG-OH or hydroxy-PEGylated lipid comprises an -OH group at the terminus of the PEG chain. Each possibility represents a separate embodiment of the present invention.

[0596] In some embodiments, the PEG lipid is a compound of Formula (P1):



or a salt or isomer thereof, wherein:

r is an integer between 1 and 100;

R is C_{10-40} alkyl, C_{10-40} alkenyl, or C_{10-40} alkynyl; and optionally one or more methylene groups of R are independently replaced with C_{3-10} carbocyclylene, 4 to 10 membered heterocyclylene, C_{6-10} arylene, 4 to 10 membered heteroarylene, $-\text{N}(\text{R}^N)-$, $-\text{O}-$, $-\text{S}-$, $-\text{C}(\text{O})-$, $-\text{C}(\text{O})\text{N}(\text{R}^N)-$, $-\text{NR}^N\text{C}(\text{O})-$, $-\text{NR}^N\text{C}(\text{O})\text{N}(\text{R}^N)-$, $-\text{C}(\text{O})\text{O}-$, $-\text{OC}(\text{O})-$, $-\text{OC}(\text{O})\text{O}-$, $-\text{OC}(\text{O})\text{N}(\text{R}^N)-$, $-\text{NR}^N\text{C}(\text{O})\text{O}-$, $-\text{C}(\text{O})\text{S}-$, $-\text{SC}(\text{O})-$, $-\text{C}(=\text{NR}^N)-$, $-\text{C}(=\text{NR}^N)\text{N}(\text{R}^N)-$, $-\text{NR}^N\text{C}(=\text{NR}^N)-$, $-\text{NR}^N\text{C}(=\text{NR}^N)\text{N}(\text{R}^N)-$, $-\text{C}(\text{S})-$, $-\text{C}(\text{S})\text{N}(\text{R}^N)-$, $-\text{NR}^N\text{C}(\text{S})-$, $-\text{NR}^N\text{C}(\text{S})\text{N}(\text{R}^N)-$, $-\text{S}(\text{O})-$, $-\text{OS}(\text{O})-$, $-\text{S}(\text{O})\text{O}-$, $-\text{OS}(\text{O})\text{O}-$, $-\text{OS}(\text{O})_2-$, $-\text{S}(\text{O})_2\text{O}-$, $-\text{OS}(\text{O})_2\text{O}-$, $-\text{N}(\text{R}^N)\text{S}(\text{O})-$, $-\text{S}(\text{O})\text{N}(\text{R}^N)-$, $-\text{N}(\text{R}^N)\text{S}(\text{O})\text{N}(\text{R}^N)-$, $-\text{OS}(\text{O})\text{N}(\text{R}^N)-$, $-\text{N}(\text{R}^N)\text{S}(\text{O})\text{O}-$, $-\text{S}(\text{O})_2-$, $-\text{N}(\text{R}^N)\text{S}(\text{O})_2-$, $-\text{S}(\text{O})_2\text{N}(\text{R}^N)-$, $-\text{N}(\text{R}^N)\text{S}(\text{O})_2\text{N}(\text{R}^N)-$, $-\text{OS}(\text{O})_2\text{N}(\text{R}^N)-$, or $-\text{N}(\text{R}^N)\text{S}(\text{O})_2\text{O}-$; and

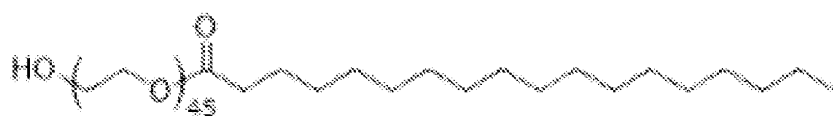
each instance of R^N is independently hydrogen, C_{1-6} alkyl, or a nitrogen protecting group.

[0597] For example, R is C_{17} alkyl. For example, the PEG lipid is a compound of Formula (P1-a):



or a salt or isomer thereof, wherein r is an integer between 1 and 100.

[0598] For example, the PEG lipid is a compound of the following formula:



9. Helper lipids

[0599] In some embodiments, the transfer vehicle (*e.g.*, LNP) described herein comprises one or more non-cationic helper lipids. In some embodiments, the helper lipid is a phospholipid. In some embodiments, the helper lipid is a phospholipid substitute or replacement. In some embodiments, the phospholipid or phospholipid substitute can be, for example, one or more saturated or (poly)unsaturated phospholipids, or phospholipid substitutes, or a combination thereof. In general, phospholipids comprise a phospholipid

moiety and one or more fatty acid moieties.

[0600] A phospholipid moiety can be selected, for example, from the non-limiting group consisting of phosphatidyl choline, phosphatidyl ethanolamine, phosphatidyl glycerol, phosphatidyl serine, phosphatidic acid, 2-lysophosphatidyl choline, and a sphingomyelin.

[0601] A fatty acid moiety can be selected, for example, from the non-limiting group consisting of lauric acid, myristic acid, myristoleic acid, palmitic acid, palmitoleic acid, stearic acid, oleic acid, linoleic acid, alpha-linolenic acid, erucic acid, phytanoic acid, arachidic acid, arachidonic acid, eicosapentaenoic acid, behenic acid, docosapentaenoic acid, and docosahexaenoic acid.

[0602] Phospholipids include, but are not limited to, glycerophospholipids such as phosphatidylcholines, phosphatidylethanolamines, phosphatidylserines, phosphatidylinositols, phosphatidyl glycerols, and phosphatidic acids. Phospholipids also include phosphosphingolipid, such as sphingomyelin.

[0603] In some embodiments, the helper lipid is a 1,2-distearoyl-177-glycero-3-phosphocholine (DSPC) analog, a DSPC substitute, oleic acid, or an oleic acid analog.

[0604] In some embodiments, a helper lipid is a non-phosphatidyl choline (PC) zwitterionic lipid, a DSPC analog, oleic acid, an oleic acid analog, or a DSPC substitute.

[0605] In some embodiments, a helper lipid is described in PCT/US2018/053569. Helper lipids suitable for use in a lipid composition of the disclosure include, for example, a variety of neutral, uncharged or zwitterionic lipids. Such helper lipids are preferably used in combination with one or more of the compounds and lipids disclosed herein. Examples of helper lipids include, but are not limited to, 5-heptadecylbenzene-1,3-diol (resorcinol), dipalmitoylphosphatidylcholine (DPPC), distearoylphosphatidylcholine (DSPC), phosphocholine (DOPC), dimyristoylphosphatidylcholine (DMPC), phosphatidylcholine (PLPC), 1,2-distearoyl-sn-glycero-3-phosphocholine (DAPC), phosphatidylethanolamine (PE), egg phosphatidylcholine (EPC), dilauryloylphosphatidylcholine (DLPC), dimyristoylphosphatidylcholine (DMPC), 1-myristoyl-2-palmitoyl phosphatidylcholine (MPPC), 1-palmitoyl-2-myristoyl phosphatidylcholine (PMPC), 1-palmitoyl-2-stearoyl phosphatidylcholine (PSPC), 1,2-diarachidoyl-sn-glycero-3-phosphocholine (DBPC), 1-stearoyl-2-palmitoyl phosphatidylcholine (SPPC), 1,2-dieicosenoyl-sn-glycero-3-phosphocholine (DEPC), palmitoyloleoyl phosphatidylcholine (POPC), lysophosphatidyl choline, dioleoyl phosphatidylethanol amine (DOPE) dilinoleoylphosphatidylcholine distearoylphosphatidylethanolamine (DSPE), dimyristoyl phosphatidylethanolamine (DMPE), dipalmitoyl phosphatidylethanolamine (DPPE), palmitoyloleoyl

phosphatidylethanolamine (POPE), lysophosphatidylethanolamine and combinations thereof. In one embodiment, the helper lipid may be distearoylphosphatidylcholine (DSPC) or dimyristoyl phosphatidyl ethanolamine (DMPE). In another embodiment, the helper lipid may be distearoylphosphatidylcholine (DSPC). Helper lipids function to stabilize and improve processing of the transfer vehicles. Such helper lipids are preferably used in combination with other excipients, for example, one or more of the ionizable lipids disclosed herein. In some embodiments, when used in combination with an ionizable lipid, the helper lipid may comprise a molar ratio of 5% to about 90%, or about 10% to about 70% of the total lipid present in the lipid nanoparticle.

10. Structural lipids

[0606] In an embodiment, a structural lipid is described in international patent application PCT/US2019/015913.

[0607] The transfer vehicles described herein comprise one or more structural lipids. Incorporation of structural lipids in the lipid nanoparticle may help mitigate aggregation of other lipids in the particle. Structural lipids can include, but are not limited to, cholesterol, fecosterol, ergosterol, bassicasterol, tomatidine, tomatine, ursolic, alpha-tocopherol, and mixtures thereof. In certain embodiments, the structural lipid is cholesterol. In certain embodiments, the structural lipid includes cholesterol and a corticosteroid (such as, for example, prednisolone, dexamethasone, prednisone, and hydrocortisone), or a combination thereof.

[0608] In some embodiments, the structural lipid is a sterol. In certain embodiments, the structural lipid is a steroid. In certain embodiments, the structural lipid is cholesterol. In certain embodiments, the structural lipid is an analog of cholesterol. In certain embodiments, the structural lipid is alpha-tocopherol.

[0609] The transfer vehicles described herein comprise one or more structural lipids. Incorporation of structural lipids in a transfer vehicle, *e.g.*, a lipid nanoparticle, may help mitigate aggregation of other lipids in the particle. In certain embodiments, the structural lipid includes cholesterol and a corticosteroid (such as, for example, prednisolone, dexamethasone, prednisone, and hydrocortisone), or a combination thereof.

[0610] In some embodiments, the structural lipid is a sterol. Structural lipids can include, but are not limited to, sterols (*e.g.*, phytosterols or zoosterols).

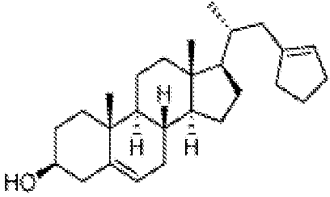
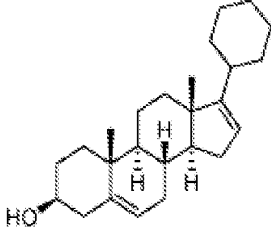
[0611] In certain embodiments, the structural lipid is a steroid. For example, sterols can

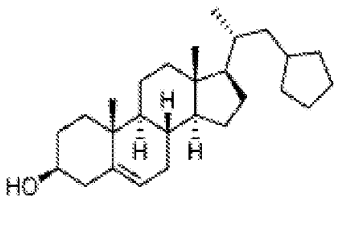
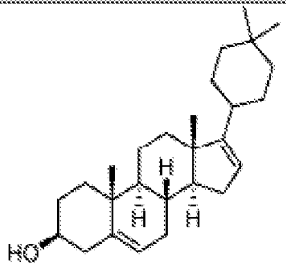
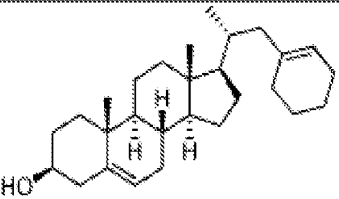
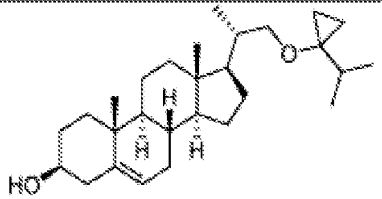
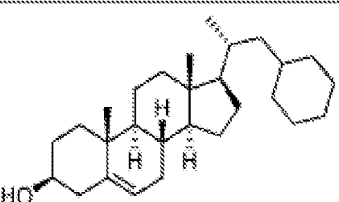
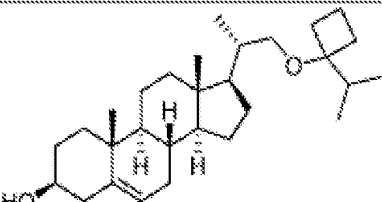
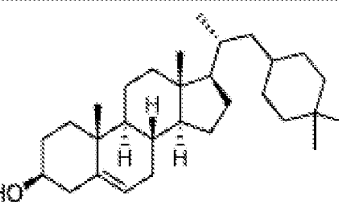
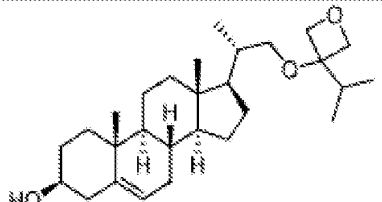
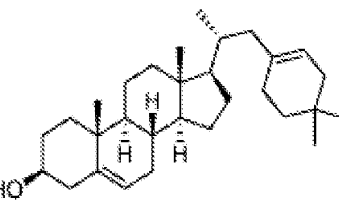
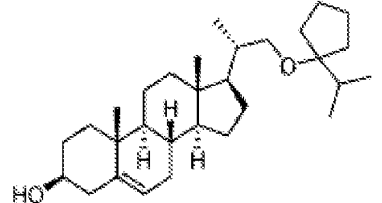
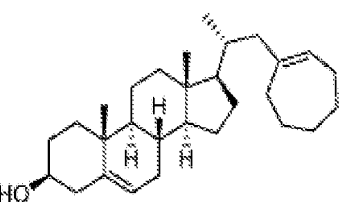
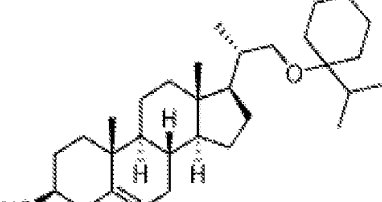
include, but are not limited to, cholesterol, β -sitosterol, fecosterol, ergosterol, sitosterol, campesterol, stigmasterol, brassicasterol, ergosterol, tomatidine, tomatine, ursolic acid, or alpha-tocopherol.

[0612] In some embodiments, a transfer vehicle includes an effective amount of an immune cell delivery potentiating lipid, *e.g.*, a cholesterol analog or an amino lipid or combination thereof, that, when present in a transfer vehicle, *e.g.*, an lipid nanoparticle, may function by enhancing cellular association and/or uptake, internalization, intracellular trafficking and/or processing, and/or endosomal escape and/or may enhance recognition by and/or binding to immune cells, relative to a transfer vehicle lacking the immune cell delivery potentiating lipid. Accordingly, while not intending to be bound by any particular mechanism or theory, in one embodiment, a structural lipid or other immune cell delivery potentiating lipid of the disclosure binds to C1q or promotes the binding of a transfer vehicle comprising such lipid to C1q. Thus, for *in vitro* use of the transfer vehicles of the disclosure for delivery of a nucleic acid molecule to an immune cell, culture conditions that include C1q are used (*e.g.*, use of culture media that includes serum or addition of exogenous C1q to serum-free media). For *in vivo* use of the transfer vehicles of the disclosure, the requirement for C1q is supplied by endogenous C1q.

[0613] In certain embodiments, the structural lipid is cholesterol. In certain embodiments, the structural lipid is an analog of cholesterol. In some embodiments, the structural lipid is a lipid in Table 16:

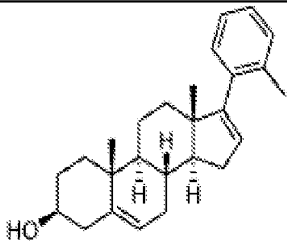
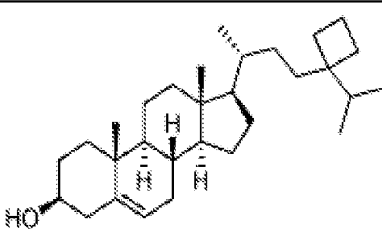
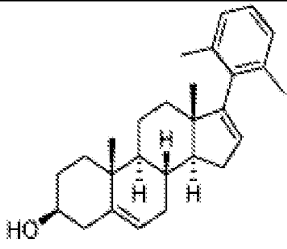
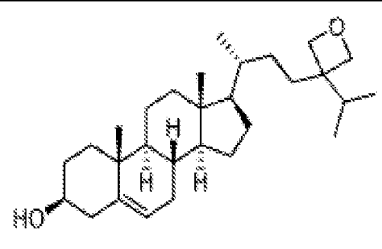
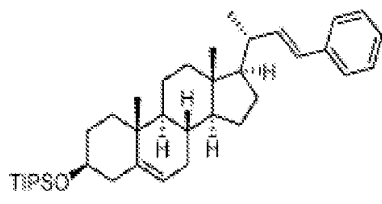
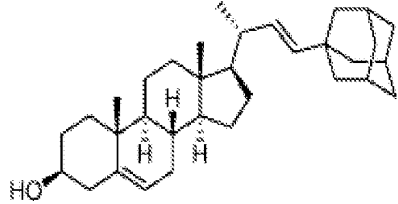
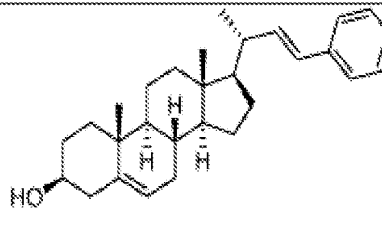
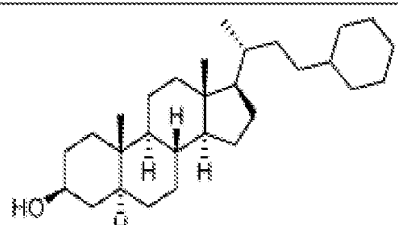
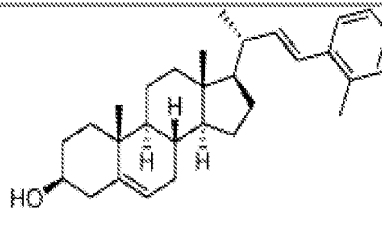
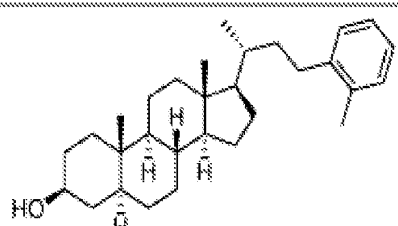
Table 16

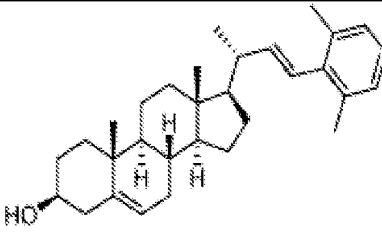
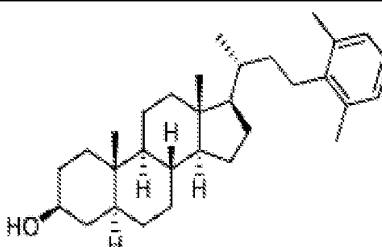
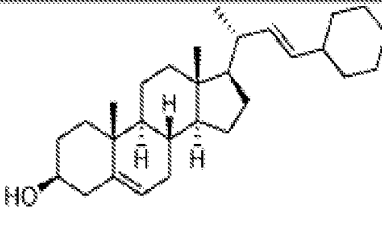
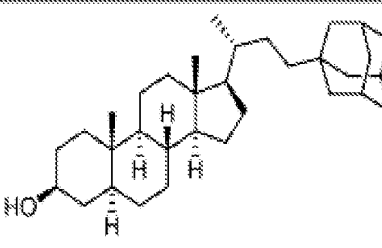
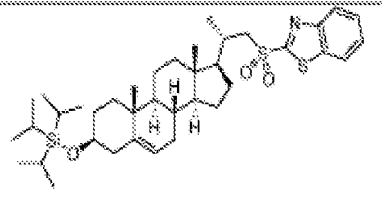
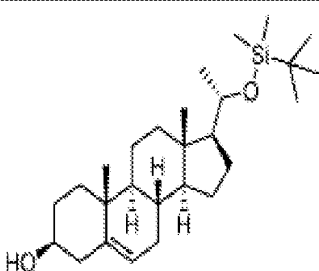
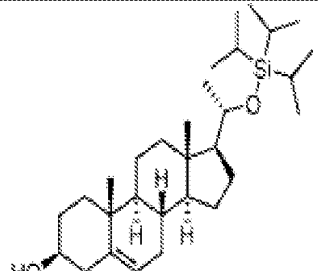
CMPD No. S-	Structure	CMPD No. S-	Structure
1		22	

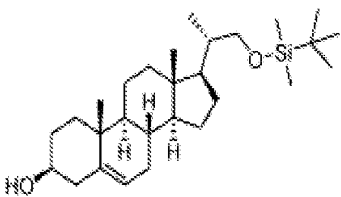
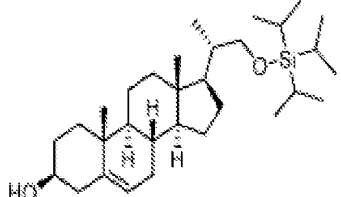
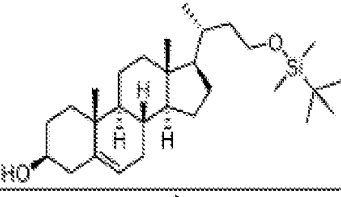
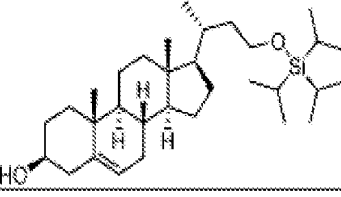
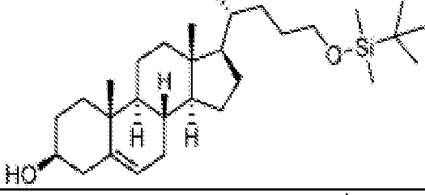
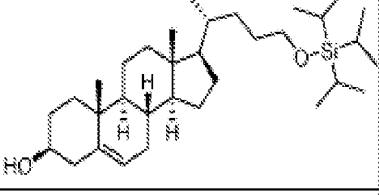
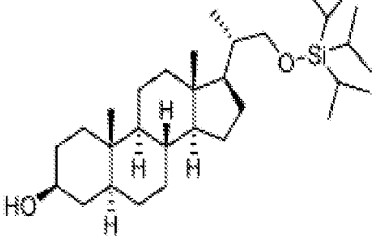
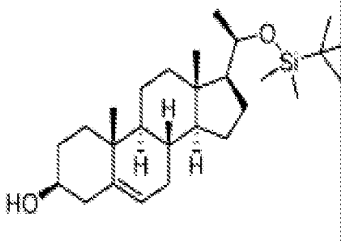
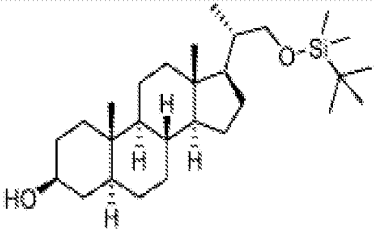
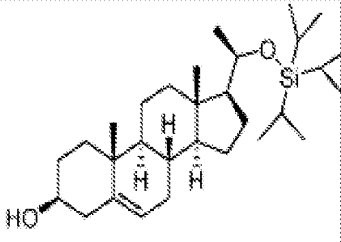
CMPD No. S-	Structure	CMPD No. S-	Structure
2		23	
3		24	
4		25	
5		26	
6		27	
7		28	

CMPD No. S-	Structure	CMPD No. S-	Structure
8		29	
9		30	
10		31	
11		32	
12		33	
13		34	

CMPD No. S-	Structure	CMPD No. S-	Structure
14		35	
15		36	
16		37	
17		38	
18		39	
19		40	

CMPD No. S-	Structure	CMPD No. S-	Structure
20		41	
21		42	
150		165	
154		169	
162		170	

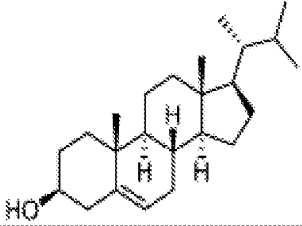
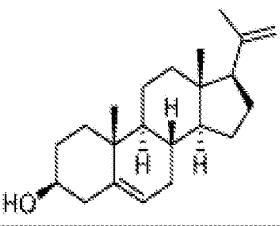
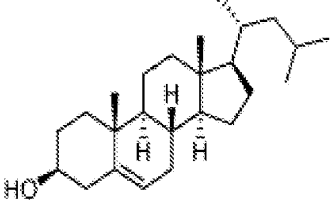
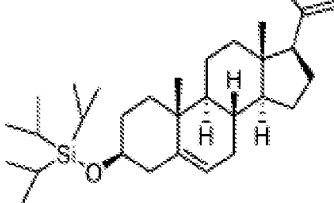
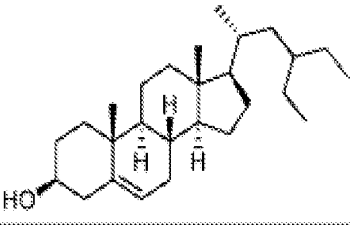
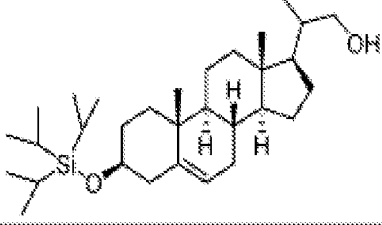
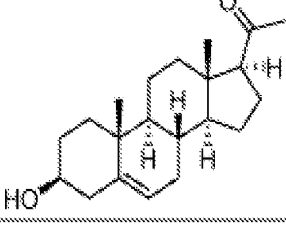
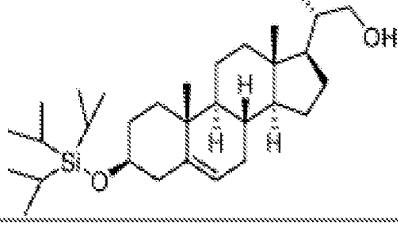
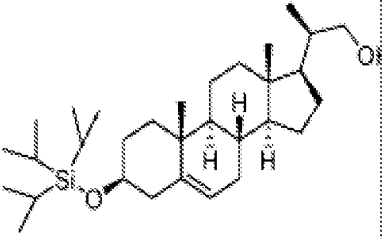
CMPD No. S-	Structure	CMPD No. S-	Structure
163		171	
164		172	
184			
CMPD No. S-	Structure	CMPD No. S-	Structure
43		47	

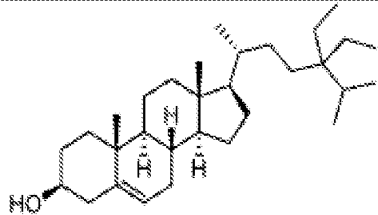
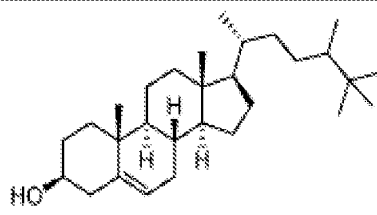
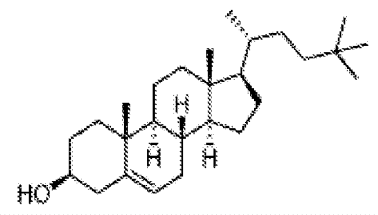
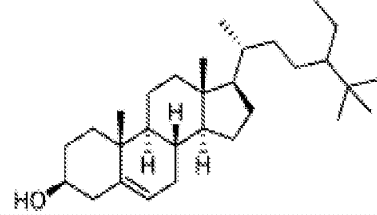
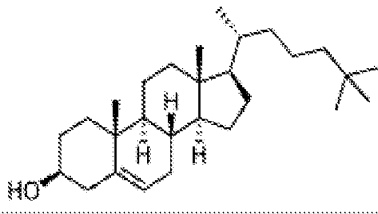
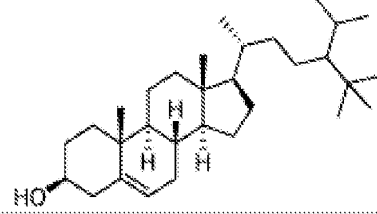
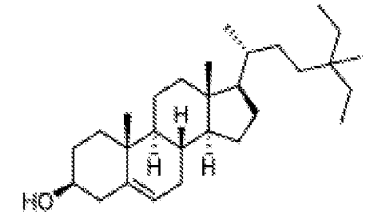
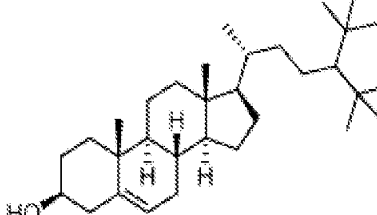
CMPD No. S-	Structure	CMPD No. S-	Structure
44		48	
45		49	
46		50	
175		177	
176		178	

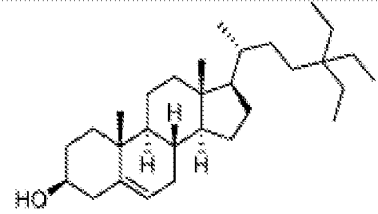
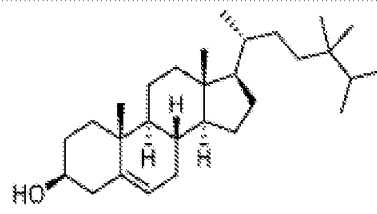
CMPD No. S-	Structure	CMPD No. S-	Structure
51		60	
52		61	
53		62	
54		63	
55		64	
56		65	
57		66	

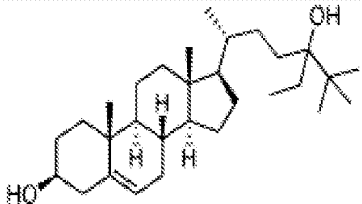
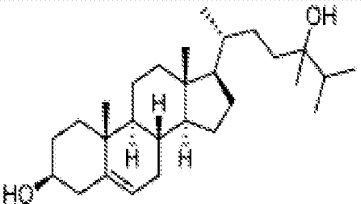
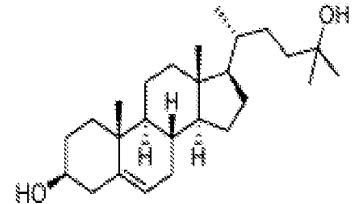
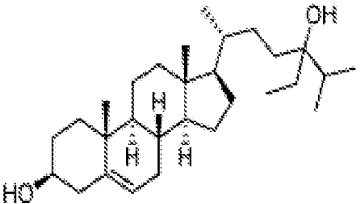
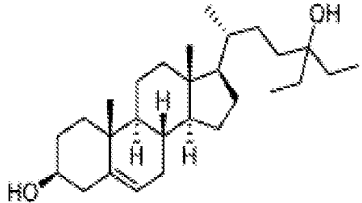
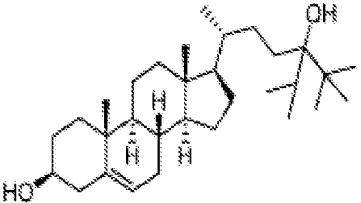
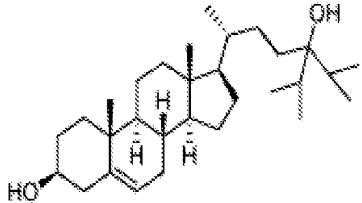
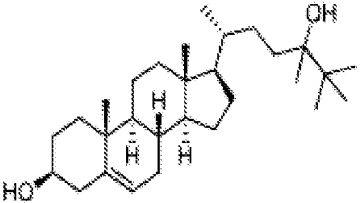
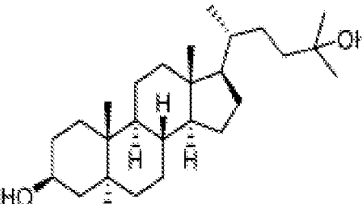
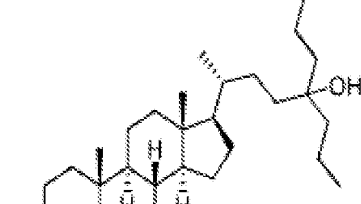
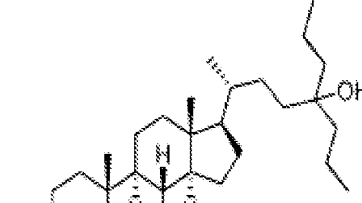
CMPD No. S-	Structure	CMPD No. S-	Structure
58		67	
59		149	
153			
CMPD No. S-	Structure	CMPD No. S-	Structure
68		71	
69		72	

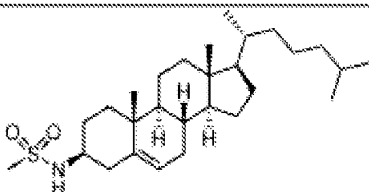
CMPD No. S-	Structure	CMPD No. S-	Structure
70		73	
74		77	
75		78	
76			
79		80	

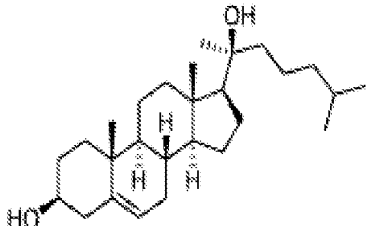
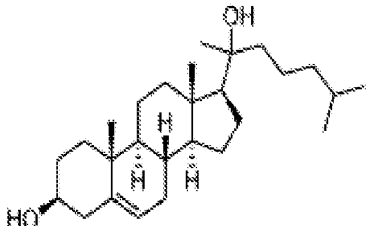
CMPD No. S-	Structure	CMPD No. S-	Structure
81		85	
82		86	
83		87	
84		152	
CMPD No. S-	Structure	CMPD No. S-	Structure
157			

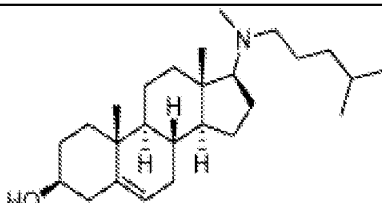
CMPD No. S-	Structure	CMPD No. S-	Structure
88		93	
89		94	
90		95	
91		96	

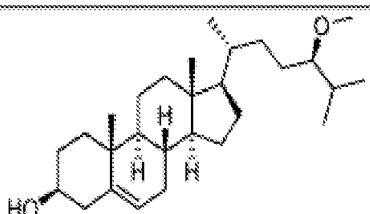
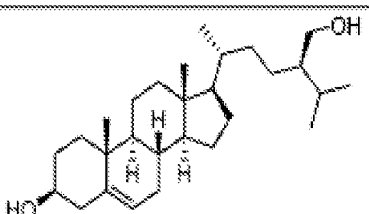
CMPD No. S-	Structure	CMPD No. S-	Structure
92		97	

CMPD No. S-	Structure	CMPD No. S-	Structure
98		102	
99		103	
100		104	
101		105	
CMPD No. S-	Structure	CMPD No. S-	Structure
180		182	
181			

CMPD No. S-	Structure
106	

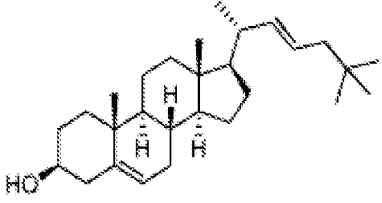
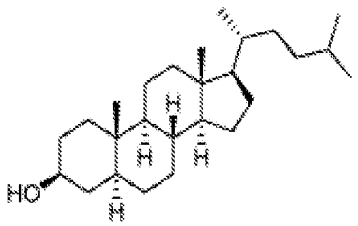
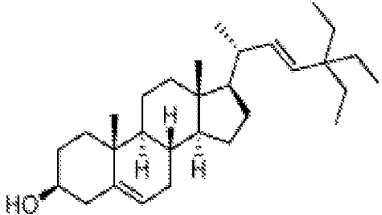
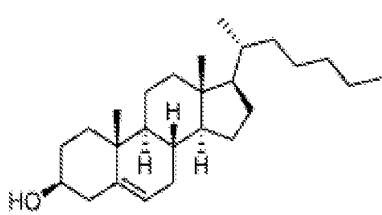
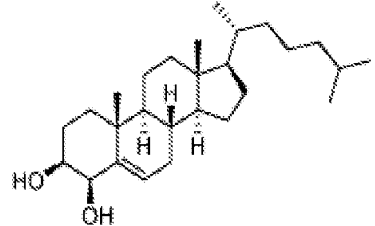
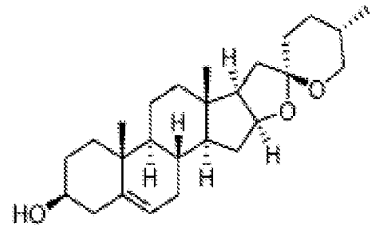
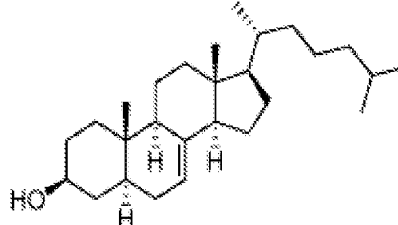
CMPD No. S-	Structure	CMPD No. S-	Structure
107		108	

CMPD No. S-	Structure
109	

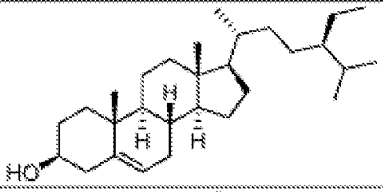
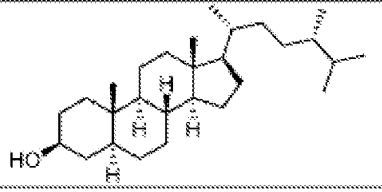
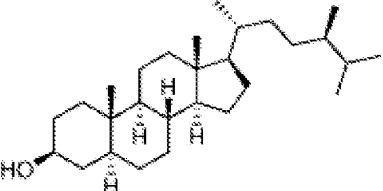
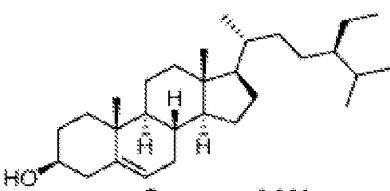
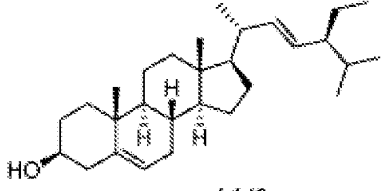
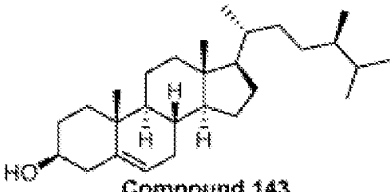
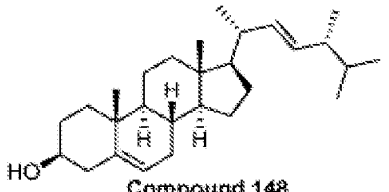
CMPD No. S-	Structure	CMPD No. S-	Structure
110		121	

CMPD No. S-	Structure	CMPD No. S-	Structure
111		122	
112		123	
113		124	
114		125	
115		126	
116		127	
117		128	

CMPD No. S-	Structure	CMPD No. S-	Structure
118		129	
119		130	
120		155	
156		167	
158		168	
160		173	

CMPD No. S-	Structure	CMPD No. S-	Structure
161		174	
166		179	
CMPD No. S-	Structure	CMPD No. S-	Structure
131		133	
132			

CMPD No. S-	Structure	CMPD No. S-	Structure
134		142	
135		143	
136		144	
137		145	
138		146	
139		147	
140		148	

CMPD No. S-	Structure	CMPD No. S-	Structure
141		151	
159			
Composition S- No.	Structure		
183	 Compound 141  compound 140  Compound 143  Compound 148		

11. LNP formulations

[0614] The formation of a lipid nanoparticle (LNP) described herein may be accomplished by any methods known in the art. For example, as described in U.S. Pat. Pub. No. US2012/0178702 A1, which is incorporated herein by reference in its entirety. Non-limiting examples of lipid nanoparticle compositions and methods of making them are described, for example, in Semple *et al.* (2010) *Nat. Biotechnol.* 28:172-176; Jayarama *et al.* (2012), *Angew. Chem. Int. Ed.*, 51:8529-8533; and Maier *et al.* (2013) *Molecular Therapy* 21, 1570-1578 (the contents of each of which are incorporated herein by reference in their entirety).

[0615] In one embodiment, the LNP formulation may be prepared by, *e.g.*, the methods described in International Pat. Pub. No. WO 2011/127255 or WO 2008/103276, the contents

of each of which are herein incorporated by reference in their entirety.

[0616] In one embodiment, LNP formulations described herein may comprise a polycationic composition. As a non-limiting example, the polycationic composition may be a composition selected from Formulae 1-60 of U.S. Pat. Pub. No. US2005/0222064 A1, the content of which is herein incorporated by reference in its entirety.

[0617] In one embodiment, the lipid nanoparticle may be formulated by the methods described in U.S. Pat. Pub. No. US2013/0156845 A1, and International Pat. Pub. No. WO2013/093648 A2 or WO2012/024526 A2, each of which is herein incorporated by reference in its entirety.

[0618] In one embodiment, the lipid nanoparticles described herein may be made in a sterile environment by the system and/or methods described in U.S. Pat. Pub. No. US2013/0164400 A1, which is incorporated herein by reference in its entirety.

[0619] In one embodiment, the LNP formulation may be formulated in a nanoparticle such as a nucleic acid-lipid particle described in U.S. Pat. No. 8,492,359, which is incorporated herein by reference in its entirety.

[0620] A nanoparticle composition may optionally comprise one or more coatings. For example, a nanoparticle composition may be formulated in a capsule, film, or tablet having a coating. A capsule, film, or tablet including a composition described herein may have any useful size, tensile strength, hardness, or density.

[0621] In some embodiments, the lipid nanoparticles described herein may be synthesized using methods comprising microfluidic mixers. Exemplary microfluidic mixers may include, but are not limited to, a slit interdigital micromixer including, but not limited to, those manufactured by Precision Nanosystems (Vancouver, BC, Canada), Microinnova (Allerheiligen bei Wildon, Austria) and/or a staggered herringbone micromixer (SHM) (Zhigaltsev, I.V. *et al.* (2012) *Langmuir*. 28:3633-40; Belliveau, N.M. *et al.* *Mol. Ther. Nucleic. Acids.* (2012) 1:e37; Chen, D. *et al.* *J. Am. Chem. Soc.* (2012) 134(16):6948-51; each of which is herein incorporated by reference in its entirety).

[0622] In some embodiments, methods of LNP generation comprising SHM, further comprise the mixing of at least two input streams wherein mixing occurs by microstructure-induced chaotic advection (MICA). According to this method, fluid streams flow through channels present in a herringbone pattern causing rotational flow and folding the fluids around each other. This method may also comprise a surface for fluid mixing wherein the surface changes orientations during fluid cycling. Methods of generating LNPs using SHM include those disclosed in U.S. Pat. Pub. Nos. US2004/0262223 A1 and US2012/0276209

A1, each of which is incorporated herein by reference in their entirety.

[0623] In one embodiment, the lipid nanoparticles may be formulated using a micromixer such as, but not limited to, a Slit Interdigital Microstructured Mixer (SIMM-V2) or a Standard Slit Interdigital Micro Mixer (SSIMM) or Caterpillar (CPMM) or Impinging-jet (IJMM) from the Institut für Mikrotechnik Mainz GmbH, Mainz Germany). In one embodiment, the lipid nanoparticles are created using microfluidic technology (*see*, Whitesides (2006) *Nature*. 442: 368-373; and Abraham *et al.* (2002) *Science*. 295: 647-651; each of which is herein incorporated by reference in its entirety). As a non-limiting example, controlled microfluidic formulation includes a passive method for mixing streams of steady pressure-driven flows in micro channels at a low Reynolds number (*see, e.g.,* Abraham *et al.* (2002) *Science*. 295: 647651; which is herein incorporated by reference in its entirety).

[0624] In one embodiment, the circRNA of the present invention may be formulated in lipid nanoparticles created using a micromixer chip such as, but not limited to, those from Harvard Apparatus (Holliston, MA), Dolomite Microfluidics (Royston, UK), or Precision Nanosystems (Van Couver, BC, Canada). A micromixer chip can be used for rapid mixing of two or more fluid streams with a split and recombine mechanism.

[0625] In one embodiment, the lipid nanoparticles may have a diameter from about 10 to about 100 nm such as, but not limited to, about 10 to about 20 nm, about 10 to about 30 nm, about 10 to about 40 nm, about 10 to about 50 nm, about 10 to about 60 nm, about 10 to about 70 nm, about 10 to about 80 nm, about 10 to about 90 nm, about 20 to about 30 nm, about 20 to about 40 nm, about 20 to about 50 nm, about 20 to about 60 nm, about 20 to about 70 nm, about 20 to about 80 nm, about 20 to about 90 nm, about 20 to about 100 nm, about 30 to about 40 nm, about 30 to about 50 nm, about 30 to about 60 nm, about 30 to about 70 nm, about 30 to about 80 nm, about 30 to about 90 nm, about 30 to about 100 nm, about 40 to about 50 nm, about 40 to about 60 nm, about 40 to about 70 nm, about 40 to about 80 nm, about 40 to about 90 nm, about 40 to about 100 nm, about 50 to about 60 nm, about 50 to about 70 nm, about 50 to about 80 nm, about 50 to about 90 nm, about 50 to about 100 nm, about 60 to about 70 nm, about 60 to about 80 nm, about 60 to about 90 nm, about 60 to about 100 nm, about 70 to about 80 nm, about 70 to about 90 nm, about 70 to about 100 nm, about 80 to about 90 nm, about 80 to about 100 nm and/or about 90 to about 100 nm. In one embodiment, the lipid nanoparticles may have a diameter from about 10 to 500 nm. In one embodiment, the lipid nanoparticle may have a diameter greater than 100 nm, greater than 150 nm, greater than 200 nm, greater than 250 nm, greater than 300 nm, greater than 350 nm, greater than 400 nm, greater than 450 nm, greater than 500 nm, greater than 550 nm,

greater than 600 nm, greater than 650 nm, greater than 700 nm, greater than 750 nm, greater than 800 nm, greater than 850 nm, greater than 900 nm, greater than 950 nm or greater than 1000 nm. Each possibility represents a separate embodiment of the present invention.

[0626] In some embodiments, a nanoparticle (*e.g.*, a lipid nanoparticle) has a mean diameter of 10-500 nm, 20-400 nm, 30-300 nm, or 40-200 nm. In some embodiments, a nanoparticle (*e.g.*, a lipid nanoparticle) has a mean diameter of 50-150 nm, 50-200 nm, 80-100 nm, or 80-200 nm.

[0627] In some embodiments, the lipid nanoparticles described herein can have a diameter from below 0.1 μm to up to 1 mm such as, but not limited to, less than 0.1 μm , less than 1.0 μm , less than 5 μm , less than 10 μm , less than 15 μm , less than 20 μm , less than 25 μm , less than 30 μm , less than 35 μm , less than 40 μm , less than 50 μm , less than 55 μm , less than 60 μm , less than 65 μm , less than 70 μm , less than 75 μm , less than 80 μm , less than 85 μm , less than 90 μm , less than 95 μm , less than 100 μm , less than 125 μm , less than 150 μm , less than 175 μm , less than 200 μm , less than 225 μm , less than 250 μm , less than 275 μm , less than 300 μm , less than 325 μm , less than 350 μm , less than 375 μm , less than 400 μm , less than 425 μm , less than 450 μm , less than 475 μm , less than 500 μm , less than 525 μm , less than 550 μm , less than 575 μm , less than 600 μm , less than 625 μm , less than 650 μm , less than 675 μm , less than 700 μm , less than 725 μm , less than 750 μm , less than 775 μm , less than 800 μm , less than 825 μm , less than 850 μm , less than 875 μm , less than 900 μm , less than 925 μm , less than 950 μm , less than 975 μm .

[0628] In another embodiment, LNPs may have a diameter from about 1 nm to about 100 nm, from about 1 nm to about 10 nm, about 1 nm to about 20 nm, from about 1 nm to about 30 nm, from about 1 nm to about 40 nm, from about 1 nm to about 50 nm, from about 1 nm to about 60 nm, from about 1 nm to about 70 nm, from about 1 nm to about 80 nm, from about 1 nm to about 90 nm, from about 5 nm to about 100 nm, from about 5 nm to about 10 nm, about 5 nm to about 20 nm, from about 5 nm to about 30 nm, from about 5 nm to about 40 nm, from about 5 nm to about 50 nm, from about 5 nm to about 60 nm, from about 5 nm to about 70 nm, from about 5 nm to about 80 nm, from about 5 nm to about 90 nm, about 10 to about 50 nm, from about 20 to about 50 nm, from about 30 to about 50 nm, from about 40 to about 50 nm, from about 20 to about 60 nm, from about 30 to about 60 nm, from about 40 to about 60 nm, from about 20 to about 70 nm, from about 30 to about 70 nm, from about 40 to about 70 nm, from about 50 to about 70 nm, from about 60 to about 70 nm, from about 20 to about 80 nm, from about 30 to about 80 nm, from about 40 to about 80 nm, from about 50 to about 80 nm, from about 60 to about 80 nm, from about 20 to about 90 nm,

from about 30 to about 90 nm, from about 40 to about 90 nm, from about 50 to about 90 nm, from about 60 to about 90 nm and/or from about 70 to about 90 nm. Each possibility represents a separate embodiment of the present invention.

[0629] A nanoparticle composition may be relatively homogenous. A polydispersity index may be used to indicate the homogeneity of a nanoparticle composition, *e.g.*, the particle size distribution of the nanoparticle compositions. A small (*e.g.*, less than 0.3) polydispersity index generally indicates a narrow particle size distribution. A nanoparticle composition may have a polydispersity index from about 0 to about 0.25, such as 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.20, 0.21, 0.22, 0.23, 0.24, or 0.25. In some embodiments, the polydispersity index of a nanoparticle composition may be from about 0.10 to about 0.20. Each possibility represents a separate embodiment of the present invention.

[0630] The zeta potential of a nanoparticle composition may be used to indicate the electrokinetic potential of the composition. For example, the zeta potential may describe the surface charge of a nanoparticle composition. Nanoparticle compositions with relatively low charges, positive or negative, are generally desirable, as more highly charged species may interact undesirably with cells, tissues, and other elements in the body. In some embodiments, the zeta potential of a nanoparticle composition may be from about -20 mV to about +20 mV, from about -20 mV to about +15 mV, from about -20 mV to about +10 mV, from about -20 mV to about +5 mV, from about -20 mV to about 0 mV, from about -20 mV to about -5 mV, from about -20 mV to about -10 mV, from about -20 mV to about -15 mV, from about -20 mV to about +20 mV, from about -20 mV to about +15 mV, from about -20 mV to about +10 mV, from about -20 mV to about +5 mV, from about -20 mV to about 0 mV, from about 0 mV to about +20 mV, from about 0 mV to about +15 mV, from about 0 mV to about +10 mV, from about 0 mV to about +5 mV, from about +5 mV to about +20 mV, from about +5 mV to about +15 mV, or from about +5 mV to about +10 mV. Each possibility represents a separate embodiment of the present invention.

[0631] The efficiency of encapsulation of a therapeutic agent describes the amount of therapeutic agent that is encapsulated or otherwise associated with a nanoparticle composition after preparation, relative to the initial amount provided. The encapsulation efficiency is desirably high (*e.g.*, close to 100%). The encapsulation efficiency may be measured, for example, by comparing the amount of therapeutic agent in a solution containing the nanoparticle composition before and after breaking up the nanoparticle composition with one or more organic solvents or detergents. Fluorescence may be used to

measure the amount of free therapeutic agent (*e.g.*, nucleic acids) in a solution. For the nanoparticle compositions described herein, the encapsulation efficiency of a therapeutic agent may be at least 50%, for example 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%. In some embodiments, the encapsulation efficiency may be at least 80%. In certain embodiments, the encapsulation efficiency may be at least 90%. Each possibility represents a separate embodiment of the present invention. In some embodiments, the lipid nanoparticle has a polydiversity value of less than 0.4. In some embodiments, the lipid nanoparticle has a net neutral charge at a neutral pH. In some embodiments, the lipid nanoparticle has a mean diameter of 50-200nm.

[0632] The properties of a lipid nanoparticle formulation may be influenced by factors including, but not limited to, the selection of the cationic lipid component, the degree of cationic lipid saturation, the selection of the non-cationic lipid component, the degree of noncationic lipid saturation, the selection of the structural lipid component, the nature of the PEGylation, ratio of all components and biophysical parameters such as size. As described herein, the purity of a PEG lipid component is also important to an LNP's properties and performance.

12. Methods

[0633] In one embodiment, a lipid nanoparticle formulation may be prepared by the methods described in International Publication Nos. WO2011127255 or WO2008103276, each of which is herein incorporated by reference in their entirety. In some embodiments, lipid nanoparticle formulations may be as described in International Publication No. WO2019131770, which is herein incorporated by reference in its entirety.

[0634] In some embodiments, circular RNA is formulated according to a process described in US patent application 15/809,680. In some embodiments, the present invention provides a process of encapsulating circular RNA in transfer vehicles comprising the steps of forming lipids into pre-formed transfer vehicles (*i.e.* formed in the absence of RNA) and then combining the pre-formed transfer vehicles with RNA. In some embodiments, the novel formulation process results in an RNA formulation with higher potency (peptide or protein expression) and higher efficacy (improvement of a biologically relevant endpoint) both *in vitro* and *in vivo* with potentially better tolerability as compared to the same RNA formulation prepared without the step of preforming the lipid nanoparticles (*e.g.*, combining the lipids directly with the RNA).

[0635] For certain cationic lipid nanoparticle formulations of RNA, in order to achieve high encapsulation of RNA, the RNA in buffer (*e.g.*, citrate buffer) has to be heated. In those processes or methods, the heating is required to occur before the formulation process (*i.e.* heating the separate components) as heating post-formulation (post-formation of nanoparticles) does not increase the encapsulation efficiency of the RNA in the lipid nanoparticles. In contrast, in some embodiments of the novel processes of the present invention, the order of heating of RNA does not appear to affect the RNA encapsulation percentage. In some embodiments, no heating (*i.e.* maintaining at ambient temperature) of one or more of the solutions comprising the pre-formed lipid nanoparticles, the solution comprising the RNA and the mixed solution comprising the lipid nanoparticle encapsulated RNA is required to occur before or after the formulation process.

[0636] RNA may be provided in a solution to be mixed with a lipid solution such that the RNA may be encapsulated in lipid nanoparticles. A suitable RNA solution may be any aqueous solution containing RNA to be encapsulated at various concentrations. For example, a suitable RNA solution may contain an RNA at a concentration of or greater than about 0.01 mg/ml, 0.05 mg/ml, 0.06 mg/ml, 0.07 mg/ml, 0.08 mg/ml, 0.09 mg/ml, 0.1 mg/ml, 0.15 mg/ml, 0.2 mg/ml, 0.3 mg/ml, 0.4 mg/ml, 0.5 mg/ml, 0.6 mg/ml, 0.7 mg/ml, 0.8 mg/ml, 0.9 mg/ml, or 1.0 mg/ml. In some embodiments, a suitable RNA solution may contain an RNA at a concentration in a range from about 0.01-1.0 mg/ml, 0.01-0.9 mg/ml, 0.01-0.8 mg/ml, 0.01-0.7 mg/ml, 0.01-0.6 mg/ml, 0.01-0.5 mg/ml, 0.01-0.4 mg/ml, 0.01-0.3 mg/ml, 0.01-0.2 mg/ml, 0.01-0.1 mg/ml, 0.05-1.0 mg/ml, 0.05-0.9 mg/ml, 0.05-0.8 mg/ml, 0.05-0.7 mg/ml, 0.05-0.6 mg/ml, 0.05-0.5 mg/ml, 0.05-0.4 mg/ml, 0.05-0.3 mg/ml, 0.05-0.2 mg/ml, 0.05-0.1 mg/ml, 0.1-1.0 mg/ml, 0.2-0.9 mg/ml, 0.3-0.8 mg/ml, 0.4-0.7 mg/ml, or 0.5-0.6 mg/ml.

[0637] Typically, a suitable RNA solution may also contain a buffering agent and/or salt. Generally, buffering agents can include HEPES, Tris, ammonium sulfate, sodium bicarbonate, sodium citrate, sodium acetate, potassium phosphate or sodium phosphate. In some embodiments, suitable concentration of the buffering agent may be in a range from about 0.1 mM to 100 mM, 0.5 mM to 90 mM, 1.0 mM to 80 mM, 2 mM to 70 mM, 3 mM to 60 mM, 4 mM to 50 mM, 5 mM to 40 mM, 6 mM to 30 mM, 7 mM to 20 mM, 8 mM to 15 mM, or 9 to 12 mM.

[0638] Exemplary salts can include sodium chloride, magnesium chloride, and potassium chloride. In some embodiments, suitable concentration of salts in an RNA solution may be in a range from about 1 mM to 500 mM, 5 mM to 400 mM, 10 mM to 350 mM, 15 mM to 300 mM, 20 mM to 250 mM, 30 mM to 200 mM, 40 mM to 190 mM, 50 mM to 180 mM, 50 mM

to 170 mM, 50 mM to 160 mM, 50 mM to 150 mM, or 50 mM to 100 mM.

[0639] In some embodiments, a suitable RNA solution may have a pH in a range from about 3.5-6.5, 3.5-6.0, 3.5-5.5, 3.5-5.0, 3.5-4.5, 4.0-5.5, 4.0-5.0, 4.0-4.9, 4.0-4.8, 4.0-4.7, 4.0-4.6, or 4.0-4.5.

[0640] Various methods may be used to prepare an RNA solution suitable for the present invention. In some embodiments, RNA may be directly dissolved in a buffer solution described herein. In some embodiments, an RNA solution may be generated by mixing an RNA stock solution with a buffer solution prior to mixing with a lipid solution for encapsulation. In some embodiments, an RNA solution may be generated by mixing an RNA stock solution with a buffer solution immediately before mixing with a lipid solution for encapsulation.

[0641] According to the present invention, a lipid solution contains a mixture of lipids suitable to form transfer vehicles for encapsulation of RNA. In some embodiments, a suitable lipid solution is ethanol based. For example, a suitable lipid solution may contain a mixture of desired lipids dissolved in pure ethanol (*i.e.* 100% ethanol). In another embodiment, a suitable lipid solution is isopropyl alcohol based. In another embodiment, a suitable lipid solution is dimethylsulfoxide-based. In another embodiment, a suitable lipid solution is a mixture of suitable solvents including, but not limited to, ethanol, isopropyl alcohol and dimethylsulfoxide.

[0642] A suitable lipid solution may contain a mixture of desired lipids at various concentrations. In some embodiments, a suitable lipid solution may contain a mixture of desired lipids at a total concentration in a range from about 0.1-100 mg/ml, 0.5-90 mg/ml, 1.0-80 mg/ml, 1.0-70 mg/ml, 1.0-60 mg/ml, 1.0-50 mg/ml, 1.0-40 mg/ml, 1.0-30 mg/ml, 1.0-20 mg/ml, 1.0-15 mg/ml, 1.0-10 mg/ml, 1.0-9 mg/ml, 1.0-8 mg/ml, 1.0-7 mg/ml, 1.0-6 mg/ml, or 1.0-5 mg/ml.

13. Targeting

[0643] The present invention also contemplates the discriminatory targeting of target cells and tissues by both passive and active targeting means. The phenomenon of passive targeting exploits the natural distributions patterns of a transfer vehicle *in vivo* without relying upon the use of additional excipients or means to enhance recognition of the transfer vehicle by target cells. For example, transfer vehicles which are subject to phagocytosis by the cells of the reticulo-endothelial system are likely to accumulate in the liver or spleen, and

accordingly may provide a means to passively direct the delivery of the compositions to such target cells.

[0644] Alternatively, the present invention contemplates active targeting, which involves the use of targeting moieties that may be bound (either covalently or non-covalently) to the transfer vehicle to encourage localization of such transfer vehicle at certain target cells or target tissues. For example, targeting may be mediated by the inclusion of one or more endogenous targeting moieties in or on the transfer vehicle to encourage distribution to the target cells or tissues. Recognition of the targeting moiety by the target tissues actively facilitates tissue distribution and cellular uptake of the transfer vehicle and/or its contents in the target cells and tissues (*e.g.*, the inclusion of an apolipoprotein-E targeting ligand in or on the transfer vehicle encourages recognition and binding of the transfer vehicle to endogenous low density lipoprotein receptors expressed by hepatocytes). As provided herein, the composition can comprise a moiety capable of enhancing affinity of the composition to the target cell. Targeting moieties may be linked to the outer bilayer of the lipid particle during formulation or post-formulation. These methods are well known in the art. In addition, some lipid particle formulations may employ fusogenic polymers such as PEAA, hemagglutinin, other lipopeptides (see U.S. patent application Ser. No. 08/835,281, and 60/083,294, which are incorporated herein by reference) and other features useful for *in vivo* and/or intracellular delivery. In other some embodiments, the compositions of the present invention demonstrate improved transfection efficacies, and/or demonstrate enhanced selectivity towards target cells or tissues of interest. Contemplated therefore are compositions which comprise one or more moieties (*e.g.*, peptides, aptamers, oligonucleotides, a vitamin or other molecules) that are capable of enhancing the affinity of the compositions and their nucleic acid contents for the target cells or tissues. Suitable moieties may optionally be bound or linked to the surface of the transfer vehicle. In some embodiments, the targeting moiety may span the surface of a transfer vehicle or be encapsulated within the transfer vehicle. Suitable moieties and are selected based upon their physical, chemical or biological properties (*e.g.*, selective affinity and/or recognition of target cell surface markers or features). Cell-specific target sites and their corresponding targeting ligand can vary widely. Suitable targeting moieties are selected such that the unique characteristics of a target cell are exploited, thus allowing the composition to discriminate between target and non-target cells. For example, compositions of the invention may include surface markers (*e.g.*, apolipoprotein-B or apolipoprotein-E) that selectively enhance recognition of, or affinity to hepatocytes (*e.g.*, by receptor-mediated recognition of and binding to such surface markers). As an example, the use of galactose as a

targeting moiety would be expected to direct the compositions of the present invention to parenchymal hepatocytes, or alternatively the use of mannose containing sugar residues as a targeting ligand would be expected to direct the compositions of the present invention to liver endothelial cells (*e.g.*, mannose containing sugar residues that may bind preferentially to the asialoglycoprotein receptor present in hepatocytes). (See Hillery A M, *et al.* "Drug Delivery and Targeting: For Pharmacists and Pharmaceutical Scientists" (2002) Taylor & Francis, Inc.) The presentation of such targeting moieties that have been conjugated to moieties present in the transfer vehicle (*e.g.*, a lipid nanoparticle) therefore facilitate recognition and uptake of the compositions of the present invention in target cells and tissues. Examples of suitable targeting moieties include one or more peptides, proteins, aptamers, vitamins and oligonucleotides.

[0645] In particular embodiments, a transfer vehicle comprises a targeting moiety. In some embodiments, the targeting moiety mediates receptor-mediated endocytosis selectively into a specific population of cells. In some embodiments, the targeting moiety is capable of binding to a T cell antigen. In some embodiments, the targeting moiety is capable of binding to a NK, NKT, or macrophage antigen. In some embodiments, the targeting moiety is capable of binding to a protein selected from the group CD3, CD4, CD8, PD-1, 4-1BB, and CD2. In some embodiments, the targeting moiety is an single chain Fv (scFv) fragment, nanobody, peptide, peptide-based macrocycle, minibody, heavy chain variable region, light chain variable region or fragment thereof. In some embodiments, the targeting moiety is selected from T-cell receptor motif antibodies, T-cell α chain antibodies, T-cell β chain antibodies, T-cell γ chain antibodies, T-cell δ chain antibodies, CCR7 antibodies, CD3 antibodies, CD4 antibodies, CD5 antibodies, CD7 antibodies, CD8 antibodies, CD11b antibodies, CD11c antibodies, CD16 antibodies, CD19 antibodies, CD20 antibodies, CD21 antibodies, CD22 antibodies, CD25 antibodies, CD28 antibodies, CD34 antibodies, CD35 antibodies, CD40 antibodies, CD45RA antibodies, CD45RO antibodies, CD52 antibodies, CD56 antibodies, CD62L antibodies, CD68 antibodies, CD80 antibodies, CD95 antibodies, CD117 antibodies, CD127 antibodies, CD133 antibodies, CD137 (4-1BB) antibodies, CD163 antibodies, F4/80 antibodies, IL-4R α antibodies, Sca-1 antibodies, CTLA-4 antibodies, GITR antibodies GARP antibodies, LAP antibodies, granzyme B antibodies, LFA-1 antibodies, transferrin receptor antibodies, and fragments thereof. In some embodiments, the targeting moiety is a small molecule binder of an ectoenzyme on lymphocytes. Small molecule binders of ectoenzymes include A2A inhibitors CD73 inhibitors, CD39 or adenosine receptors A2aR and A2bR. Potential small molecules include AB928.

[0646] In some embodiments, transfer vehicles are formulated and/or targeted as described in Shobaki N, Sato Y, Harashima H. Mixing lipids to manipulate the ionization status of lipid nanoparticles for specific tissue targeting. *Int J Nanomedicine*. 2018;13:8395–8410. Published 2018 Dec 10. In some embodiments, a transfer vehicle is made up of 3 lipid types. In some embodiments, a transfer vehicle is made up of 4 lipid types. In some embodiments, a transfer vehicle is made up of 5 lipid types. In some embodiments, a transfer vehicle is made up of 6 lipid types.

14. Target cells

[0647] Where it is desired to deliver a nucleic acid to an immune cell, the immune cell represents the target cell. In some embodiments, the compositions of the invention transfect the target cells on a discriminatory basis (*i.e.*, do not transfect non-target cells). The compositions of the invention may also be prepared to preferentially target a variety of target cells, which include, but are not limited to, T cells, B cells, macrophages, and dendritic cells.

[0648] In some embodiments, the target cells are deficient in a protein or enzyme of interest. For example, where it is desired to deliver a nucleic acid to a hepatocyte, the hepatocyte represents the target cell. In some embodiments, the compositions of the invention transfect the target cells on a discriminatory basis (*i.e.*, do not transfect non-target cells). The compositions of the invention may also be prepared to preferentially target a variety of target cells, which include, but are not limited to, hepatocytes, epithelial cells, hematopoietic cells, epithelial cells, endothelial cells, lung cells, bone cells, stem cells, mesenchymal cells, neural cells (*e.g.*, meninges, astrocytes, motor neurons, cells of the dorsal root ganglia and anterior horn motor neurons), photoreceptor cells (*e.g.*, rods and cones), retinal pigmented epithelial cells, secretory cells, cardiac cells, adipocytes, vascular smooth muscle cells, cardiomyocytes, skeletal muscle cells, beta cells, pituitary cells, synovial lining cells, ovarian cells, testicular cells, fibroblasts, B cells, T cells, reticulocytes, leukocytes, granulocytes and tumor cells.

[0649] The compositions of the invention may be prepared to preferentially distribute to target cells such as in the heart, lungs, kidneys, liver, and spleen. In some embodiments, the compositions of the invention distribute into the cells of the liver or spleen to facilitate the delivery and the subsequent expression of the circRNA comprised therein by the cells of the liver (*e.g.*, hepatocytes) or the cells of spleen (*e.g.*, immune cells). The targeted cells may function as a biological “reservoir” or “depot” capable of producing, and systemically excreting a functional protein or enzyme. Accordingly, in one embodiment of the invention

the transfer vehicle may target hepatocytes or immune cells and/or preferentially distribute to the cells of the liver or spleen upon delivery. In an embodiment, following transfection of the target hepatocytes or immune cells, the circRNA loaded in the vehicle are translated and a functional protein product is produced, excreted and systemically distributed. In other embodiments, cells other than hepatocytes (*e.g.*, lung, spleen, heart, ocular, or cells of the central nervous system) can serve as a depot location for protein production.

[0650] In one embodiment, the compositions of the invention facilitate a subject's endogenous production of one or more functional proteins and/or enzymes. In an embodiment of the present invention, the transfer vehicles comprise circRNA which encode a deficient protein or enzyme. Upon distribution of such compositions to the target tissues and the subsequent transfection of such target cells, the exogenous circRNA loaded into the transfer vehicle (*e.g.*, a lipid nanoparticle) may be translated *in vivo* to produce a functional protein or enzyme encoded by the exogenously administered circRNA (*e.g.*, a protein or enzyme in which the subject is deficient). Accordingly, the compositions of the present invention exploit a subject's ability to translate exogenously- or recombinantly-prepared circRNA to produce an endogenously-translated protein or enzyme, and thereby produce (and where applicable excrete) a functional protein or enzyme. The expressed or translated proteins or enzymes may also be characterized by the *in vivo* inclusion of native post-translational modifications which may often be absent in recombinantly-prepared proteins or enzymes, thereby further reducing the immunogenicity of the translated protein or enzyme.

[0651] The administration of circRNA encoding a deficient protein or enzyme avoids the need to deliver the nucleic acids to specific organelles within a target cell. Rather, upon transfection of a target cell and delivery of the nucleic acids to the cytoplasm of the target cell, the circRNA contents of a transfer vehicle may be translated and a functional protein or enzyme expressed.

[0652] In some embodiments, a circular RNA comprises one or more miRNA binding sites. In some embodiments, a circular RNA comprises one or more miRNA binding sites recognized by miRNA present in one or more non-target cells or non-target cell types (*e.g.*, Kupffer cells or hepatic cells) and not present in one or more target cells or target cell types (*e.g.*, hepatocytes or T cells). In some embodiments, a circular RNA comprises one or more miRNA binding sites recognized by miRNA present in an increased concentration in one or more non-target cells or non-target cell types (*e.g.*, Kupffer cells or hepatic cells) compared to one or more target cells or target cell types (*e.g.*, hepatocytes or T cells). miRNAs are thought to function by pairing with complementary sequences within RNA molecules,

resulting in gene silencing.

[0653] In some embodiments, the compositions of the invention transfect or distribute to target cells on a discriminatory basis (*i.e.* do not transfect non-target cells). The compositions of the invention may also be prepared to preferentially target a variety of target cells, which include, but are not limited to, hepatocytes, epithelial cells, hematopoietic cells, epithelial cells, endothelial cells, lung cells, bone cells, stem cells, mesenchymal cells, neural cells (*e.g.*, meninges, astrocytes, motor neurons, cells of the dorsal root ganglia and anterior horn motor neurons), photoreceptor cells (*e.g.*, rods and cones), retinal pigmented epithelial cells, secretory cells, cardiac cells, adipocytes, vascular smooth muscle cells, cardiomyocytes, skeletal muscle cells, beta cells, pituitary cells, synovial lining cells, ovarian cells, testicular cells, fibroblasts, B cells, T cells, reticulocytes, leukocytes, granulocytes and tumor cells.

15. Pharmaceutical compositions

[0654] In certain embodiments, provided herein are compositions (*e.g.*, pharmaceutical compositions) comprising a therapeutic agent provided herein. In some embodiments, the therapeutic agent is a circular RNA polynucleotide provided herein. In some embodiments the therapeutic agent is a vector provided herein. In some embodiments, the therapeutic agent is a cell comprising a circular RNA or vector provided herein (*e.g.*, a human cell, such as a human T cell). In certain embodiments, the composition further comprises a pharmaceutically acceptable carrier. In some embodiments, the compositions provided herein comprise a therapeutic agent provided herein in combination with other pharmaceutically active agents or drugs, such as anti-inflammatory drugs or antibodies capable of targeting B cell antigens, *e.g.*, anti-CD20 antibodies, *e.g.*, rituximab.

[0655] With respect to pharmaceutical compositions, the pharmaceutically acceptable carrier can be any of those conventionally used and is limited only by chemico-physical considerations, such as solubility and lack of reactivity with the active agent(s), and by the route of administration. The pharmaceutically acceptable carriers described herein, for example, vehicles, adjuvants, excipients, and diluents, are well-known to those skilled in the art and are readily available to the public. It is preferred that the pharmaceutically acceptable carrier be one which is chemically inert to the therapeutic agent(s) and one which has no detrimental side effects or toxicity under the conditions of use.

[0656] The choice of carrier will be determined in part by the particular therapeutic agent, as well as by the particular method used to administer the therapeutic agent. Accordingly,

there are a variety of suitable formulations of the pharmaceutical compositions provided herein.

[0657] In certain embodiments, the pharmaceutical composition comprises a preservative. In certain embodiments, suitable preservatives may include, for example, methylparaben, propylparaben, sodium benzoate, and benzalkonium chloride. Optionally, a mixture of two or more preservatives may be used. The preservative or mixtures thereof are typically present in an amount of about 0.0001% to about 2% by weight of the total composition.

[0658] In some embodiments, the pharmaceutical composition comprises a buffering agent. In some embodiments, suitable buffering agents may include, for example, citric acid, sodium citrate, phosphoric acid, potassium phosphate, and various other acids and salts. A mixture of two or more buffering agents optionally may be used. The buffering agent or mixtures thereof are typically present in an amount of about 0.001% to about 4% by weight of the total composition.

[0659] In some embodiments, the concentration of therapeutic agent in the pharmaceutical composition can vary, *e.g.*, less than about 1%, or at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, or about 50% or more by weight, and can be selected primarily by fluid volumes, and viscosities, in accordance with the particular mode of administration selected.

[0660] The following formulations for oral, aerosol, parenteral (*e.g.*, subcutaneous, intravenous, intraarterial, intramuscular, intradermal, intraperitoneal, and intrathecal), and topical administration are merely exemplary and are in no way limiting. More than one route can be used to administer the therapeutic agents provided herein, and in certain instances, a particular route can provide a more immediate and more effective response than another route.

[0661] Formulations suitable for oral administration can comprise or consist of (a) liquid solutions, such as an effective amount of the therapeutic agent dissolved in diluents, such as water, saline, or orange juice; (b) capsules, sachets, tablets, lozenges, and troches, each containing a predetermined amount of the active ingredient, as solids or granules; (c) powders; (d) suspensions in an appropriate liquid; and (e) suitable emulsions. Liquid formulations may include diluents, such as water and alcohols, for example, ethanol, benzyl alcohol and the polyethylene alcohols, either with or without the addition of a pharmaceutically acceptable surfactant. Capsule forms can be of the ordinary hard or soft shelled gelatin type containing, for example, surfactants, lubricants, and inert fillers, such as lactose, sucrose, calcium phosphate, and corn starch. Tablet forms can include one or more of

lactose, sucrose, mannitol, corn starch, potato starch, alginic acid, microcrystalline cellulose, acacia, gelatin, guar gum, colloidal silicon dioxide, croscarmellose sodium, talc, magnesium stearate, calcium stearate, zinc stearate, stearic acid, and other excipients, colorants, diluents, buffering agents, disintegrating agents, moistening agents, preservatives, flavoring agents, and other pharmacologically compatible excipients. Lozenge forms can comprise the therapeutic agent with a flavorant, usually sucrose, acacia or tragacanth. Pastilles can comprise the therapeutic agent with an inert base, such as gelatin and glycerin, or sucrose and acacia, emulsions, gels, and the like containing, in addition to, such excipients as are known in the art.

[0662] Formulations suitable for parenteral administration include aqueous and nonaqueous isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and nonaqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In some embodiments, the therapeutic agents provided herein can be administered in a physiologically acceptable diluent in a pharmaceutical carrier, such as a sterile liquid or mixture of liquids including water, saline, aqueous dextrose and related sugar solutions, an alcohol such as ethanol or hexadecyl alcohol, a glycol such as propylene glycol or polyethylene glycol, dimethylsulfoxide, glycerol, ketals such as 2,2-dimethyl-1,3-dioxolane-4-methanol, ethers, poly(ethyleneglycol) 400, oils, fatty acids, fatty acid esters or glycerides, or acetylated fatty acid glycerides with or without the addition of a pharmaceutically acceptable surfactant such as a soap or a detergent, suspending agent such as pectin, carbomers, methylcellulose, hydroxypropylmethylcellulose, or carboxymethylcellulose, or emulsifying agents and other pharmaceutical adjuvants.

[0663] Oils, which can be used in parenteral formulations in some embodiments, include petroleum, animal oils, vegetable oils, or synthetic oils. Specific examples of oils include peanut, soybean, sesame, cottonseed, corn, olive, petrolatum, and mineral oil. Suitable fatty acids for use in parenteral formulations include oleic acid, stearic acid, and isostearic acid. Ethyl oleate and isopropyl myristate are examples of suitable fatty acid esters.

[0664] Suitable soaps for use in certain embodiments of parenteral formulations include fatty alkali metal, ammonium, and triethanolamine salts, and suitable detergents include (a) cationic detergents such as, for example, dimethyl dialkyl ammonium halides, and alkyl pyridinium halides, (b) anionic detergents such as, for example, alkyl, aryl, and olefin sulfonates, alky, olefin, ether, and monoglyceride sulfates, and sulfosuccinates, (c) nonionic

detergents such as, for example, fatty amine oxides, fatty acid alkanolamides, and polyoxyethylenepolypropylene copolymers, (d) amphoteric detergents such as, for example, alkyl- β -aminopropionates, and 2-alkyl-imidazoline quaternary ammonium salts, and (e) mixtures thereof.

[0665] In some embodiments, the parenteral formulations will contain, for example, from about 0.5% to about 25% by weight of the therapeutic agent in solution. Preservatives and buffers may be used. In order to minimize or eliminate irritation at the site of injection, such compositions may contain one or more nonionic surfactants having, for example, a hydrophile-lipophile balance (HLB) of from about 12 to about 17. The quantity of surfactant in such formulations will typically range, for example, from about 5% to about 15% by weight. Suitable surfactants include polyethylene glycol, sorbitan, fatty acid esters such as sorbitan monooleate, and high molecular weight adducts of ethylene oxide with a hydrophobic base formed by the condensation of propylene oxide with propylene glycol. The parenteral formulations can be presented in unit-dose or multi-dose sealed containers, such as ampoules or vials, and can be stored in a freeze-dried (lyophilized) condition requiring only the addition of a sterile liquid excipient, for example, water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described.

[0666] In certain embodiments, injectable formulations are provided herein. The requirements for effective pharmaceutical carriers for injectable compositions are well-known to those of ordinary skill in the art (see, *e.g.*, *Pharmaceutics and Pharmacy Practice*, J.B. Lippincott Company, Philadelphia, PA, Banker and Chalmers, eds., pages 238-250 (1982), and *ASHP Handbook on Injectable Drugs*, Toissel, 4th ed, pages 622-630 (1986)).

[0667] In some embodiments, topical formulations are provided herein. Topical formulations, including those that are useful for transdermal drug release, are suitable in the context of certain embodiments provided herein for application to skin. In some embodiments, the therapeutic agent alone or in combination with other suitable components, can be made into aerosol formulations to be administered via inhalation. These aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like. They also may be formulated as pharmaceuticals for non-pressured preparations, such as in a nebulizer or an atomizer. Such spray formulations also may be used to spray mucosa.

[0668] In certain embodiments, the therapeutic agents provided herein can be formulated as inclusion complexes, such as cyclodextrin inclusion complexes, or liposomes. Liposomes

can serve to target the therapeutic agents to a particular tissue. Liposomes also can be used to increase the half-life of the therapeutic agents. Many methods are available for preparing liposomes, as described in, for example, Szoka *et al.*, Ann. Rev. Biophys. Bioeng., 9, 467 (1980) and U.S. Patents 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

[0669] In some embodiments, the therapeutic agents provided herein are formulated in time-released, delayed release, or sustained release delivery systems such that the delivery of the composition occurs prior to, and with sufficient time to, cause sensitization of the site to be treated. Such systems can avoid repeated administrations of the therapeutic agent, thereby increasing convenience to the subject and the physician, and may be particularly suitable for certain composition embodiments provided herein. In one embodiment, the compositions of the invention are formulated such that they are suitable for extended-release of the circRNA contained therein. Such extended-release compositions may be conveniently administered to a subject at extended dosing intervals. For example, in one embodiment, the compositions of the present invention are administered to a subject twice a day, daily or every other day. In an embodiment, the compositions of the present invention are administered to a subject twice a week, once a week, every ten days, every two weeks, every three weeks, every four weeks, once a month, every six weeks, every eight weeks, every three months, every four months, every six months, every eight months, every nine months or annually.

[0670] In some embodiments, a protein encoded by an inventive polynucleotide is produced by a target cell for sustained amounts of time. For example, the protein may be produced for more than one hour, more than four, more than six, more than 12, more than 24, more than 48 hours, or more than 72 hours after administration. In some embodiments the polypeptide is expressed at a peak level about six hours after administration. In some embodiments the expression of the polypeptide is sustained at least at a therapeutic level. In some embodiments, the polypeptide is expressed at least at a therapeutic level for more than one, more than four, more than six, more than 12, more than 24, more than 48, or more than 72 hours after administration. In some embodiments, the polypeptide is detectable at a therapeutic level in patient tissue (*e.g.*, liver or lung). In some embodiments, the level of detectable polypeptide is from continuous expression from the circRNA composition over periods of time of more than one, more than four, more than six, more than 12, more than 24, more than 48, or more than 72 hours after administration.

[0671] In certain embodiments, a protein encoded by an inventive polynucleotide is produced at levels above normal physiological levels. The level of protein may be increased as compared to a control. In some embodiments, the control is the baseline physiological

level of the polypeptide in a normal individual or in a population of normal individuals. In other embodiments, the control is the baseline physiological level of the polypeptide in an individual having a deficiency in the relevant protein or polypeptide or in a population of individuals having a deficiency in the relevant protein or polypeptide. In some embodiments, the control can be the normal level of the relevant protein or polypeptide in the individual to whom the composition is administered. In other embodiments, the control is the expression level of the polypeptide upon other therapeutic intervention, *e.g.*, upon direct injection of the corresponding polypeptide, at one or more comparable time points.

[0672] In certain embodiments, the levels of a protein encoded by an inventive polynucleotide are detectable at 3 days, 4 days, 5 days, or 1 week or more after administration. Increased levels of protein may be observed in a tissue (*e.g.*, liver or lung).

[0673] In some embodiments, the method yields a sustained circulation half-life of a protein encoded by an inventive polynucleotide. For example, the protein may be detected for hours or days longer than the half-life observed via subcutaneous injection of the protein or mRNA encoding the protein. In some embodiments, the half-life of the protein is 1 day, 2 days, 3 days, 4 days, 5 days, or 1 week or more.

[0674] Many types of release delivery systems are available and known to those of ordinary skill in the art. They include polymer based systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones, polyesteramides, polyorthoesters, polyhydroxybutyric acid, and polyanhydrides. Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are lipids including sterols such as cholesterol, cholesterol esters, and fatty acids or neutral fats such as mono-di-and tri-glycerides; hydrogel release systems; sylistic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples include, but are not limited to: (a) erosional systems in which the active composition is contained in a form within a matrix such as those described in U.S. Patents 4,452,775, 4,667,014, 4,748,034, and 5,239,660 and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patents 3,832,253 and 3,854,480. In addition, pump-based hardware delivery systems can be used, some of which are adapted for implantation.

[0675] In some embodiments, the therapeutic agent can be conjugated either directly or indirectly through a linking moiety to a targeting moiety. Methods for conjugating therapeutic agents to targeting moieties is known in the art. See, for instance, Wadwa *et al.*, J,

Drug Targeting 3:111 (1995) and U.S. Patent 5,087,616.

[0676] In some embodiments, the therapeutic agents provided herein are formulated into a depot form, such that the manner in which the therapeutic agent is released into the body to which it is administered is controlled with respect to time and location within the body (see, for example, U.S. Patent 4,450,150). Depot forms of therapeutic agents can be, for example, an implantable composition comprising the therapeutic agents and a porous or non-porous material, such as a polymer, wherein the therapeutic agents are encapsulated by or diffused throughout the material and/or degradation of the non-porous material. The depot is then implanted into the desired location within the body and the therapeutic agents are released from the implant at a predetermined rate.

16. Therapeutic methods

[0677] In certain aspects, provided herein is a method of treating and/or preventing a condition, *e.g.*, an autoimmune disorder or cancer.

[0678] In certain embodiments, the therapeutic agents provided herein are co-administered with one or more additional therapeutic agents (*e.g.*, in the same pharmaceutical composition or in separate pharmaceutical compositions). In some embodiments, the therapeutic agent provided herein can be administered first and the one or more additional therapeutic agents can be administered second, or vice versa. Alternatively, the therapeutic agent provided herein and the one or more additional therapeutic agents can be administered simultaneously.

[0679] In some embodiments, the subject is a mammal. In some embodiments, the mammal referred to herein can be any mammal, including, but not limited to, mammals of the order Rodentia, such as mice and hamsters, or mammals of the order Logomorpha, such as rabbits. The mammals may be from the order Carnivora, including Felines (cats) and Canines (dogs). The mammals may be from the order Artiodactyla, including Bovines (cows) and Swines (pigs), or of the order Perssodactyla, including Equines (horses). The mammals may be of the order Primates, Ceboids, or Simoids (monkeys) or of the order Anthropoids (humans and apes). Preferably, the mammal is a human.

17. Sequences

Table 17. IRES sequences.

SEQ ID NO	IRES	Sequence
1	EMCV-A	ccccctctccctccccctaacgttactggccgaagccgcttggaaataaggccggtgtgcgttt

		gtctatatgttatttccaccatattgcccgtcttttggcaatgtgaggggcccgaaacctggccctgtct tcttgacgagcattcctaggggtcttccctctcgccaaaggaatgcaaggctgtgtgaatgtcgtg aagggaagcagttcctctgggaagcttctgaagacaaacacgtctgtagcgacctttgcaggcag cgggaacccccacctggcgacaggtgctctcgcccaaaagccacgtgtataagatacacctg caaaaggcggcacaaccccagtgccacgtgtgtgagttggatagttgtggaagagtc aaatggctc tcctcaagcgtattcaacaaggggctgaaggatgccagaagggtacccattgtatgggatctgat ctggggcctcgggtgcacatgctttacatgtgttagtcgagggttaaaaaacgtctaggccccccgaa ccacggggacgtgtgtttctcttggaaaaacacgatgataatgtggccacaacc
2	EMCV-B	ctccccctcccccttactatactggccgaagccacttgggaataaggccgggtgtgcgtttgtcta catgctattttctaccgcattaccgtcttatggaatgtgagggtccagaacctgacctgtcttctga cgaacactcctaggggtcttccctctcgacaaaggagtgtaaggctgtgtgaatgtcgtgaaggga agcagttcctctgggaagcttctaaagacaaacacgtctgtagcgacctttgcaggcagcggaa ccccccacctgtgtacaggtgctctcgcccaaaagccacgtgtataagatacacctgcaaaag ggcgacacacccagtgccacgtgtgtgagttggatagttgtggaagagtc aaatggctcctca agcgtattcaacaaggggctgaaggatgccagaagggtacccattgtatgggatctgatctggg gctcgggtgcacgtgctttacacgtgtgtgagtcgagggtgaaaaaacgtctaggccccccgaacca cggggacgtgtgtttctcttggaaaaacacgattacaat
3	EMCV-Bf	ttgccagtcgtcgtatctgcaggctgggtccgtgactaccactcccccttcaacgtgaaggct acgatagtgccaggggcggtactgcccgaagtgcacccccaaacaacaacaacccccctc ccctcccccttactatactggccgaagccacttgggaataaggccgggtgtgcgtttgtctacat gctattttctaccgcattaccgtcttatggaatgtgagggtccagaacctgacctgtcttctgacg aacactcctaggggtcttccctctcgacaaaggagtgtaaggctgtgtgaatgtcgtgaaggga cagttcctctggaagcttctaaagacaaacacgtctgtagcgacctttgcaggcagcgggaacc ccccacctgggtgacaggtgctctcgcccaaaagccacgtgtataagatacacctgcaaaaggc ggcacaaccccagtgccacgtgtgtgagttggatagttgtggaagagtc aaatggctcctcaag cgtattcaacaaggggctgaaggatgccagaagggtacccattgtatgggatctgatctggggc ctcgggtgcacgtgctttacacgtgtgtgagtcgagggtgaaaaaacgtctaggccccccgaaccacg gggacgtgtgtttctcttggaaaaacacgattacaat
4	EMCV-Cf	ttgccagtcgtcgtatctgcaggctgggtccgtgactaccactcccccttcaacgtgaaggct acgatagtgccaggggcggtactgcccgaagtgcacccccaaacaacaacaacccccctc cctccTcccccttaacgttactggccgaagccgttgggaataaggccgggtgtgcgtttgtctatat gttattttccaccatattgcccgtcttttggcaatgtgaggggcccggaacctggccctgtcttctgac gagcattcctaggggtcttccctctcgccaaagggaatgcaaggctgtgtgaatgtcgtgaaggaa gcagttcctctggaagcttctgaagacaaacacgtctgtagcgacctttgcaggcagcgggaac ccccacctggcgacaggtgctctcgcccaaaagccacgtgtataagatacacctgcaaaagg cggcacaaccccagtgccacgtgtgtgagttggatagttgtggaagagtc aaatggctcctcaa gcgtattcaacaaggggctgaaggatgccagaagggtacccattgtatgggatctgatctgggg cctcgggtgcacatgctttacatgtgttagtcgagggttaaaaaacgtctaggccccccgaaccacgg ggacgtgtgtttctcttggaaaaacacgatgataat
5	EMCV pEC9	cccccccctaacgttactggccgaagccgttgggaataaggccgggtgtgcgtttgtctatatgttat tttccaccatattgcccgtcttttggcaatgtgaggggcccggaacctggccctgtcttctgacgagc attcctaggggtcttccctctcgccaaagggaatgcaaggctgtgtgaatgtcgtgaagggaagcag ttcctctggaagcttctgaagacaaacacgtctgtagcgacctttgcaggcagcgggaaccccc cacctggcgacaggtgctctcgcccaaaagccacgtgtataagatacacctgcaaaaggcggc acaaccccagtgccacgtgtgtgagttggatagttgtggaagagtc aaatggctcctcaagcgt attcaacaaggggctgaaggatgccagaagggtacccattgtatgggatctgatctggggcctc gggtgcacatgctttacatgtgttagtcgagggttaaaaaacgtctaggccccccgaaccacgggga

		cgtggttttctttgaaaaacacgatgataat
6	Picobimavirus	gtaaaftaatgctatttcaaaaatttaaacagaaaggagagatgttatgaaccgggtttacaagggtt catacatcgaaaatagcactacctggggcagccgacacactaacatcgtctgtttaaccagaagtg ttactgaaaggaggtattta
7	HCV QC64	acctgcccctaataaggggcgacactccgccaatgaatcactcccctgtgaggaaactactgtcttcac gcagaaagcgtctagccatggcggttagtatgagtgctgtacagcctccaggccccccccctccgg gagagccatagtggtctgctgggaaccgggtgagtacaccggaattgccgggaagactgggtccttfc ttggataaaccactctatgcccggacatttggcggtgccccgcaagactgctagccgagtagc gttgggttgcgaaggccttggtagtgcctgatagggtgcttgcgagtgccccgggaggtctcg tagaccgtgcac
8	Human Cosavirus E/D	ctacaagctttgtgtaaaaaacttttgggttttctcaagcttctcactacaggccccaaagat gtcctgaaggtaacccgtgtatctgaggatgagcaccatcgactaccggacctgcaaaatttgc aaacgcatgtgtatcccagccccctctctcggggagggggtttgctactcagcacaggatct gatcaggagatccacctccgggtgctttacacggggcggtgatttaaaattgcccaggcctggc gcacaacctaggggactagggtttcttattttaaagctgtcaat
9	Human Cosavirus F	gtcttaggacgacgcatgtgtatcccagccccgctacattggcggggttttgaagcacca gacactggatctgatcaggaggaggtagctgtttacagccccctctaaaaattgcccaggctcc ggccaccaacctaggggactagggtttcttattttaaattgtcatt
10	Human Cosavirus JMY	acatgggggagactgcatgtggcagtcctgaaacgtgtgtttgacgtctacattatagcagtggtg gtggagtactgcaaagatgtaccgtgctttacacgggttttgacccacacggcgtgttgacgt cgtaggggcagcagggtttattttcatfataaattcttactttctagctgcatgagtttcatgcatgacgg agtgtactcccggttcttcttgacagggtgcctccacgccccttggatcttaagtgaccaagt actgtgtgtggaggtgaagatagagagctcttgggaatgtcatgtggctgtgcccagggtgtga gcgatgccattcgtgtgtgcggtttctctctgtgtgacacgagcctcacaggccaaaagccccg tcgaaaggacccgaatgggtgagtgacctgactccccctgcatagtttggattaggaacttg aggaaftctgtataaattctatcacatcaggccccaaagatgtcctgaaggtaacctgtgtatctg aggatgagcaccacgactacccggacttgcattagcagacacatgtgtgtgcccagccccacct ctcagagggtgggtgtgtcactcagcacaggatctgatcaggagccccgctgtgtgtttaca ctcagcggggttaaaaattgcccaggcctggcacaacaacctaggggactagggtttctattt ttgtaattatgtcaat
11	Rhinovirus NAT001	gtgacaatcagccagattgttaacgggtcaagcacttctgtttcccggtaccctgtatagcttacc cgaggcgaaaagttaggttatcggtatccgcaaagtgcctacgagaagcctagtagcactttgaa gcctatgctgtgctcactcaactgtttaccagcagtagacctggcagatgaggctagatgtcccc accagcgatggtgatctagcctgctggtgctgctgcacactctattagtgtaagccagaaagt gacaagggtgaagagcctattgtgctcactttgagtcctccggccccgtaattggctaatcctaa ccccgtagctgtgtcatgtaatccaacatgtctgcagtcgtaattgggcaactatgggatggaaacaa ctactttgggtgtcgtgtttctgtttttcttattgcttcttatggtgacaactgtagtattacattgtta cc
12	HRV14	ttaaaacagcggatgggtatccaccattcgacccattgggtgtagtactctggtactatgtactttg tacgctgttttcccccaaccaccttcttaaaattccaccatgaaacgttagaagcttgacattta aagtacaataggtggcgccatatacaatgggtgtctatgtacaagcacttctgtttccaggagcgag gtataggctgtaccactgccccaaagcctttaaccgttatccgccaaccaactacgtaacagttagt accatctgtttctgactggacgttcgatcagggtgattttccctccactagtttggctgatgaggcta ggaaftccccacgggtgacctgttctagcctgctgtggcgccaaccagcttatgctggacgc ccttttaaggacatggtgtgaagactcgcattgtcttgggtgtgagtcctccggccccgtgaatcg ctaaccttaaccttagagccttatgcccagatccagtggtgtgaagctgtaattgagcaattccggg acgggacccgactactttgggtgtccgtgtttctcattttctcatattgtcttatggtcacagcatatata

[illegible]

		cttggctcagcatatctgaggcgacggttccgctgctccagaccagtcacagcaatggacgggcccagtggtgcgtagctgcgttccgggtttccggcgcatgtttggcgaaacgctgaggtaaagggttggtgccc aacgcccgttaatttgggtatacctcaagaccacccagggaatgccaggggagggtaccccacttcggt gggactcagaccctgggctaattgtctacgggtggttctctctgcttccacttctcttttctggcatg
18	Salivirus NG-J1	tatggcaggcgggcgttggtagcggcctggccacccacagcaagaatccatcatctgctectca cccccaattttcccttttctccctgcacacattacgcttactcgcattgtgcattgagtggtgcatgtgt tgaacaaacagctacactcacatggggggcggggtttccggccctacggcctctcgcgaggccac cccttccctcccttataactacagtgcttggtaggtgaagcctcgtatcccccgaggagctgctc acgtggcaactgtggggaccagacaggttatcaaggcaccgcgttcttccgcttcagggagtat ccctactagtgaattctagcggggctctgcttgggtgccaaacctccccaaatgcgcgtgcgggag tgctcttcccaactcaccctagtatcctctcatgtgtgtccttggctcagcatatcagacagtgtcc gctgtccagaccagtcagtaatggacgggccaagtgcgtgtagctgttccggcgttgcgggg gcatgttgggtgaacgggtggggtgaagggtgggtgcccacgcccgtacttgggtgacacctcaa gaccacccagggaatgccaggggaggtaccaccactcaggggtgggactcagccctgggctaattgt ctacgggtggttcttcttgccttccacttcttcttctgttcaag
19	Human Parechovirus 1	tttgaagggggtctctagagagcttggcgcgtgggcttatacccgacttgcgtgagtttctctagg agagccctttccagccctgaggcggtgggtcaataaaagcctcaaacgtaactaaccttaaga agatcatgtaaacctatgcctggtctccactatcgaaggcaacttgcataaagaagagtggggtc aagacgcttaagcatagagacagtttcttcttaacccacattgtgtgggggtggcagatggcgctg ccataactctaatagtgagataccacgctgtggaccttatgtctcacagccatcctctagtaagt gtgagacgtctggtgacgtgtgggaacttattggaacaacatttgcgcaagcatcctactgcc agcgggaaaaacacctggtaacagggtgcctctggggccaaaagccaagggttaacagacccttag gattgggttcaaacctgagatgttgggaagataatttagtacctgctgctgtagttatgcaaacact agtgtgaaggcccatggaaggatgccagaaggtaaccgtaggtgaacaagtgcacactatggatctg atttggggccagatacctctatcttgggtatctgggttaaaaaacatctaattgggccaaaacccggggg ggatccccgggttctcttattctatcaatgccact
20	Crohivirus B	gtataagagacaggtgtttgcttctcggactggcatcttgggaccaaccccccttttccagc catgggttaaatggcaataaaggacgtaacaacttgaaccattgaagcttgaatttgaaccact aagcttctgtgcacataatgaaccatcaagctgttagtcccagcaggaggttgcattgctgtgac gaaatggggctcgacccccatagtaggatactgtatttgcattccattgtggacctgcaactcta cacatagaggcttgtcttgcattcaaacctgagtagtgtagcttagaccctatagtagcggga ggaccgttgttctcctaataaccctacataataggctaggtgggcatgcccaatttgaagatcca gactgggggtcgggtcgggcagggttagatccctgttagctactgcctgatagggtgggtctcaac catgtgtagtttaaattgagctgttcataacc
21	Yc-3	actgaagatcctacagtaactactgccccaatgaacgccacagatgggtctgctgactactat cttagtgcctagttgaggttgaagtgcgggtttttagaagaaccagtttctgaacattatcatccc agcatctattctatacgcacaagatagatagtcacagcagacacatctgtgctactgcttgatagag ttgcgctggtcaacttagattggtataaccagttgagtggtgcaa
22	Rosavirus M-7	tatgcatacctggacggccctaacctcggctgtggccttctgcgatttcagcgctaccaggcttctg gtctcgcaggcggttgaattagtaggtgcactgtctaaagtgaagacagcagtgctctctgtgaaaagt tgatgacactctcagggtttagcgtatcactcaaggctagcggatttccccgtgtggtaacacacg cctctaggcccagaaggcacgggtgtgacagcaccctttagtggctgtgttccccaccagcac ctgatttgggtattctctagtaacggacaagcatggctgctcttaagcattcagtgctgcgggggc tgaaggatgccagaaggtaaccgcaggtaacgataagctcactgtggatctgactgtggggctgc gggctgggtgtctttccaccagccaaaaccgttaaacggtagtcgagttaaaaaacgtctag gccccacccccagggtatgggggttcccttaaacctcacaagttaac
23	Shanbavirus A	tgaaaaggggcgccaggggtgtgtgtgtgttactaaataccaccatcgccctgaccttccctttcc

		cctgtggctcagggtcacttagccccctcttgggtaccagtagtttctaccctgggcacagggt taactatgcaagacgggaacaacaatctcttagtccccctcgcgtagtgggtcgacccccatgt gtaggagtggataaaggacggagtgagccgatacggggaagagtgtgcggtcacaccttaattc catgagcgtcggaagaaggagctgtgaacaatggcgacctgaacctgacacatggagctcca caggcatggctactgtagactacgcagcctgggtggagtggtataccctgggtgagccgcca gtgaatgggagttcactggtaacacacactgcctgtaggggtcaggggcctcctgccccgcgta atgaggtagaccatatgcc
24	Pasivirus A	gcggctggataftctggccgtgcaactgctttgaccagtggtctgggtaacttagccaagtgtc cttcccccttccattatattgtttatggctttgtctgtgttttagttatataagatccttccgcc gatataagacctcgacagtctagtgtaggaggtgtgtatataattgtccccagaagagtgcacct gacacatagaaacatgagtagatgtgtatccgtggaggatgccgggactggattccataccc attgccatcccaaacaggagggtataccactatgtgcagctctgcagtgagggtctgcagatt tagtcatactgcctgtaggggtgtgggctgcactctgggtactcaggctgtttatataat
25	Pasivirus A 2	gctggactttctggctgcgcaactgcttttaaccagtggtctgggtacttagccaaaaccccccttc cccgtaacctagtgtgtgtgtattattttgtgtgtgtgttaaattttatataagatccttccgccc atataagacctcgacagtctagtgtaggaggtgtgtatataatgtccccagaagagtgcacctg acacatagaaacatgagtagatgtgtatccgtggaggatgccgggactggattccataccc ttgccatcccaaacaggagggtataccgctatgtgcgctctacagtgaggatctgtagattta gtcatactgcctgtaggggtgtgggctgcactctgggtactcaggctgtttatataat
26	Echovirus E14	ttaaaacagcctgtgggtgttcccatccacaggggccactggcgccagcactctgtattgcgg taccttagtgcgctgtttatataaccgtccccaaacgtaacttagacgcatgtcaacgaagacca atagtaagcgcagcacaccagctgtgttccggtaacgacttctgtaccccgaccgagatcaa taagctactacgtggctgaaggagaaaacgttctgttaccgaccaattactcaagaacctagta acaccatgaagggtgcgaggtttctgcctcgacaccccagtgtagatcaggtcgtatgagtcac cgcatccccacgggtgacctgtggcggtgctgcgctggcggtgctccatggggaacccat gggacgcttcaactgacatgggtgcgaagagtctatgagctaattgttagtctcctggccccctga atggcgctaatcctaactgcggagcagataccacacaccagtgggcagctctgtatgacgggca actctgcagcggaaacgactactttgggtgtcctgtttcttctatccttatactgctgtctatgggtg acaattgagagattgtaccatatagtctattggaftggccatccgggtgacaaatagagcaattgtgtat ttgtttgtgtgttctgtccattaaattacaagggtctaaacaccccttaattattatagcattcaacaca acaaa
27	Human Parechovirus 5	gtacattagatgcgtcatctgcaactttagtcaataaattacctccaatgtcattaccaacattccctac cttttcaactaacactaagacaacaagtacatgcctgggtctccactattcgaaggcaacttgcaat aagaagagtggaaataagacgttaaaagcatagagctagtattcttttcaacccacaagtttgtg gggtggcagatggcgtgccaactctattagtgagataccatgctgtgtgattctatgtctacaca gccaatccttagtaagtgataaggtgtctgtgtatgtgggaactcacatgaaccaitaatttaccg taaggtaacctatagccagcgggaatcacatctgtgtgacagatgcctctggggccgaaagccaagg tttaacagacctataggtattgttcaaaacctgaattgatgtgtgtgtatagtaacctgtgtatct ggtaacagtgtaacactagtgttaaggccacgaaggatgccagaaggtaccgtaggtaaca agtgacactatggatctgctgtgggcccagctacctctatcatgtgagttgggttaaaaaacgtctag tgggccaacccaggggggtatccctgtttcttttacctaatcaaacccact
28	Aichi Virus	tttgaagagggtgggtggggcctcggccccctcaccctcttttccggtgtgtgtgttccggacc accgttactcattcagcttcttccggaacctgttcggagggaattaaacgggcaccataactcccccc accccccttttgaactaagtagtgtgtctgtgtatctgactccacggaaacggaccgatccgtgg tgaaaaaacagctagggtccacatcctccttccctggaggggcccccgccccccacatcctcc ccccagcctgacgtatcacaggctgtgtgaagccccgcgaagctgctacgtggcaattgtgg gtcccccttcatcaagacaccaggcttcttctcttaaggctagccccggcggtgaattcaggttg

		ggcaactagtggtgactgtgcgctcccaatctcgccgaggagtgctgtcccaagccaaac ccctggcccttactatgtgctgcaagcatactgagaagggttccgctgtggtgccaacctg gtgacagggtgccccagtggtcgtaaccttcttccgtctcggacggtagtgattggttaagatttgg gtaagggtcatgtgccaacgccctgtgctgggafgaaacctctactgccetagggaatgccaggcag gtacccaccctcgggtggggtcgtgagcctgggctaattgtctacgggtagtttcaattccaatcctt tatgtcggagtc
29	Hepatitis A Virus HA16	ttcaagagggtctccgagtttccggaacccctcttgaagtcacatggtgaggggacttgatac ctcaaccgctgttgcctaggtataggctaaatttcccttccctgtcttccctatttcttgtttgt ttgtaaatattaattcctgcagggtcagggtctttaatctgtttctctataagaacctcaattttcacgc tttctgtctcttcttccagggtcttcccttgccttaggtctgtgctggttgcgcccggcggtgta actccatgattagcatgagctgttaggagctaaattggggacgcagatgttgggacgtgcgcttg cagtgtaacttggcttcatgaacctcttgcattctccacaagggttaggtctacgggtgaacctctt aggctaatactcaatgaagagatgccttggatagggttaacagcgccggtatgtgtgagttgtaa gacaaaaaccattcaacgcgggaggtggtctcctcactggtgagtgaggggaattgattgt cagggtgtgtctaggttaattcagacctctgtgtgtagggcaaacatttggccttaaatgg gatcctgtgagagggggtccctcattgacagctggactgttcttggggccttatgtgtgtgttgc ctgagggtactcaggggcatttaggttttctcattcttaataata
30	Phopivirus	gggagtaaacctcaccacggttgcggtggttacggctacctaattttggatgtaaatattaactctg cagggtcaggtctctgaattatgtccacgctagtggtcactctctaccataagtgacgccttagcg gaaccttctacacttgatgtggttaggggtacattatttccctgggcttcttggcccttttccctg cactatcattcttctcgggtctcagcatgccaatgtccgacccgtgctgcccgcgggggttaa ctccatggttagcatgagctgttaggccctaaaagtgtgacactgggaactggactattgaagcat acactgttaactgaacatgtaactccaatcgatctctacaaggggtaggctacgggtgaacccc ttaggttaatactcatattgagagatactctgtaggttaagggtgctgataatgtgtgagtttaacga caaaaaccattcaacagctgtgggccaacctcaggttagatgttggagccaagtgcgtagg gggtgtgtggaatgctcagtggaagggtccctccgaaaggtcgtagggttaatcaggggca gttaggtttccacaattacaattgaa
31	CVA10	gtcttccgatctgggtgttcccaccacaggggccactgggcccagcactctgattccacgga atcttgtgctgctgttttacaaccttcccaattgtacgtagaagcaatacacactactgatcaata gtaggcatggcgccgcatgtcatgataagcacttctgttccccggactgagatcaataga ctgtctacgcgggtgaaggagaaaacgttctgtaaccggctaaactactcgagaacactagtagca ccatgggaagctgagggtgttctgctcagcacttccccgtgtagatcaggtcgtatgagtcactgca atccccacgggacggtgagtggtgctgctgttggcggcctgctatggggcaaccataggac gtctaatgttgacatggtgcgaagagctattgagctagttagtagtctccggccctgaatgcg gctaactctaactgaggagcatgcttcaaccaggaggtggtgtgtctgtaacgggttaactctg cagcgggaaccgactacttgggtgtcgtgttcttcttatccttatttggctgttattggtgacaatc acgggaattgttgccatagctattggtggttgcctcgggtgttaacagagctattgtatactatttg ttggatttactccctatcatacaaatctctgaacacttgtgtcttatactgaacttaaacacacgaaa
32	Enterovirus C	ttaaaacagctctgggtgttccaccacaggggccacgtggcgccagtacacgggtaccac ggtaacctgtgacgctgttttatactccctccccgtaactagaagcacgaacacaagttcaata gaagggggtacagaccagtaccacgaacaagcacttctgttccccgggtgaggtcacatag actgtccccacggtcaaaagtactgatccgttatccgtcacgtacttggaaagcctagtaccac cttggaaactacatgctgttgcgtcagcactcgaacccggaggtgtagcttaggtgatgagctg gacgttccccactggtgacagtgttccaggctgctgttggcgccctacgtgtgttccaaaaccaca ggacgctagtagtaacaagggtgtgaagagccactgagctacgtgagaatctccggccctg aatgcggtaatccaaccacggagcaggtaatcgcaaccagcggtgagcctgtgtgaacggtg taagctgtggcggaaccgacttgggtgtcgtgttcttcttattttatgggtgctgtctattggt

		gacaatcatagattgtatcataaagcaaattggattggccatccggagtgagctaaactatctatttc tctgagtggtggattcgtttaccacacattctgaacaatcagccctcattagtgttacccgtgtaataaga cgatatcatcacg
33	Enterovirus D	ttaaaacagctctgggggtgttccaccccagaggcccacgtggcggttagtactccgggtacccc gggtacccctgtacgacctgtttatactcccttcccaagtaactttagaagaaataaactaatgttcaac aggagggggtacaaaccagtaccaccacgaacacacacttctgtttcccggtgaagttgcatag actgtacccacgggtgaaagcgatgaatcgtttacccgcttaggtacttgcagaagcctagatcat cttggaaatcttcgatgctgttgcgacagcactctaccccgagtgtagcttgggtcgatgagcttggga caccacacacggcgacgtgggtccaggctgctgttggcgccctaccatgggttagcaccatggga cgctagttgtgaacaaggtgcgaagagcctattgagctacctgagagtcctcggccctgaatgc ggctaatacccaaccacggagcaaatgctcacaatccagtgaagtggttgcgtaatgcgcaagtct gtggcggaaccgactacttgggtgtccgtgttctttttattattatggctgcttaagtgacaatct gagattgtatcatatagctattggattagccatccgggtgatatcttgaattttgcataacttttcaca aatcctacaacattacactacatttctcttgaataattgagacaactcata
34	Enterovirus J	ttaaaatagcctcagggtgttccacccctgaggcccacgtgggttagtactctgttattacgggtac ctttgtacgacctattttatacccccttcccaagtaatttagaagcaagcacaaaccagttcagtagta agcagtacaatccagttactgtaatgaacaagtacttctgttaccgcgggaagggtctatcgggttaagct gtacccacggctgaagaatgacctaccgttaaccggctacacttctgcagaagcctagtaatgcgg ttgaagttttattgacgtttacgctcagcacactacccgctgtgtattttggctgtagtgcacggcac tccccacggcgacggctggcggtgctgctgttggcgccaaccaaggagtgcaagctccttggga cgtcataattacagacatgggtgtgaagagcctattgagctaggtggtagctcctcggccctgaatgc ggctaatacctaactccggagcatatcgggtgcgaaccagcacttgggtgttgaataacgtaagtctg gagcgggaaccgactacttgggtgtcgtgttctgttttaactttatggctgcttattggtgacaattt aacattgttaccatatagtctgttgggttggccatccggattttgtataaaacatttctcgtgcttga cctttaacacatttggtaacttcttaaatccctttattagtccttaataactaaga
35	Human Pegivirus 2	aactgttgtgttagcaatgcgcatattgctacttgggtacgcctaattggtaggcgcccggcgacc ggccccgcaagggtcctagtaggacgtgtgacaatgccatgagggtacatgacactgggtgag cggaggcgacaccgaagtcgggtgaactcgaactccagtgcgaccacctggctgtgctgttcag gagggtcatgccacgggaacgctgacgtgcgaagggtatgggtccctgcaactggtgcatgcg cggcaccactcgtacagcctgatagggtgcggtggggccccccagtgtagctgcgtggag cgcaac
36	GBV-C GT110	tgacgtgggggggtgatTTTccccccggcactgggtgcaagcccagaaaccgacgcct atctaagtagacgcaatgactcggcgccgactcggcgaccggccaaaagggtggtggtggtgga tgacagggttggtaggtcgtaaatccgggtcatcctggttagccactataggtgggtcttaagagaa ggfcaagattcctcttacgctcggcgagaccgcgcacgggtccacaggtgttggccctaccggt gtgaataagggtcccgacatcagge
37	GBV-C K1737	gacgtgggggggtgatcccccccTTTggcactgggtgcaagcccagaaaccgacgccta tttaaacagacgttaagaaccggcgccgacccggcgaccggccaaaagggtggtggtggtgat gccagggttggtaggtcgtaaatccgggtcatccttggtagccactataggtgggtcttaagggttgg ttaagggtccctctggcgctgttggcgagaaagcgacgggtccacaggtgttggccctaccggtgt gaataagggtcccgacgtcaggctcgtgttaaacggagccactaccacctgggcaacaacg cccacgtacgggtccacgtgcccttcaatgtctctcttgaccaataggcttagccggcgagttgaca aggaccagtgggggctggcggttaggggaaggaccctgcgctgcccctccgggtgagtg ggaaatgc
38	GBV-C Iowa	tgacgtgggggggtgatccGccccccggcactCgggtgcaagcccataaaaccgacgccta tctaagtagacgcaatgactcggcgccgactcggcgaccggccaaaagggtggtggtggtggtg gacagggttggtaggtcgtaaatccgggtcatcctggttagccactataggtgggtcttaagagaag

		gtcaagactcctctgtgctgcggcgagaccgcgcacgggtccacaggtgctgcccctaccgggtg tgaataaggccccgacgtcaggtcgtcgttaaacaggagcccgctcaccacacctgggcaaacgac gcccacgtacgggtccacgtgccttca
39	Pegivirus A 1220	tgtagcaatgcgcataatgctactcgggtacgcctaattggtaggcgcccgccgacggccccgc aaggccctagtaggacgtgtgacaatgccatgcgggatcatgacactggggtagcgaggca gcaccgaagtcgggtgaactcgaactccagtgccaccacctggcttggtcgttcattgaggggcat gcccacgggaacgctgacgtgcaagggtgggtccctgactggtgcatgcgcgccacca ctccgtacagcctgatagggtggcgggcgccccccagtgtagctgcgtggagcgcaac
40	Pasivirus A 3	attttctggcgtgtagctgtttgaccagtggtctggttacttagccaaatcccccttctcacc cttttaattgatggtctgtgtgtttgtttgtctgtctaataatataagatcttcccccgatata gacctcgacagctgtgtgtaggaggtgtgtgttattatgtttccccagaagagtgcacctgacac atagaaccatgagtagatgtgtatcgtggaggatcgcccggaactggtatccatataccattgcc atcccaacaagcgagggtataccactatgtgcgctgttcagtggaatctgcaatttagtcat actgctgatagggtgtggcctgactctgggtactcaggctgttcataaat
41	Sapelovirus	ccccccacccctaaagggtgtgtatccacataccccaccctcccttcaaaagtggacggacaact ggattttgactaacggcaagctcgaatggtatgtttggatacgtttaaacggcagtagcgtggcga gctatggaaaaatcgcaattgtcgtatgcatgttagtgacgcgttcggcgctcctcttgggtgatt cggcgactggttacaggagagtaggacgtgagctatgggcaaacctctacagtattacttagagg gaatgtgcaattgagactgacgagcgtctcttgagatgtggcgcatgctcttggtcattacatagt gagcttcagggttgggaaacctggactggcctatactacctgatagggtcgcggtggcgccct gtaactagtagtcagttgaaaccccc
42	Rosavirus B	gtctcttagtgtctatgcttcagagagcgggtgaactgacaccgttgcctttgacagccctcgtgc cggctttccgggttcgcagacggttggcatcatggttagttaggctaagatagtggtatgtagt gaacagttttgattgtttggagttttgtagcgtatgtagtgtgtgtggacctccccacgtggtaa cacgtgccccacaggccaaaagccaaggtgtgaaagcaccctactagtcacagactcaccat ctgggaactcctctatgaaaaatcttagtaacttttgaftcggctattcatcaacctcttagtcaagg gctgaaggatgccgggaagggtaccgcaggtaacgataagctcagtggtatctgacggggc tttgggtgcgaccgtctgtccggcgtagccagagtaaaaaacgttaggccttcaccaccaaggg attgggtttcccaatcatttgaagtact
43	Bakunsa Virus	ttttgaacgcacctcgagcgatataccggggacccccctcccttttcttctacatttctccaaa tttccctcttccctgttattttgtttgatttcttgacatgactcggacgatctatctcatttgcctt gtctgtccacacagtggtcatgtcgaaagatcatcaacactggacgtgtactgtatgtggccaaacg tgcccacaggggaaaccatgcgggtcgtgtagcggcggtggagctgtgtgacccctctccct gctcataaactttggtaggtgaagggttcaagcgacgttgcctgagggcgcatccggatgtg gggaaccaacaactaggctgaatggccgacctcagggtgagtagtagggctgtgcaccaa aagggaactcgtatcgtatccggcctgtagcctagtgagtggtgactcgtatgtgggaatctacg actggcctagtagcagggtgtagcccggttccacgcccacctgtgttagggacaccccccc
44	Tremovirus A	tttgaagaggcctcggaggtgtccggaggctcttcttgaccaacccatactggggggtgtgtg ggaccgtacctggagtgacgggtatatatgcatcccgcatggcaaggcgctgctaccttgcctt tgacgcatggtatgctcatcatttgccttggtaagccccatagaacaggcggtcagtgccga aaatcccttgcgttcacagaaccatcctaaccatgggtgtagtatgggaatcgtgtatgggatga ttaggafctctcgtagagggtatggtgtccattcaaatccaggagtagtctgctctgacattgg gacatttgatgaaccggacgtgttcagtatccgggtgtctgtattgttacggtgtatccgtctgg cacactgaaagggtatttttggtaatccttctactgcctgatagggtggcggtgccggccacga gagatgaagggtagcaatttaac
45	Swine Pasivirus 1	gcttttgaccagtggtcgtggttacttagccaagtcccttctctattttcactagtttatgtgtgtt gtctgtttgtttgttaaatgtatataagatccttcccgccgacacagacctcgacagctgtgtgta

		ggaggggttggtgataattaatttggcccaaaagagtgaccgtgatacgtggaaaccatgagtagatgtgtatccgtggagagatgcgccgggactggattccataaccattgccatcccaacaacggagggtataccaccacagtgccgcgtttgcagtgggaaictgcaaatftagtcatactgcctgatalaggggtgggcctgcactttgggggtactcaggctgttcataaat
46	PLV-CHN	acatgggggtatgtttgtctgtctgttttgttgaacaatatataagatcctttccgccgatagacctcgcacagtctagtgtaggaggattgggtgataagtaacttgcgccagaagagtgaccgtgacacatagaacctatgagtagatgtgtatccgtggaggatgcgccgggactggattccataccattgccatcccaacaacggagggtataccactatgtgcgcgtttgcagtgggagcctgcaaatftagtcatactgcctgatalaggggtgtggcctgcactctgggggtactcaggctgtttatataat
47	Pasivirus A (longer)	tgaanaagttggtgtgcagctggatttccggctgtgcaactgctttgaccagtggtctgtgggttacttagccaaattcctttcccttatccctattggtttgtgtgtgtgtgtgtttgtttgtttgttcttaactatatacagatccttccgccgatacagacctgcacagctgtgtgtaggagggtgtgtgttataatttgcaccaaaagagtgaccgtgacacgtggaaaccatgagtagatgtgtatccgtggaggatgcgccgggactggattccataccattgccatcccaacaacggagggtataccaccacagtgccgcgtttgcagtgggaaictgcaaatftagtcatactgcctgatalaggggtgtggcctgcactttgggggtactcaggctgtttatataat
48	Sicivirus	gtgtcattaaggtgtgtttggaagttcgaattagctgtttgtgtgtgattagtagacccctggagggtacccaattcggatctgaccagggaccgtgactataccgctcggtaattcgggtttaaacaatgacgtcaccacacaattacttttctcattttatcttcatctgtcttcttattaccgattacactcgatttcttggatgttctgtggagatttccctgttactctggacctcattattgtttgtttcaccagcgagctgtcccaattgcttattatttgcgtttacaacttctgtcctaataattttctgtgtgafcggggtgatgtgagctccggggtactcctgccattcaac
49	Hepacivirus K	gggaacaatgttcctgcgcgggaacgactctagecatgagtcctagtagcagtgctgtgccaccattagcacaaaaaccactgactgagccacaccctccgggaatcctgagtagacaggacattcgtcctggacgacgcagtgagcctccatgcccagaaaaattgggtataccacgggtaagggggtggccaccagcgggaatctgggggctgtcactgactatggtacagcctgatagggtgctgtccgcagcgtcagtggtatgcggctgttcatggaaac
50	Hepacivirus A	acctcctgtctaggcacgggtgcgtttgcagcgttttgcgtttgcatgcgtacacgcgtctgtccaacgaggagggaactcacatcaccatgtgtcactcccccatafgagggttcacccccgtttacacgggaatggggtaaccataccaaagtacgggtatgcgggtctcctctaggggccccccggcagggtcgaaggagctgggaattcgtgaattcgtgagtacacgaaaaatcgcgcttgaacgtctttgaccttggagccgaaatttggcggtccccacgaagggaaggcggggcggtgttggggccgccccctttatccccgggtctgataggatgtcttgcgagggcacctgcgggtctctgtagaccataggac
51	BVDV1	gtatacagagaatttgcctaggacctgtttacaatafgggcaatctaaaattataattaggcctaaggacaaaatcctcctcagcgaaggccgaaaaagaggctagccatgccccttagtaggactagcaaaataaggggggtagcaacagtggtgtgagttcgttggatggctgaagccctgagtagcagggtagtcgtcagtggttcgacgcttcggaggaacaagcctcgagataccacgtggacgagggcattgccacagcacacttaacctggacgggggtcgttcagggtgaaaacggtttaaccaaccgctacgaatacagcctgatagggtgtcgcagaggcccactgtattgtactgaaaatctctgtctgtacatggcac
52	Border Disease Virus	gtatacgggtagtagctcatccccgtatacaaaattggataccccaaactcgattgggttagggagccctcttagcgcaggccgaaccgtgttaaccatacagctagtaggactagcagacgggaggactagccatcgtgtgtgagatccctgagcagcttaaatcctgagtagcaggatagtcgtcagtagttcaacgcaggcacggttctgccttgagatgctacgtggacgagggcattgcccaagacttgccttaactctggcgggggtgtgcggagggtgaaaacacctaacgggtgttgggggttacagcctgatagggtgtctgcagaggccacgaatatggctagtataaaaatctctgtctgtacatggcac
53	BVDV2	gtatacagagattagctaaagtactgtatatgtgattggacgtcaacaaattttaattggcaacgtagggaaccttccccctcagcgaaggccgaaaaagagcctagccatgccccttagtaggactagcaaaaagt

		agggggactagcggtagcagtgagttcgttgatggccgaacccctgagtacaggggagtcgtc aatggttcgacactccattagtcgaggagtcctcgagatgcatgtggacgagggcatgccacgg cacatcttaacccatgcgggggttgcagtggtgaaagcgctaactgtggcgttatggacacagcct gataggggtgtagcagagacctgctattcgcctagtataaaactctgctgtacatggcac
54	CSFV-PK15C	gtatacggaggttagttcattctcgtatgcattatggacaaatcaaaattcaatttgggttcaggccctc ccaccagcgacggccgaactgggctagccatgcccatagtaggactagcaaacggagggacta gccgttagtgggcgagctccctgggtgttctaagtcctgagtagcagacagtcgtcagtagttcgacg tgagcagaagcccacctcgagatgctatgtggacgagggcatgcccagacgcaccttaacct agcgggggtcgttaggggtaaatcacaccacgtgatgggagtcgacctgatatgggtgctgcag aggctcactattaggctagtataaaactctgctgtacatggcac
55	SF573 Dicistrovirus	aaaaccgacccagagatcagaagtcgttgacgcgatctttattagaggacgttgcgtggcgc gagctttaattagcagacgcaaaaataaacaacaaatgctgacgcgagacttaattgtcagac gattggccaaatccgatgtgatcttgcgtctccagattgcccgaataggagtagtag
56	Hubei Picorna- like Virus	ccccaaaaccccccttaaacicaacactgtagtggaattcattttccgttgcaaaacaaacattac taccgcatttatgtaggctctgtgtttctatgcgacggttacattaatctctactctgaccactagtt ataaaaccgaagacctgaatgaacgattttcttctttcaacctctaacgaacctgacggcttga gaaacctgaagtagtaattatgtttaaaagaaaggaaagcgaacgcgatgactctacatccctat tcataaccgttgcaccacatgtgagcgtatgcgaggtcgggactgcagtattagggggaacgagct acatggagagtaattatctctccctctacgggagtcctatgtgagctgtagaagcgggtggca cctctcgttacctcgctgtacatgatcc
57	CRPV	aaaagcaaaaatgtgatcttgcctgttaaatacaattttagagggttaataaattacaagtagtgctat tgtatttaggttagctatttagctttacgttcaggatgcctagtggcagccccacaatatecagggaag ccctctctgcgggttttcagattaggtagtcgaaaaacctaagaaattacct
58	Salivirus A BN5	ttctctcttcgacgccttacggcaggcgggtcgcgggacggcttcggcctacccgcgacaag aatgccgtcatctgtcttatcaccataattcttcccttccccgcaaccatcacgcttactcgcgca cgtgtgagtggaacgtgcgtgttccaaacagttacactcacacccttggggcgggtttgtcccgcc ctcgggttctcgggaacccctccctctctctcctttctatccgcttcaatttccataactacagt gctttgtaggtgaagcatcctgaccccccggaagctgccaacgtggcaactgtggggatccag gcaggttatcaaaaggcacccggtcttccgcttcagagatccctgcgggtgaattccgacagg gctctgcttggtgccaaacctccccaaatgcgcgtgcgggagtgcttccccaaactcatcttagt aacctctcatgtgtgtgtgtgtagcatatctgaggcgacgttccgctgtccagaccagtcagc aatggacgggcccagtggtgctgtgctgttccggttccggcgcatgtttggcgaacgcgtgagg taagttgggtgcccgaatcccgaatttggtagacacctcaagaccaccagggaatgccaggga ggtagccaccactcgggtggatctgacctgggctaattgtctacgggtggttcttcttgcctcactctc tttttctggcatg
59	Salivirus A BN2	tatggcaggcgggctgttgacggcttcggccccaccacagcaagaatccatcatctgtctca cccccatgtttcccttcttccctgcaaccgttacgcttactcgcaggtgcatttgagtggtgcacgt gttgataaacagctacactcacatggggcggggtttcccgccctgcggcctctcgcgaggccc accctcccccttctccataactacagtgccttggtaggtgaagcatcctgatccccgcggaagct gtcacgtggcaactgtggggacccagacaggttatcaaaaggcacccggtcttccgcttcagg agtatccctgctagtgaattctagtagggctctgttgggtgccaacctccccaaatgcgcgtgcg ggagtgctcttccccaaactcaccttagtactctcatgtgtgtgttggtagcatatctgagacgat gttccgctgtccagaccagtcagtaattggacgggcccagtggtgtagtcttccggtttcc ggcgcatgttggtaaccggtgggtaagggttggtgtgccaacgcccgtacttgggtatacct caagaccaccagggaatgccagggaggtacccgcttcacagcgggatctgacctgggcta tgtctacgggtgttcttctgtcttccacttcttctactgttc
60	Salivirus A	tttcgaccgcttattggcaggcgggctgttgacggcttcggccccaccacagcaagaatccat

	02394	catctgtcctcacccecaatttctccctcctcctcctgcaaccattacgcttactcgcattgcatgag tggtgcacgtgttgaaacaacagctacactcacgtggggcggttttccgcccctcggcctc gagaggcccacccctcctcctccataactacagtccttgtaggtgaagcactcgtatcccc gaggagctgctcgtggcaactgtggggaccagacaggtatcaaaaggcaccggcttcc gctccaggagatccctgctagtgaattctagtggtgctcgtggtccaaacctccccaaatg cgcgctggggagtgctcttcccaactcaccttagtatcctcctcatgtgtgcttggtcagcatat ctgagacgatgtccgctgtccagaccagtcagcaatggacgggacagtgctgtagtgcctt ccggctgtccggcgcatgttggtgaaccgggtgggtgaagggtggtgtgcccacgccgtactt tggtgacaactcaagaccaccaggaatgccaggaggtaccccgctcagggcggtatctga ccctgggctaattgtctacgggtgttcttctgcttcaattcttcttctgctc
61	Salivirus A GUT	tatggcaggcggtgtgtgacgggttgcggcccaaccacagcaagaatgccatctgtctcac ccccatttcccttcttccctgcaatcaccgcttactcgcattgcatgagtggtcattgtgtt gaacaaacagctacactacatggggcggttttccgcccctacggcctcgcgaggcccac ccttccctcccttataactacagtccttggtgaggtgaagcactcgtatcccccggaagctgct cacgtggcaactgtggggaccagacaggttatcaaaaggcaccggcttccgcttcaggagc atccccactagtgaaattctagtggtgctcgtctgtgccaacctccccaaatgcgctgcggg agtgtcttcccaaccatctagtatcctcctcatgtgtgtgcttggtcagcatatctgagacgact tcgctgtccagaccagtcagtaattggacgggacagtgctgtagtcttccggttgcgg gctcatgttggtgaaccgggtgggtgaagggtggtgtgcccacgcccgtacttgggtgacactc aagaccaccaggaatgccaggaggtaccccgctcacggcggtatctgacctgggctaatt gtctacgggtgttcttctgcttccacttcttctt
62	Salivirus A CH	ttcctcgaaccattacgcttaatcgcattgagtgcatgagtggtgcatgtgtgaacaaacagctaca atcacatggggcggttttccgcccacggcttctcgcaggcccacccctccttcttccat aactacagtccttgtaggtgaagcatcccgatcctccgaggagctcctcagtggtgcaactgtgg ggaccagacaggttatcaaaaggcaccggcttctcgcctcaggagtatccctgtagcgaatt ctagtagggctcgtgtgtgccaacctctccaaatgcgctgcgggagtgcttctccccaaatc acccagtatcctcctcatgtgtgtgctgtcagcagatctgagacgagtggtccgtgtccagacca gtccagtaattggacgggacagtgctgtagtctcctcggcttgcggcgcatgttggtgaac cgggtgggtgaagggtgtgtgcccacgcccgaatcaggggatacctcaaggcaccaggaat ggcaggaggtatcccgctcacagcggtatctgacctggggtaaatgtctgcgggggtcct ctggcccattctcagtaatttcagg
63	Salivirus A SZ1	ctgtctcaccatcttcccttcttctcaccgttacgcttactcgcattgagtggtgca cgtgttgaaacaacagctacactcacatggggcggttttccgcccctcggcctcgcgag gcccaccctcctcctcctccataactacagtccttgtaggtgaagcactcgtatcccccgga agctgtcacgtggcaactgtgggaccagacaggttatcaaaaggcaccggcttccgcttcc aggagtatccctgtagtgaaattctagtagggctctgcttggtgccaacctccccaaatgcggt gaggagtgcttcccccaactcaccttagtatcctcctcatgtgtgtgcttggtcagcatatctgaga cgatgttccgctgtccagaccagtcagtaattggacgggacagtgctgtagtcttccggctt gtccggcgcatgttggtgaaccgggtgggtgaagggtgtgtgcccacgcccgtacttgggtgat acctcaagaccaccaggaatgccaggaggtaccccgctcacagcggtatctgacctggg ctaattgtctacgggtgttcttcttcttccacttcttctactgttcatg
64	Salivirus FHB	acatgggggtctgcggacggcttgcggccaccgcgacaagaatgccgtcatctgtctcatta cccgtattccttcccttcccgcaaccaccacgcttactcgcgcaggtgtgagtgccagtcgt tgtcaaacagctacaccacacccttccggggcggtttgtccgcccctcgggttctcgcggaa ccccccctcctctctcttctatccgcccctcacttccataactacagtccttgtaggtgagc acctgaccccccggaagctgctaactgtggcaactgtggggtacaggcaggttatcaagg caccgggttcttccctcaggagtatcttccgggtgaattccggtagggctgtctgtgtgcca

		acctccccaaatgcgcgtgcgggagtgctctcccaactcatcttagtaacctctcatgtgtgtg cttggtcagcatatctgagcgacgttccgtgtccagaccagtcagcaatggacgggccagt gtgcgtagtcgttccgggtttccggcgcatgtttggcgaaacgctgaggttaagggtgtgtgtccc aacgcccgtaatgttggtgataacctcaagaccaccagggaatgccaggagggtacccacttcggt gggatctgacctgggctaattgtctacgggtgttcttctgtctccacttctcttttctggcatg
65	CVB3	tfaaacagcctgtgggtgtatccacccacaggcccaattggggtgtagcactctggtatcacggt acctttgtgcgctgtttatataccccctcccccaactgtaacttagaagtaacacacaccgatcaaca gtcagcgtggcacaccagccagctttgtatcaagcacttctgttaccggactgagtatcaataga ctgctcacgcggtgaaggagaaagcgttcgttatccggccaactacttcgaaaaacctagtataca ccgtggaaagttgcagagtgtttcgtcagcactaccagtgtagatcaggtcgatgagtcaccgc attccccacggcgaccgtggcggtgctgctgttggcgccctgcccattggggaaccatggg acgctctaatacagacatggtgcgaagagtcattgagctagttgtagtctccggcccctgaatg cggctaatacctaactgcggagcacacacctcaagccagaggcgagtggtcgtaacgggcaac tctgcagcggaaaccgactacttgggtgtccgtgtttcattttattctatactgctgcttatggtag aatgagagatcgttaccatatagtattggattggccatccggtgactaatagagctattatatacc cttgttgggtttataccacttagcttgaaagagggttaaacattacaattcatgttaagtgaatacag caaa
66	CVB1	tfaaacagcctgtgggtgttccacccacaggcccaattggggtgtagcactctggtatcacggt cctttgtgcgctgtttatataccccctcccccaattgtaatttagaagttcacacaccgatcattagca agcgtggcacaccagccatgtttgtatcaagcacttctgttaccggactgagtatcaatagaccg ctaacgcggtgaaggagaaaacgttcgttaccggccaactacttcgaaaaacctagtaacacca tggaagttgcggaggtgttcgtcagcactaccagtgtagatcaggtcgatgagtcaccgcgttc cccacggggtgacgtggcggtgctgctgttggcgccctgacctaggggaaccctgtaggacg ctctaatacagacatggtgcgaagagtcattgagctagttgtaactctccggcccctgaatgcgg ctaactaactgcggagcacataacctcaaacaggggcgagtggtgtcgtaacgggcaactctg cagcggaaaccgactacttgggtgtccgtgttctattttattctatactgctgcttatggtgacaatt gacaggtgtgtacatatagtattggtggtccatccggtgactaacagagcaattatatactcttgg ttgggtttataccacttagcttgaaagagggttaaacactacatctcatcattaaactaatacaacaa a
67	Echovirus 7	tfaaacagcctgtgggtgttccacccacaggcccaattggggtgtagcactctggtatcacggt tacctttgtgcgctgtttatataccccctcccccaattgtaacttagaagaacacacaccgatcaaca gcaagcgtggcacaccagccatgtttgtatcaagcacttctgttaccggactgagtatcaataga ctgctcacgcggtgaaggagaaagcgttcgttatccggccaactacttcgagaacctagtaac accatggaaagttgcggaggtgttcgtcagcactaccagtgtagatcaggtcgatgagtcaccg cttccccacggcgaccgtggcggtgctgctgttggcgccctgacctatggggaaccatagg acgctctaatacagacatggtgcgaagagtcattgagctagctggtattctccggcccctgaatg cggctaatacctaactgtggagcacatgcccctaataccaagggtagtggtcgtaatgagcaattcc gcagcggaaaccgactacttgggtgtccgtgttctcttattctgtactgctgcttattggtgacaat tgagagattgttaccatatagtattggtggtccatccggtgactaatagagctattgtatctctt gttgattgttaccacttaatttgaagaaatcaggacactacgctacattttactattgaacaccgca aa
68	CVB5	tfaaacagcctgtgggtgttccacccacaggcccaattggggtgtagcactctggtatcacg gtacctttgtgcgctgtttatgccccctcccccaattgaaacttagaagttacacacaccgatcaa cagcggggtggtacataccagccgctgtgtatcaagcactcctgtttccccggaccgagtatcaat agactgctacgcggtgaaggagaaaacgttcgttaccggctaactacttcgagaacctagta gcatcatgaagttgcgaagcgtttcgtcagcacatcccagtgtagatcaggtcgatgagtcac cgactccccacggcgaccgtggcggtgctgctgttggcgccctgacctaggggcaaccctg

		aggacgcctcaatacagacatggcggaagagtcgattgagctagtagtctccggcccctg aatccggctaataactgcggagcacataccctcaaccaggggcgattggtcgtaacgggt aactctgcagcgggaaccgactacttgggtgctcggtgttccctttattcttataatggctgcttatgggtg acaattgaaagattgtaccatatagtattggattggccatccgggtgctaacagagctattatatac ctcttgggtgattgtaccacttgatctaaagggaagcaagacactacaattcatcatataaattgaaca cagcaaa
69	EVA71	ttaaaacagcctgtgggtgacccactcacaggggccactggcgcaagcactctggcacttgc gtaccttgggtgcgctgtttatataccctcccccaatgaaattagaagcagcaaacccgatcaata gcaggcataacgctccagttatgtcttgatcaagcacttctgttccccggactgagtafcaatagac tgctcacgcggtgaaggagaaaaacgttctgtatccggctaactacttcggaagcctagtaacac catggaagttgcggagagtttcgttcagcacttcccagtgtagatcaggtcgatgagtcaccgcat tccccacggcgaccgtggcggtggctgctgtggcggtgctccatggggtaaccatgggac gctctaatacggacatggtgtgaagagctactgagctagtagtagtctccggcccctgaaatgcg gctaateccaactgcggagcacacgcccacaagccagtggtgtagtgcgaacgggcaactct gcagcgggaaccgactacttgggtgtcggtgttccctttattcttaagtgtgctgtatggtgacaatt aaagagttgtaccatatagtattggattggccatccgggtgtgcaacagagcgatcgtttacattt attggtttgtaccattgacactgaagctctgtgatcccttaattttatcttaaccctcaacacagccaa ac
70	CVA3	ttaaaacagcctgtgggtgtaccacccacaggggccactggcgctagcacactgggtattacg gtaccttgggtgcgctgtttatacccccccccaactcgaacttagaagtaagcaaacccgatca atagcaggtgcggcgaccagtcgcatctgatcaagcacttctgtaaccccgaccgagtagtcaa tagactgctcacgcggtgaaggagaaaaacgttctgtaccgggctaactacttcgagaaacccagt agcatcatgaaagttgcagagtttctgctcagcactacccccgtgtagatcaggccgatgagica ccgcaactccccacggcgaccggtggcggtgctgctgtggcggtgctgcttatggggcaacca taggacgctctaatacggacatggtgcgaagagctattgagctagtagtagtctccggcccctg aatgcggctaatacctaactgcggagcacatacccttaatacaagggcagtggtgctgaacgggta actctgcagcgggaaccgactacttgggtgtcggtgttcccttttaattttactggctgttatggtgac aattgaggaaattgttccatatagtattggattggccatccgggtgactaacagagctattgtgtcca attgttggaatttaccggctcacactcacagtcgtaagaacccctcattacgtgttatttcaactcaa gaaa
71	CVA12	ttaaaacagcctgtgggtgtaccacccacaggggccactggcgctagcactctggtactacg gtaccttgggtgcgctgttttaagccctacccccactcgtaaacttagaaggcttctcacactcgatc aatagtaggtgtggcacgcagtcacaccgtgatcaagcacttctgttaccgggtctgagtacca ataagctgctaacgcggctgaaggggaaaaacgatcgttatccggctaactacttcgagaaccca gtaccacatgaacgttcaggggtgttctgctcgccacaacccagtgtagatcaggtcgatgagt caccgtattccccacggcgaccgtggcggtggctgctgtggcggtgctgcccattggggtagccc atgggacgctctaatactgacatggtgcgaagagctattgagctagtagtctccggcccct gaatgcggctaatacctaactgcggagcacatacccttaatacaagggcagtggtgctgaacggg caactctgcagcgggaaccgactacttgggtgtcggtgttccctttattcttacttggctgttatggt gacaattgaaaagttgtaccatatagtattggattggccatccgggtgacaaatagagctattgtata tcttttgggttacgtaccccttaattacaaagtgtttcaacttgaatacatcctaactaaattg tagaaa
72	EV24	ttaaaacagcctgtgggtgacccacccacaggggccacaggggcgctagcactctgggtatcacg gtaccttgggtgcgctgtttattacccctcccccaatgaaaattagaagcaatgcacaccgatcaac agcaggcggtggcgaccagtcacgctcgtatcaagcacttctgttccccggaccgagtagtcaata gactgctcacgcggtgaaggagaaaggttctgtatccggctaaccacttcgagaacccagtaa caccatgaaagttgcagggtgttctgctcagcacttcccagtgtagatcaggtcgatgagtcacc

		gcggtcccccacggggcgaccgtggcggtggctgctggtggcgccctgcctatgggtaacccatag gacgctctaatacagacatggtgcgaagagttaattgagctggttagtatccctccggccctgaat gcggctaatccaaactgcggagcgacgtgctccaatccagggggtgcatgctgtaacgggtaac tcgacagcggaaccgactacttgggtgtccgtgttccctttattctatactggctgcttatggtgaca atcgaggaaattgtaccatactatgctattggattggccatccgggtgctaacagagcgattatatacctc ttgttggatttatgcagctcaataccaccaactttaacacattgaaatatafcttaagttaaacacag caaa
348	AP1.0	atttcggggtacggccctggagccactccggctcctaagatttagaagtttagcacacccgcc cactaggggcccccctccaggggggcaacgggcaagcactctgtttcccggtatgctgata ggctgtaaccacggctgaaacagagattatcgctatccgcttactacttcgagaagcctagtaatg atgggtgaaattgaatccgttatccgggtgtctccccacaccagaaactcatgtaggggttgcca tcccggtacggcgacgtagcgggcatccctgcgtgcatgaggccctcttagaggagcggatg atafggactctgtcgtgaagagcctattgagctagtgtcgcactcctccgccccgtgaatcggtcta atcctaaccgccggagcagggtgggtccaatccaggccctggccgtgctgaatcgtaagctgtggga cggaaaccgactacttccgggaaggcggtgttccattgttccattgtgtgtttatgggtgacaactctg ggtaaacgttctattgcgtttattgagagattcccaacaattgaacaacgagaactacctgtttatta aatftacacagagaagaattaca
349	CK1.0	gtggccacgcgccgggcccagatacttcccttactccttccgggactgttggggagggaacacaac agggtccctctgttttccattccttcccttcccaaccccaaccgcgtatctgttggcgccgcaa gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccaggctcctcgtgctacgggtgcg ggaggtctcccccactgtgtgaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgtcgggatgcgacgggtaaccgggtccccagtggtgtgtagtcgacttccagggtc ctcctgttggcggtgttcagaaactgtctcaggtgaagtgggggtgtgccaatccctcaaaagggttg attctttaccaccttaggaatgtcctggagggtacccagcaacagctgggatctgaccggaggct aatgtctacgggtgtgttcttttcttttccacacaactctactgtgacaactcactgactatccact tgtctgtTcacG
350	PV1.0	aacaaaaggctacaccacttgggtacggcccgccaccttgtggcgcaagacatttagaaga atagcataccgccactagggccctgcagccagcagggtaacgggcaagcacttctgtcccc ggtagaacgggtataggctgtaccacggccgaaaactgaactatcgttaccgactcgtacttcg caaaagcttagtaggaactggaaagtgcgattatgaccggagtggtccccactccagaaac gcgtgatgagggtgcccacccgaccatggcgacatggtgggcatccctgcgtggcacgcgg cctctaaggagataactgcctcactgtgtaaccgaagagccccgtgagctacggtttattcctcg cctccctgaatgcggctaatectaaccatgagcagttgccatagatccataggtggactgtcgt acgcgtaagtgtggcggaaccgactacttgggatggcgtgttccctgtttctccattgtgtgtgt atgggtgacaagtatagatctcgtatctatagcgtttcttgagagatttccaacatttattcaagtgt caattctgtgtttaagcagtagtgaacc
351	SV1.0	tctgtctcaccatcttcccttcttctgcacccgttacgcttactcgcattgcatgagtggtgca cgtgcttgaaacacagctacactcacaftggggcggggtttcccgccctgcggcctctcgcgag gccacccctccctctcctccataactacagtgcttggtaggtaagcatcctgacccccggga agctgctcacgtggcaactgtgggaccagacaggttatcaaggcaccggcttctcgccttc aggagtatccctgtagtgaaattctagtagggctctgcttgggtgccaacctccccaaatgcgcgt gcgggagtgctcttccccaaactcaccctagatcctctcatgtgtgtcttggtagcatatctgaga cgatgttccgtgtcccagaccagtcagtaattggacgggcccagtgctgtagtcttccggctt gtccggcgcatgttggtaaccgggtgggtaagggtgtgtgccaacgcccgtacttgggtgat acctcaagaccaccaggaatgcaggggaggtacccgcttcacagcgggatctgacctggg ctaattgtctacgggtgggtcttcttcttccacttcttctactgttccacc

352	Caprine Kobuvirus 5Δ40	gtggccacgcccgggcccacgatacttcccttcaactccttcgggactgttggggagggaacacaaac agggtccccctgttttccattcccttcccccttttcccaaccccaaccgcccgtatctggtggcgga gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccagggtcctcctgctgacgggtgcg ggagtgctccacccaactgttgaagcctgtccaacgcgtcgtcctggcgaagactatgacgtcgc atgttccgctcgggatgcgacccgggtaaccgggtccccagtggtgtgtagtcgacgttccagggtc ctcctggttggcgttgcagaaaactgttccaggtaagtgggggtgtgccaatccctacaaagggtg atttttaccaccttaggaatgtctcggagggtacccagcaacagctgggatctgaccggagggt aattgtctacgggtgtgttcttttcttttccacacaactctactgtgacaactactgactatccact tgctcttgtgtccttctgtctgttcaagttccttgattgttttactgtcttttactgtcttttcttca caactctgtcagttcaaaagtc
353	Caprine Kobuvirus 5Δ40/3Δ122	gtggccacgcccgggcccacgatacttcccttcaactccttcgggactgttggggagggaacacaaac agggtccccctgttttccattcccttcccccttttcccaaccccaaccgcccgtatctggtggcgga gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccagggtcctcctgctgacgggtgcg ggagtgctccacccaactgttgaagcctgtccaacgcgtcgtcctggcgaagactatgacgtcgc atgttccgctcgggatgcgacccgggtaaccgggtccccagtggtgtgtagtcgacgttccagggtc ctcctggttggcgttgcagaaaactgttccaggtaagtgggggtgtgccaatccctacaaagggtg atttttaccaccttaggaatgtctcggagggtacccagcaacagctgggatctgaccggagggt aattgtctacgggtgtgttcttttcttttccacacaactaaagtc
354	Caprine Kobuvirus 5Δ40/3Δ86_Dista 1	gtggccacgcccgggcccacgatacttcccttcaactccttcgggactgttggggagggaacacaaac agggtccccctgttttccattcccttcccccttttcccaaccccaaccgcccgtatctggtggcgga gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccagggtcctcctgctgacgggtgcg ggagtgctccacccaactgttgaagcctgtccaacgcgtcgtcctggcgaagactatgacgtcgc atgttccgctcgggatgcgacccgggtaaccgggtccccagtggtgtgtagtcgacgttccagggtc ctcctggttggcgttgcagaaaactgttccaggtaagtgggggtgtgccaatccctacaaagggtg atttttaccaccttaggaatgtctcggagggtacccagcaacagctgggatctgaccggagggt aattgtctacgggtgtgttcttttcttttccacacaactctactgtgacaactactgactatccact tgctctaaagtc
355	Caprine Kobuvirus 5Δ40/3Δ 122_Kozak	gtggccacgcccgggcccacgatacttcccttcaactccttcgggactgttggggagggaacacaaac agggtccccctgttttccattcccttcccccttttcccaaccccaaccgcccgtatctggtggcgga gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccagggtcctcctgctgacgggtgcg ggagtgctccacccaactgttgaagcctgtccaacgcgtcgtcctggcgaagactatgacgtcgc atgttccgctcgggatgcgacccgggtaaccgggtccccagtggtgtgtagtcgacgttccagggtc ctcctggttggcgttgcagaaaactgttccaggtaagtgggggtgtgccaatccctacaaagggtg atttttaccaccttaggaatgtctcggagggtacccagcaacagctgggatctgaccggagggt aattgtctacgggtgtgttcttttcttttccacacaactgcccacc
356	Caprine Kobuvirus 5Δ40/3Δ86 Proximal	gtggccacgcccgggcccacgatacttcccttcaactccttcgggactgttggggagggaacacaaac agggtccccctgttttccattcccttcccccttttcccaaccccaaccgcccgtatctggtggcgga gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccagggtcctcctgctgacgggtgcg ggagtgctccacccaactgttgaagcctgtccaacgcgtcgtcctggcgaagactatgacgtcgc atgttccgctcgggatgcgacccgggtaaccgggtccccagtggtgtgtagtcgacgttccagggtc ctcctggttggcgttgcagaaaactgttccaggtaagtgggggtgtgccaatccctacaaagggtg atttttaccaccttaggaatgtctcggagggtacccagcaacagctgggatctgaccggagggt aattgtctacgggtgtgttcttttcttttccacacaacttctactgtcttttcttccacaactctgtca gttcaaaagtc
357	Parabovirus	tgaaccgttacgcaccactcagttggtgttgggtgacccaatgatggaacaaaaggctacaccac ttgggctacggcccgcgcaccttgggtgcaaaagacattagaagaatagcataccgcccactag ggccctgcagccagcagggttaaccgggcaagcaattctgtctccccggtagaacgggtataggctgt

		<p> acccacggccgaaaactgaactatcggtacccgactcgtacttcgcaaagcttagtaggaaactg gaaagttcgagttattgacccggagtggtccccactccagaaacgctgatgagggtgccacc ccgaccatggcgacatgggtggccatccctgcgctggcacgcggcctctaaggagataactcgt cctactggtaaccgaagagccccgtgagctacgggttattcctccgctccctgaatgcggctaact ctaaccatgagcagttgccatagatccataggtggactgtcgtaacgcgtaagttgtggcgga accgactacttgggatggcggtgttctgttttccattgtgtgtatggtgacaagttatagatct cgatctatagcgtttcttgagagatttccaaacatttattcaagtcgtacaattctgtgttaagcagta cagtgttaagg </p>
358	Parabovirus 5Δ48	<p> aacaaaaggctacaccacttgggctacggcccgccaccttggcgcaaagacattagaaga atagcataccgccactagggccctgcagccagcagggtaacgggcaagcacttctgtctccc ggtagaacggtataggctgtacccacggccgaaaactgaactatcggtacccgactcgtacttcg caaagcttagtaggaaactggaaagttcagttattgacccggagtggtccccactccagaaac gcgtgatgagggtgtccaccccgaccatggcgacatgggtggcctccctgcgctggcacgcgg cctctaaggagataactcgtcctactggtaaccgaagagccccgtgagctacgggttattcctccg cctccctgaatgcggctaactcaccatgagcagttgccatagatccataggtggactgtcgtta acgcgttaagttgtggcggaaccgactacttgggatggcggtgttctgttttccattgtgtgt atggtgacaagttatagatctcgatctatagcgtttcttgagagatttccaaacatttattcaagtcgta caattctgtgttaagcagtacagttaagg </p>
359	Parabovirus 5Δ67	<p> tgggctacggcccgccaccttggcgcaaagacattagaagaatagcataccgccactag ggccctgcagccagcagggtaacgggcaagcacttctgtctccccgtagaacggtataggctgt acccacggccgaaaactgaactatcggtacccgactcgtacttcgcaaagcttagtaggaaactg gaaagttcgagttattgacccggagtggtccccactccagaaacgctgatgagggtgccacc ccgaccatggcgacatgggtggccatccctgcgctggcacgcggcctctaaggagataactcgt cctactggtaaccgaagagccccgtgagctacgggttattcctccgctccctgaatgcggctaact ctaaccatgagcagttgccatagatccataggtggactgtcgtaacgcgtaagttgtggcgga accgactacttgggatggcggtgttctgttttccattgtgtgtatggtgacaagttatagatct cgatctatagcgtttcttgagagatttccaaacatttattcaagtcgtacaattctgtgttaagcagta cagtgttaagg </p>
360	Parabovirus 3Δ60	<p> tgaaccgttacgcaccactcagttgggtgttgggtgcaccaatgatggaacaaaaggctacaccac ttgggctacggcccgccaccttggcgcaaagacattagaagaatagcataccgccactag ggccctgcagccagcagggtaacgggcaagcacttctgtctccccgtagaacggtataggctgt acccacggccgaaaactgaactatcggtacccgactcgtacttcgcaaagcttagtaggaaactg gaaagttcgagttattgacccggagtggtccccactccagaaacgctgatgagggtgccacc ccgaccatggcgacatgggtggcctccctgcgctggcacgcggcctctaaggagataactcgt cctactggtaaccgaagagccccgtgagctacgggttattcctccgctccctgaatgcggctaact ctaaccatgagcagttgccatagatccataggtggactgtcgtaacgcgtaagttgtggcgga accgactacttgggatggcggtgttctgttttccattgtgtgtatggtgacaagttatagatct cgatctatagcgtttcttgagagatttccaaacatttattcaagtcgtacaattctgtgttaagcagta cagtgttaagg </p>
361	Apodemus Picornavirus	<p> ttgaaaggggtgcggatcatggcggtttctcgccatgatatccgcacattgcaaacccatattgca taccactgggtatgcattatggggaggccctttcaccctcccccccaattaccttttccccct agtaaccatacgtttactcagcgtactactccgggtacgtgatgaagaaggctacggagatt ctgggctacggccctggagccactccggctcctaaagatttagaagtttgagcacaccgcccc ctaggggcccccatccaggggggcaacgggcaagcacttctgttccccggtatgatctgatagg ctgtaaccacggctgaacacagagattatcggtatccgcttactacttcgagaagcctagtaatgatg ggtgaaattgaatccgttatccgggtgtctccccacaccagaacctatgatagggtgtgccatcc cggctacggcgacgtagcgggcatccctgcgctggcatgaggcctcttaggagacgggatgata tggatctgtcgtgaagagcctattgagctagtgctgactcctccgccccggtgaatgcggctaact </p>

		ctaaccgggagcaggtgggtccaatccaggggcctgctgtaagtcgtaagtcgggacg gaaccgactactttcgggaaggcgtgttcattgttcatttattgtgtttatggtgacaactctgg gtaaacgttctattgcgtttattgagagattcccaacaattgaacaaacgagaactacctgttttattaa atttacacagagaagaattaca
362	Apodemus Picornavirus 5Δ105	ccccccccccaattaccttttccccctctagtaaccatacgtttactcagcgttaactactccggg ttacgtgatgaagaagggtacggagattctcgggctacggccctggagccactccggctccta aagatttagaagtttgagcacaccgcccactagggcccccatcagggggggcaacgggca gcacttctgtttccccggtatgatctgataggctgtaaccacggctgaaacagagattatcgttatcc gcttactacttcgagaagcctagtaattgatgggtgaaattgaatccgttgatccgggtgtctcccca caccagaactcatgatgagggttgcacccggctacggcgacgtagcgggcatccctgcgct ggcatgaggcctcttaggaggacggatgatagtgctgtcgtgaagagcctattgagctagtgtc gactcctccgccccgtaaatcggctaactcctaaccgggagcagggtgggtccaatccaggggc ctggcctgtcgtaatcgttaagtcgtggacggaaccgactacttccgggaaggcgtgtttccattgt tcatttattgtgtttatggtagaactctgggtaaacgttctattgcgtttattgagagattcccaaca attgaacaaacgagaactacctgttttattaaatttacacagagaagaattaca
363	Apodemus Picornavirus 5Δ201	attctcgggctacggccctggagccactccggctcctaagatttagaagtttgagcacaccgcc cactagggcccccatccaggggggcaacgggcaagcacttctgtttccccggtatgatctgata ggctgtaaccacggctgaaacagagattatcgttatccgcttactacttcgagaagcctagtaattg atgggtgaaattgaatccgttgatccgggtgtctccccacaccagaactcatgatagggttcca tccggctacggcgacgtagcgggcatccctgcgctgcatgaggcctcttaggaggacggatg atatggatctgtcgtgaagagcctattgagctagtgtcgaactcctccgccccgtaaatcgggcta atcctaaccgggagcagggtgggtccaatccaggggcctgctgctgtaattgcgttaagtcgtggga cggaaaccgactacttccgggaaggcgtgtttccattgttccatttattgtgtttatggtagaactctg ggtaaacgttctattgcgtttattgagagattcccaacaattgaacaaacgagaactacctgttttatta aatttacacagagaagaattaca
364	Kobuvirus SZAL6	tttacaccctctttccgggtggtccggaccagaccacggtactcattcagctacttcgggtactgt ttcggagggaattaaacgggcacccctacccaagggttacatgggaccatattctctccctgtaac tttaagttttgtcccggtattcttgactccaggcggatgtgtgtcggccgtcctgtgaacaaacagct agacactttctccccctctctgggtgctcctggcagtcactccctccccccagcgttaacatgcc ccgctggagtgatgcacctggaagtcgtgacgtgggttagtaacttcggtgaaaaccactataa tgacaactggttgaccccaactcaaggactcagtcgttctccttaaggctagccgggccac atgaatttcagctggcaactagtgagtcacatgtcccgcaacctcggctgcggagtgctgttc ccaagcgtatgccttctctgtgaagagtgccctggcaagcacatctgagaagtcgttcgctgc gtcgtgcaacctggcgacaggtgaccagtgctgtagacttctcggattcgtccggctcttct ctaggaaacatgcgtgtaagggtcatgtgccaagccctgcgcgggtgttcttactgccttagg aatgtgcggcagggtacccctacttcggtaggagatctgagcggtagctaattgtctacgggtagttc attccatcttctctcagggtcgacatc
365	Kobuvirus SZAL6 5Δ158	ttgactccaggcggatgtgtgtgcccgtcctgtgaacaaacagctagacactttctccctccct ctgggtgtctccggcagtcactccctccccccagcgttaacatgccccgctggagtgatgcact ggaagtcgtggacgtgggttagtaacttcggtgaaaaccactataatgacaactggtgaccccc acactcaaaggactcagtcgttctccttaaggctagccggccacatgaattgcagctggcaac tagtgagtcaccatgtcccgcaacctcggctgcggagtgctgttcccaagcgtatgccttcttc tgtaaagtgctgcctggcaagcacatctgagaagtcgttcgctgcgtgccaacctggcgac agggtgaccagtgctgtagacttctcggattcgtccggctcttctctaggaacatgcgtgttaa gggttcagtgccaagccctgcgcgggtgttcttctactgcctagggaatgtgcggcagggtaccc ctacttcggtaggagctgagcggtagctaattgtctacgggtagtttccattctcttctcaggt cgacatc

366	Kobuvirus SZAL6 5Δ76	gaattaaacgggacccctaccaagggttacatgggacatattcctcctccctgtaactttaagtt ttgtgcccggtattcttgaactccaggcggatgtgtgtcgcctcctgtgaacaaacagctagacact ttctccctccctctgggctgctccggcagtcactccctcccccagcgtaacatgccccgtg gagtgatgcacctggaagtcgtggacgtgggttagtaactcgggtgaaaaccactataatgacaa ctgggtgacccccacactcaaaggactcgagctttctcccttaaggctagccccggccacatgaatt tgagctggcaactagtgagtcacacatgtcccgcaacctcggctgcggagtgctgttccccaa cgatgacctctctgttaagagtgcgctggcaagcacatctgagaagtcgttcgctgcgtgtg ccaacctggcgacaggtgacccagtggtgtagactttccggattcgtccggctctctctagga aacatgctgtgaagggtcatgtgccaaagccctgcgcgggtgttcttactgcccagggaatgtg ccgcaggtacccctactcgttagggtatctgagcggtagctaattgtctacgggtagttcattcca ttctcttcagggtcgacatc
367	Kobuvirus SZAL6 3Δ37	ttcacacctcttttccgggtgtcgggaccagaccacggctactccattcagctacttoggtaactg ttcggagggaattaaacgggacccctaccaagggttacatgggacatattcctcctccctgtaac tttaagtttgtgccgtattcttgactccaggcggatgtgtgtcgcctcctgtgaaacacagct agacactttctccctccctctgggctgctccggcagtcactccctcccccagcgtaacatgcc ccgctggagtgatgcacctggaagtcgtggacgtgggttagtaactcgggtgaaaaccactataa tgacaactgggtgacccccacactcaaaggactcgagctttctccttaaggctagccccggccac atgaattgcagctggcaactagtgagtcacacatgtcccgcaacctcggctgcggagtgctgttc cccaagcgatgacctctctgttaagagtgcgctggcaagcacatctgagaagtcgttcgctgc gtcgtgccaacctggcgacaggtgacccagtggtgtagactttccggattcgtccggctctct ctaggaaacatgctgtgaagggtcatgtgccaaagccctgcgcgggtgttcttactgcccagg aatgtgcggcaggtacccctactcgttagggtatctgagcggtagctaattggacatc
368	Salivirus SZ1	tctgtctcaccccatcttccctcttctcgtcacctgtacgttactcgcaltgcafttgagtggtgca cgtgtgtgaacaaacagctacactcacatggggcggggttttccggccctgcggcctctcgcgag ggccacccctccctctcctccataactacagtgcttggtagtgtaagcatcctgaccccccgga agctgtcacgtggcaactgtgggacccagacaggtatcaaaggcacccggcttcttcgcttc aggagtatccctgtagtaattctagtaggctctgtgtgtgtgccaacctccccaaatgcgcgt gcgggagtgctcttcccccaactcaccttagtatcctctcatgtgtgtgtgtgtagcatatctgaga cgatgttcgctgtccagaccagtcagtaattggacgggacagtggtgtagtgcgttccggctt gtccggcgcatgtttgtgaaccgggtgggtaagggtgtgtgtgccaacgcccgtactttgtgat acctcaagaccaccaggaatgcaggggaggtacccgcttcacagcgggatctgacctggg ctaattgtctacgggtgttcttctgtcttccactcttctactgttcatg
369	Crohivirus B	gtataagagacaggtgtttgcctgtcttctcgactggcatcttgggaccaaccccccttccccagc catgggttaaatggcaataaaggacgttaacactttgtaaccattaagcttgaattttgtaaccact aagcttgtgcacataatgtaaccatcaagctgttagtccagcaggaggttgcagctgttagcc gaaatgggctcgcacccccatagtaggatacttgatttgcattccattgtggacctgcaactcta cacatagaggcttgtcttgcataaacacctgagtagagtgtagcttagacctatagtagggga ggacggttgtttctcaataaacctacataataggctaggtggcatgcccaatttgcagatcca gactgggggtcgtctgtggcagggttagatccctgttagctactgcctgatagggtgtgtctaac catgtgtagtttaattgagctgttcataacc
370	Crohivirus B 5Δ51	cccccttttccagccatgggttaaatggcaataaaggacgttaacactttgtaaccattaagctt tgttaatttgaaccactaagcttgtgcacataatgtaaccatcaagctttagtccagcaggagg tttgcagctttagccgaaatgggctcgacccccatagtaggatacttgatttgcattccattgt ggacctgcaactctacatagaggcttgttcttgcataaacacctgagtagagtgtagcttag acctatagtagggaggacggttgttctcaataaacctacataataggctaggtggcatgcc caatttgcagatccagactgggggtcgtgtgtggcagggttagatccctgttagctactgctg atagggtgtgtctcaacctgtgttagtttaattgagctgttcataacc

371	CVB3	<p> ttaaaacagcctgttggtgatccaccacagggcccatfaggcgctagcactctggtatcacgg tacctttgtgcgectgtttataccccctccccaaactgtaacttagaagtaacacacaccgatcaac agtcagcgtggcacaccagccacgttttgatcaagcacttctgttaccgccgactgagtatcaatag actgctcacgcggtgaaggagaaagcgttcgttatccggccaactacttcgaaaaacctagtaac accgtgggaagttgcagagtgtttcgtcagcactacccagtgtagatcaggtcgtatgagtcaccg cattccccacggcgaccgtggcggtgctgcgttggcgcgctgccatggggaaacccatgg gacgctctaatacagacatgtgtcgaagagtctattgagctagtgtgtagtcctcggccccctgaat gcggctaatacctaactgcggagcacacaccctcaagccagagggcagtggtgtcgttaacgggcaa ctctgcagcgggaaccgactactttgggtgtccgtgttcattttatctatactggctgcttatgtgga caattgagagattgtaccatatagtctattggattggccatccgggtgaccaatagagctattatatact cttgttgggttataccacttagcttgaagagggtaaaacattacaattcattgttaagtgaatacag caaa </p>
372	CVB3 3Δ91	<p> ttaaaacagcctgttggtgatccaccacagggcccatfaggcgctagcactctggtatcacgg tacctttgtgcgectgtttataccccctccccaaactgtaacttagaagtaacacacaccgatcaac agtcagcgtggcacaccagccacgttttgatcaagcacttctgttaccgccgactgagtatcaatag actgctcacgcggtgaaggagaaagcgttcgttatccggccaactacttcgaaaaacctagtaac accgtgggaagttgcagagtgtttcgtcagcactacccagtgtagatcaggtcgtatgagtcaccg cattccccacggcgaccgtggcggtgctgcgttggcgcgctgccatggggaaacccatgg gacgctctaatacagacatgtgtcgaagagtctattgagctagtgtgtagtcctcggccccctgaat gcggctaatacctaactgcggagcacacaccctcaagccagagggcagtggtgtcgttaacgggcaa ctctgcagcgggaaccgactactttgggtgtccgtgttcattttatctatactggctgcttatgtgga caattgagagattgtaccatatagtctattggattggccatccgggtgaagcaaa </p>
373	SAFV	<p> cacttatttaattcggccttttggacaagccctcgggtgaagaacctctctcttttcgacgtgtgtgg aattgccatcatttcgacgaaagtgtctatcatgctctcccgattatgtgatgtttctgcccctgtgg gcggagcattctcgggtgagaacacttgaatcttttcttggaaacttgggtcccccggtctaagcc gcttgggaatgacaggggttatttctgtatcttattctacttttgcgggtctatccgtaaaaagggtac gtgctgccccttctctctggagaattcacacggcggtcttctcgtctcaacaagtgtgaatgca gcatgcgggaacggtgaagaaaacagtttctgtggaatttagagtgacacatgaaacagctgt agcgacctcacagtagcagcggaactccctcttggcgacaagagcctctgcggccaaaagcccc gtggataagatccactgtgtgagcgggtgcaacccagcaccctggttgatgacattctctatgg aaccagaaaatgttttctcaagccctccggtgagagaagccaagaatgtctgaaggtacccccgc gtgcgggatctgacagagaccaattggcggtgtttacactgtcactttgtttaaataattgtcac agcttctcaaaccgaagtgtcttgggtttccaatttgttga </p>
374	SAFV 5Δ46	<p> cctctctcttttcgacgtgtgtggaattgccatcatttccgacgaaagtgtctatcatgctctcccgatfa tgtgatgttttctgcccgtgtggcgagcattctcgggttgagaaacctgaatcttttcttggaaac ctgtgtcccccggtctaagccgttggaaatgacaggggttatttctgtatcttatttctacttttgcg gttctatccgtaaaaagggtacgtgtgccccttctctctgagaaattcacacggcggttcttccgt ctctcaacaagtgtgaatgcagcatgcggaaacgggtgaagaaaacagtttctgttgaaatttaga gtgcacatcgaaacagctgtagcgacctcacagtagcagcggaactccctcttggcgacaagag cctctgcggccaaaagccccgtggataagatccactgtgtgagcgggtgcaacccagcaccct ggttcgatgatcattctctatggaaccagaaaatgttttctcaagccctccggtagagaagccaag aatgtctgaagggtacccccgtgcgggatctgatcaggagaccaattggcggtgtttacactgt cactttgtttaaataattgtcacagcttctcaaaccaagtgtgttgggtttccaatttgttga </p>
375	SAFV 5Δ93	<p> gtgctatcatgctctcccgattatgtgatgttttctgcccgtgtggcgagcatttctgggttgaga aacctgaatcttttcttggaaacttgggtcccccggtctaagccgttggaaatgacaggggtattt tctgtatcttatttctacttttgcgggtctatccgtaaaaagggtacgtgtgtgccccttctctctgga gaattcacacggcggttcttccgtctctcaacaagtgtgaatgcagcatgcgggaacgggtgaaga </p>

		<p>aaacagttttctgtggaatttagagtgacatcgaaacagctgtagcgacctcacagtagcagcg gactccccctcttgcgacaagagcctctcgcgccaaaagccccgtggataagatccactgctgtg agcgggtgcaacccagcaccctgggtcgtatgacattctctatggaaccagaaaaatggtttctcaa gccctccggttagagaagccaagaatgtcctgaaggtagcccgctgctggggtatgacagggaga ccaattggcgggtgctttacactgtcactttgggttaaaaatgtcacagcttctccaaaccaagtgtgt tggtttccaattttgtga</p>
376	SAFV 3Δ47	<p>cacttatttaattcgccctttgtgacaagccctcggtgaaagaacctctctcttttcgacgtgtgtg aattgccatcatttcgacgaaagtgtatcatgctccccgattatgtgatgttttctgcctgtgtg gcggagcattctcggttgagaaccttgaatcttttcttggaaccttggttcccccggtctaagcc gcttggaatatgacagggttattttctgtatcttattctacttttgcgggttctatccgtaaaaagggtac gtgtgcccccttctctctgagaaatcacacggcggtcttccgtctcaacaagtgtgaatgca gcatgccggaacgggtgaagaaaacagtttctgtggaatttagagtgacacatgaaacagctgt agcgacctcacagtagcagcggactccctcttggcgacaagagcctctcgcgccaaaagcccc gtggataagatccactgctgtgagcgggtgcaacccagcaccctggttcgatgacattctatgg aaccagaaaatgttttctcaagccctcggtagagaagccaagaatgtctgaaggtagcccg gtgcggggtctgacagagaccaattggcggtgctttacactgtcactttggttaattgtga</p>
377	SAFV Kozak	<p>cacttatttaattcgccctttgtgacaagccctcggtgaaagaacctctctcttttcgacgtgtgtg aattgccatcatttcgacgaaagtgtatcatgctccccgattatgtgatgttttctgcctgtgtg gcggagcattctcggttgagaaccttgaatcttttcttggaaccttggttcccccggtctaagcc gcttggaatatgacagggttattttctgtatcttattctacttttgcgggttctatccgtaaaaagggtac gtgtgcccccttctctctgagaaatcacacggcggtcttccgtctcaacaagtgtgaatgca gcatgccggaacgggtgaagaaaacagtttctgtggaatttagagtgacacatgaaacagctgt agcgacctcacagtagcagcggactccctcttggcgacaagagcctctcgcgccaaaagcccc gtggataagatccactgctgtgagcgggtgcaacccagcaccctggttcgatgacattctatgg aaccagaaaatgttttctcaagccctcggtagagaagccaagaatgtctgaaggtagcccg gtgcggggtctgacagagaccaattggcggtgctttacactgtcactttggttaaaaatgtcac agcttctccaaaccaagtgtgtgtgtttccaattttgtgacggcc</p>
378	GLuc CK dCTG1	<p>gtggccacgccccggccaccgatacttcccttcaactcttctgggactgttggggagggaacacaac agggtctccccgttttccattccttcccccttttcccaacccaaccgcccgtatctggtggcgcaa gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccagggtcctctgctgacgggtgcg ggagtgtccccaccaactgttgaagcctgtccaacgcgtcgtctggaagactatgacgtgc atgttccgctgcggtacggaccgggttaaccgggttccccagtggtgtgtgtagtcgacgttcagggtc ctctggttggcgtgtccagaaactgttcaggtaagtgggtgtgtcccaatccctacaaagggtg attcttaccaccttaggaatgtctcggaggtacccagcaacagctgggactgacggagggt aattgtctacgggtgtgttcttttcttttccacaaactctacGTctgacaactcactgactatcca ctgtctctaaagtc</p>
379	GLuc CK dCTG1_2	<p>gtggccacgccccggccaccgatacttcccttcaactcttctgggactgttggggagggaacacaac agggtctccccgttttccattccttcccccttttcccaacccaaccgcccgtatctggtggcgcaa gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccagggtcctctgctgacgggtgcg ggagtgtccccaccaactgttgaagcctgtccaacgcgtcgtctggaagactatgacgtgc atgttccgctgcggtacggaccgggttaaccgggttccccagtggtgtgtgtagtcgacgttcagggtc ctctggttggcgtgtccagaaactgttcaggtaagtgggtgtgtcccaatccctacaaagggtg attcttaccaccttaggaatgtctcggaggtacccagcaacagctgggactgacggagggt aattgtctacgggtgtgttcttttcttttccacaaactctacGTGTacaactcactgactatcc actgtctctaaagtc</p>

380	GLuc CK dCTG1_2_3	gtggccacgcccgggcccacgatacttccctcactccttcgggactgttggggagggaacacaa agggctccccctgttttccattccttcccccttttcccaacccaaccccgctatctggtggcgcaa gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccaggctcctcctgctacgggtgcg ggagtgctccacccaactgttgaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttcgctcgggatgcgacccgggtaaccgggttccccagtggtgtgtagtcgacttccaggctc ctcctgggtggcgtgttccagaaaactgttcaggtaagtgggggtgtgccaatccctacaaagggtg atttttaccaccttaggaatgctcggagggtacccagcaacagctgggatctgaccggaggct aattgtctacgggtgtgtttcttttcttttccacacaactctacGTcGTacaactcacGTactat ccacttgcctctaaagtc
381	GLuc CK dAll	gtggccacgcccgggcccacgatacttccctcactccttcgggactgttggggagggaacacaa agggctccccctgttttccattccttcccccttttcccaacccaaccccgctatctggtggcgcaa gacacacgggtcttccctctaaagcacaattgtgtgtgtgtccaggctcctcctgctacgggtgcg ggagtgctccacccaactgttgaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttcgctcgggatgcgacccgggtaaccgggttccccagtggtgtgtagtcgacttccaggctc ctcctgggtggcgtgttccagaaaactgttcaggtaagtgggggtgtgccaatccctacaaagggtg atttttaccaccttaggaatgctcggagggtacccagcaacagctgggatctgaccggaggct aattgtctacgggtgtgtttcttttcttttccacacaactctacGTcGTacaactcacGTactaC TactGTctctaaagtc
382	CK SZ1-L1S	gggggtggggggggcctcggccccctcacctcttttccgggtggccacgcccgggcccacgcat acttcccttcaactccttccgggactgttggggagggaacacacagggctccccctgttttccattcctt cccccttttcccaacccaaccccgctatctggtggcgccagacacacgggtcttccctctaaa gcacaattgtgtgtgtgtccagggtcctcctcgtacgggtgcgggagtgctccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcgtgttccgctgcggatgcggacc gggtaaccgggttccccagtggtgtgtagtcgacttccagggtcctcctggttggcgtgttccagaaa ctgcttcaggtaagtgggggtgtgccaatccctacaaagggtgaaccacccagggaatgccaggga ggtaaccgcgttcacagcgggatctgaccctgggctaattgtctacgggtgttcttctgttccactt ctttactgttgcacc
383	CK Aichi Scan (AV-S)	gggggtggggggggcctcggccccctcacctcttttccgggtggccacgcccgggcccacgcat acttcccttcaactccttccgggactgttggggagggaacacacagggctccccctgttttccattcctt cccccttttcccaacccaaccccgctatctggtggcgccagacacacgggtcttccctctaaa gcacaattgtgtgtgtgtccagggtcctcctcgtacgggtgcgggagtgctccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcgtgttccgctgcggatgcggacc gggtaaccgggttccccagtggtgtgtagtcgacttccagggtcctcctggttggcgtgttccagaaa ctgcttcaggtaagtgggggtgtgccaatccctacaaagggtgatttttaccaccttaggaatgct ccggagggtacccagcaacagctgggatctgaccggaggctaattgtctacgggtgttcttctgttccatt tccaatcctttatgtcggagtc
384	CK Aichi Loop (AV-L1)	gggggtggggggggcctcggccccctcacctcttttccgggtggccacgcccgggcccacgcat acttcccttcaactccttccgggactgttggggagggaacacacagggctccccctgttttccattcctt cccccttttcccaacccaaccccgctatctggtggcgccagacacacgggtcttccctctaaa gcacaattgtgtgtgtgtccagggtcctcctcgtacgggtgcgggagtgctccacccaactgttgt aagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcgtgttccgctgcggatgcggacc gggtaaccgggttccccagtggtgtgtagtcgacttccagggtcctcctggttggcgtgttccagaaa ctgcttcaggtaagtgggggtgtgccaatccctacaaagggtgaactgccttaggaatgccaggca ggtaaccacctcgggtgggatctgagcctgggctaattgtctacgggtgttcttcttcttca cacaactctactgtgacaactcactgactatccactgtctcttctgtgcttctgttcttcaagtt ccttgattgttttactgtcttctactgtcttcttctcaaatccttgcctcagttcaaagtc

385	CK SZ1-L2	<p>gggggtggggggggcctcgccccctaccctctttccgggtggccacgcccgggcccacgat acttcccttcaactccttcgggactgttggggagggaacacaacagggtcctccctgtttccattcctt ccccctttcccaacccaaccccggtatctgttggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtccagggtcctctcgtacggtgcgggagtgctccaccaactgttgt aagcctgtccaacgcgtgatcccccggaagctgctcacgtggcaactgtggggacccagaca ggttatcaaaggcaccgggtcttccgcttcaggagatccctgctagtgaattctagtagggctct gcttgctgtgtccagaaactgtctcaggtaagtgggggtgcccattccctacaaaaggttgattctt caccaccttaggaatgtctccggagggtacccagcaacagctgggatctgaccggaggctaattgt ctacgggtgtgtttctttttttcacacaactctactctgacaactcactgactatccactgtctct ctgtgctttctgtctgttcaagttccttgattgttttgactgcttttactgcttttctctcacaatcc ttgctcagttcaaagtc</p>
386	CK Aichi TriLoop (AV-L2)	<p>gggggtggggggggcctcgccccctaccctctttccgggtggccacgcccgggcccacgat acttcccttcaactccttcgggactgttggggagggaacacaacagggtcctccctgtttccattcctt ccccctttcccaacccaaccccggtatctgttggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtccagggtcctctcgtacggtgcgggagtgctccaccaactgttgt aagcctgtccaacgcgtgtcctggcaagcatatctgagaagggtgtccgctgtgtgctgccaact ggtgacagggtgccccagtggtcgtaaccttcttcgtctcggacggtgcgtgtccagaaactgct tcaggtaagtgggggtgtgccaatccctacaaaaggttgattctttaccaccttaggaatgtctccg agggtacccagcaacagctgggatctgaccggaggctaattgtctacgggtgtgtttcttttttt tcacacaactctactgtgacaactcactgactatccactgctctctgtgctttctgtctgtgttcaa gtttcttgattgttttgactgcttttactgcttttctctcacaatcctgtcagttcaaagtc</p>
387	CK Scan Deletion (ΔS)	<p>gggggtggggggggcctcgccccctaccctctttccgggtggccacgcccgggcccacgat acttcccttcaactccttcgggactgttggggagggaacacaacagggtcctccctgtttccattcctt ccccctttcccaacccaaccccggtatctgttggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtccagggtcctctcgtacggtgcgggagtgctccaccaactgttgt aagcctgtccaacgcgtgtcctggcaagactatgacgtgcgtgtccgtgcggatgccgacc gggtaacgggttcccagtggtgtagtgcgacttccagggtcctcgtgtgtggtgttcagaaa ctgttcaggtaagtgggggtgtgccaatccctacaaaaggttgattctttaccaccttaggaatgt ccggagggtacccagcaacagctgggatctgaccggaggctaattgtctacgggtgtgtg</p>
388	CK Loop Deletion ($\Delta L1$)	<p>gggggtggggggggcctcgccccctaccctctttccgggtggccacgcccgggcccacgat acttcccttcaactccttcgggactgttggggagggaacacaacagggtcctccctgtttccattcctt ccccctttcccaacccaaccccggtatctgttggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtccagggtcctctcgtacggtgcgggagtgctccaccaactgttgt aagcctgtccaacgcgtgtcctggcaagactatgacgtgcgtgtccgtgcggatgccgacc gggtaacgggttcccagtggtgtagtgcgacttccagggtcctcgtgtgtggtgttcagaaa ctgttcaggtaagtgggggtgtgccaatccctacaaaaggttgattcttttcttttcacacaactct actgctgacaactcactgactatccactgtctctgtgtgctttctgtctgtgtcaagttcgtgattgt ttttgactgcttttactgctttttcttcacaatcctgtcagttcaaagtc</p>
389	CK Triloop Deletion ($\Delta L2$)	<p>gggggtggggggggcctcgccccctaccctctttccgggtggccacgcccgggcccacgat acttcccttcaactccttcgggactgttggggagggaacacaacagggtcctccctgtttccattcctt ccccctttcccaacccaaccccggtatctgttggtggcggcaagacacacgggtctttccctctaaa gcacaattgtgtgtgtgtccagggtcctctcgtacggtgcgggagtgctccaccaactgttgt aagcctgtccaacgcgtgttccagaaactgttcaggtaagtgggggtgtcccattccctacaaa agggttgattctttaccaccttaggaatgtctccggagggtacccagcaacagctgggatctgaccg gaggctaattgtctacgggtgtgtttcttttttttcacacaactctactgtgacaactcactgact atccactgtctctgtgtgctttctgtctgtgttcaagttccttgattgttttgactgcttttactgctttt cttctcacaatcctgtcagttcaaagtc</p>

413	RhPV	gataaaagaacctataatccctcgacacccgctcacaccgctctatctgctcattagggaatt acggctcctttttgtggatacaatctctgtatacagataacttattgttaattcattgacctttacgcaa tcctgcgtaaaatgctggataggggtgacttcggatttcgagcctatattggtttgaaggaccttta agtcctactatactacattgtactagcgttaggccagtaggcccgtaagatattataactattttatta tattttatcacccccacattaatccagttaaagctttataactataagtaagccgtgcccgaacgtt aatcggctcgttagtgctgaacaactgttagtttaattttccaaaattttttcacaatttttagtaaga ttttagcttgccftaagcagctttatatctctgtatatttttaaagtttataggagcaaggtcgttta ctcgcaatagctattttatttttaggaataattatcaectcgtaatttttaattataactagctttatct atttata
414	Halastavi arva (1x mut)	ctgattctaaccttgcctgtatgggtccctaacgggttcatttaacatgcatgaggggtgctataacc gcatcattctaaagcgattcaatgcttcatttaggaattttgtgacgattaaaaggtaaccccaaa aaacaaaaccaatctactgtatttctgttttaactgaccactgcgacccaaatttctgctcttataca aagtatgtgtgtcttttgggtgtacaacctgagaactgtctacaactacatactacgagggaagaa attcgggttaagccgtgcttctcagcttagtatatactatctGgacacacctcttcatctctaatccc catctagctcctgacagagacgtgctgtatttaacaaataacccccctgttaataagagacaaagta caatcaagctaagttctctggaggtcctgttaggaacttagccattgtgatagagtcataagctatgt gcatagacagctctagctcaccatttcttccaacccatctttcatcagcttaactctatgaatccga tgcaaaaaccattctaactcttatgtgctttccaagccaaatgagagctcactctttgagccgcta tttaattggacaataaacgttttataggtgtacatcatattgttaaaaaaaa
415	Oscivirus	cctcggctccctcttccgtcgcgcgccacgacgttaaatgcgggtgtgtgtgcttaggtgccacac cactgctatttgggtccccccttccctatatafgtttgtttgttttcaatttctttaggattggcacctc ctfatgccaaatctaaatcgtggaggtacccaggcttctgtgtcttaacagaaactccacgtccaggt catagaaactggttggtaggctgctgtagtagtcatttgcctagtagtcccttgaacagggtgct ccggttactgctggtattccgggtgtaggtcgccatggtgtaacaccatcctgcatgtgtgtgaa ccagtaccgcaaggatagcaaggatgaacactgttgacgaaatggtgaagtatcaattcaccttc atggccgggaagggtcacgtggcaatcatgccaccaggtaacctcctctgggaggtacgtgaggt gggctaagcagacctgccaatgtggctgaactttccctattgttttactttgtaacattatagttgtgt tagtgatttgtgtgtgtgcccctgtgagctatatacagataagttcgcagctagaagtaactccttcg acatcggctgtattggaa
416	Cadicivirus B	caccaaccccttgacctgtaattgtagtgacagagtgctcctctgttcccggttaaccgtgttccagg acacgattgtaactctgcgcctcaccagcgtgctgacgtctgcataaggaacgtgcttccc catgtctctatcaattcttggtagtgaccgcccagtgctcctctatgggttcttctctcatgggt tctttgtgcatgcgaatgtcaccttaattggaggtctttaattagatactcttctcatcttgatag tgtcgggaatttgattcctagctctgcaaaacaacccactgatgaattcaactttcaaccgcacaa acataatcaggtttttaaattgaattgtctaaattctaaatttagttatttaagtagtttgcattctgact cgatgtaaaatgtcatacaagtcttcttcttcttcttctacattgaagttgacattagcagctgctg cacagcttfcaggttttggatcgacatcgcaactccaccacctctcttcttagtgttgatgcg gctaactctaaccgagagcaataaaacccagggttattgtctgaacgcgcaagctcttgacgggaac cgactatacacacacctctttacccttttagtaacaccttggtagc
417	PSIV (2x mut for Xba1)	gcaaaatgggtacgtagtttaaccactgcgtatcaggattgcaggccacgaagggtattgcatatct ttctatgcggtattacggcttaaaacccgttgatctgtTgtttgactgcctgtactagtggccatt ttatttaggttagagaccctgatagtaggaggttacaaactcttaaaaattgttgacccggaaa agatgggtgacccctgaagtagtggatcAagaagatctatgcgtgcatagtaatccaggtgttcc tgttttaggtgacctctgaaagtagatgacctggaaagtcacgtagtgcccaataagcacgttt gggcagcgtgctatcacaggctgacatccgaggagcccttgttttagctggctggaagcca atgatcttaagtagataagtgctgtgtgtgtgtcaacagaaagctttagtagctcttctgtgag aaagaacacatgcattctatgctctcaattctatttttttttgggcgaaggaaagctctcacgc gagtacgaatagccaacctttat
418	PSIV IGR	GCTGACTATGTGATCTTATTAATAATTAGGTTAAATTTTCGAG GTTAAAAATAGTTTAAATATTGCTATAGTCTTAGAGGTCTT GTATATTTTATACTTACCACACAAGATGGACCGGAGCAGCCC TCCAATATCTAGTGTACCCTCG
419	PV Mahoney	ATGAGTCTGGACATCCCTCACCGGTGACGGTGGTCCAGGCT GCGTTGGCGGCCTACCTATGGCTAACGCCATGGGACGCTAG

		TTGTGAACAAGGTGTGAAGAGCCTATTGAGCTACATAAGA ATCCTCCGGCCCCCTGAATGCGGCTAATCCCAACCTCGGAGC AGGTGGTCACAAACCAGTGATTGGCCTGTCGTAACGCGCA AGTCCGTGGCGGAACCGACTACTTTGGGTGTCCGTGTTTCC TTTTATTTTATTGTGGCTGCTTATGGTGACAATCACAGATTG TTATCATAAAGCGAATTGGATTGGCC
420	REV A	GGGGTCGCCGTCCTACACATTGTTGTGACGTGCGGGCCAG TTCGAATCTGTAATAAAAGCTTTTCTTCTATATCCTCAGAT TGGCAGTGAGAGGAGATTTTGTTCGTGGTGTGGCTGGCCT ACTGGGTGGGGTAGGGATCCGGACTGAATCCGTAGTATTTT GGTACAACATTTGGGGGCTCGTCCGGGATTCTCCCATCG GCAGAGGTGCCTACTGTTTCTTCGAACTCCGGCGCCGGTAA GTAAGTACTTGATTTTGGTACCTCGCGAGGGTTTGGGAGGA TCGGAGTGGCGGGACGCTGCGGGGAAGCTCCACCTCCGCT CAGCAGGGGACGCCCTGGTCTGAGCTCTGTGGTATCTGATT GTTGTTGAACCGTCTCTAAGACGGTGATACTATAAGTCGTG GTTTGTGTGTTTGTGTTACCTTGTGTTTGTTCGTCACTTGT CGACAGCGCCCTGCGAATTGGTGTACCCACACCGCGCGGCT TGCGAATAATACTTTGGAGAGTCTTTTGCCTCCAGTGTCTTC CGTTTGTACTCGTCTCTCTCCCTCTCCGGCCGGGATGGG
421	Tropivirus A	tgtcgcatgttccaacatcaaaattctgggagagtgcgaactccttaacactgccttgcctcgac ggagccgttgttatagtgtagcgagatacaaacactaaaccacttgcctcgacgggaacc ccttaccctttttttttatagtagtaaaagtgaatctttagatgaatgtcatagaaaactgcaaatgagt accacgtctaacatgagagaaatgatactggagaatccaagtttagaagtaactacgaatccagc ggaaacaagggaattctgagcttctaataaggcgtttaagactattgcaaaattctgtgtcgtaagt atatttcattgctgtagaacgtgtaaccactcggctagtagaagcattgttagcactatfatgaa actccacactatcctttctggagaagcacacaaactacatggtaagctagaccatfatcttaagcg gtgagtacactgcaacctgttaacaatgctgtatgactactttttagatcttgagcaatattgtgag gtggacatgtccaaaggtaagtgtgttgggaatggagggtccatttcccggtcacagtagtctact agtattgggtgatagccttgcggcgatcaaccatgtatttaacgtgtgactttcac
422	Symapivirus A	ttgggaatcccaatgcttcttcaacaccgctgactatcggtggtcgcttgcgtcacaacact agtcacttcccccttaactactaccaagacttctaactacccttacctattttgtctaaatttcaa acttttattctacgcgtcttataacatctttctattttagtgtagtttgtgattgtgtgtgtattt atttaattgggactagtgacggcgcccggttgggtatccgctccctttaaattgttgaagcactct tgacattataacctatcatttagttactgtttagtagatcgtatttctgaatcgaacattfatgcaattct tctcgccgagactgtctagagataaagtctcgtatatttagttagcgtgtgtatagtagagactta gatacttcacactgaggacgcttttctctatcctttgacctgattcaggccagtggtgagtaatg attgatgtgagtgccctacaatttcttaagactggtgatagcctcgcgccgctcgccattata caactgaatagcgggtgaaactctct
423	Sakobuvirus A FFUP1 (1x mut)	tcacgcgcttttccggtggtcaccaccggttagggagcgccagcgttcggttcgctaccaggt gacacactccttccccctcccatccggttcccatcctctggactggtttctcctcagattgacca gcagctgggagctgttaccagacgttggacagtaagtcccgatgactatagggctgggtgcta gtgcttggtaagcactcaacgccatcctaatgtgtacctcggttgcctcctggtcgtggtgacc ggctgtttcttctcccttggctcCagacgggctggtgtcctaccaccgcttgcagtcagacctcc ccctgcgactcgaacgccctgtcccgagcagggttagtattgtgtgtagatctgcatgtgacac cccatccactggttagagcaggaagtgtccctagctaacgcggcaagtattacttccgctacacgt ccttgagattcctcgacactctggaactagggtgactgtgggttgggaaaacccactgtgctctg tactgcctgatagggtcgcggtgaccagtggttagcaggtgttttggat
424	Rosavirus C NFSM6F	caggagatctccatgaataatctttccaccctcttagcgtctatgtaggagcgggttgag ccccgttgcaccagcgtcagagtgtgtcgtagcaggcttctgctctcggcccatgcccggccaca cctccattagtagtgaaagggtgtaagtacatgtgaaaagggttctaataatgagctgaatgtag cgattacctaaggtagcggatccccacgttggaacacgtgctctcaggccaaaagccaaagg tgttaaaagcacccttaggttagccactaccccggtgctcagttctcttagaagattcacttagta gtgtgtgcactggcaactcctaagcagagctagtgagtggtggttaaggatgcctgaaggtaacgg

		caggtaacgtaagacactgtgcatctgacaggggctcagtgctgaagctttacagaggtagct cgagttaaaaacgtctatgccccccccacgggagtggggttccccacaccaatttagatt gcact
425	Rosavirus 2 GA7403	ccaggcatggcggttaaacatgcattccctccctagtaacctcccttcgccccctccccacgttgta cccccccgagatggctgctaaggcgctgtctctacagcagtcctgttctgggtgttataagtg ctttctttccactccactccctgctatggggagcggaacggcctgtctcggctgtctctgtca gatcttccacctccaggctttctggactcgcaggggtggagtagtaggcgactgtctaaagtga aggtagcagtggttggcggaagagttgtggacctactttgagttgtagcgatcatccagagctag cggatctccccacgggtaacgcgtgctctagggccaaaggcacgggtgttcacagcacccttt ggatggcggggtgccccctccgcacttaagtagaaaaacagcttagtagtcaataacatgg ctttctcaagcattcagtgctacatgggactgaaggatgccagaaggaccgcaggcaacga taagctcactgtggtatctggtggccctggccagggtgctafacacctggttaaaccaaatct ggtagtcagggttaaaaaacgtctaagtcacccccccggggacgggggttcccttaaacctt caactgacacc
426	Rhinovirus A	cgaattccggacatctcttctgggggcgagcgtcaccgtgccccctatggaggcaactgtgcctc taatcggtgacccactgagaaaaatttcttctacgtggctaaacaatgcaactttataataacacaaa tttaatgcttaatcttaacacaaagatttgaacatatgtttgaaagtggcacacttcaaacattgcat agttgctaggggtgaagtcctttaaagggttgacagagatcttctctttatgagcggttaggagt atcttctgataattatgtgtcgtcaactcactccagatgtatgacgggtgtactaagcgattggaa ctagtcataacctcttgaattttgtattgagagcttagcagggggtatattaccgctaaagggtgac acactcgtgaggggtgctttggtgtgtgtatatttccgcccacttgcattgggggtgctaaattct aatgctgtgaataaacattttctgaatacattctctacatttggagtcataatagaggtaagccactca ggtagccctgacatgactgtgagctgagagtggtgtaattatctaatatttggcgactttctaaatct ctgttttttagtgggacaatttatggtataaaa
427	Rafivirus LPXYC222841	gtgtccgggaagcgactcaagctttgactgagctctacaccttcatccgtaacatcttaagttagt tgccatggacctctagtgcaactgccatcaccgggggtgtattggactggttttccacaatccattca tcctgagggaattttgcttttactaggtatgtgccaccacacgcttatctgtgctattgtgtcaacc atgttcttaagtagttgtgccgtgggtgagtagataaccacaacaatccgataaagcatctcgaa ggatgtgagtaattggagtgatgtgtacagagaccacaaactgaaccaagagagacacagtg aggatgttaaagggggaactcttgaaggggcatgtcccgcaattcctactgactgacacggggg gttgggtcgtggattttagcaaatcctgttactgggtgtagccttgtgcacttcaattgttctgtta taagtctctga
428	Rafivirus WHWGGF74766	tgcgaattattgcacagtccttttccccatctgtgtgtgtgatggggtaagccgcagagtaata cctactctgtgcgaacacactcactctttctatctactttatcatgtaataataagtaggggaacata ttcaattcatattgttcatctcactgaacccgcatgaaggactgcattgcatatctggacgaagtgac gtggaattttggacattatgtgattggacaccattacgctttgtgctctacggagatgttaaccataa tcttaagtagtagtaccacagcacaagaggataaagtggcatacacgacaacgggtgtgtctgc accttagtaattgtggtgttcccttggagcgtgctgaaactctgtgggttaaagacacacattagta caaatgtgggggaactcactgaaggggcatgtccgtgtactgggtgtgcccgaaggtgggggtc gctttctggagaacttagtagttctgttattgggtgtagccttgcggcggtatcaactcagattttaa tcggtgttttgcatt
429	Poecivirus BCCH-449	actacacaatcgcaacacgcgcaagttttagtttgaattggcgtgcaaatgtcaaatcaagcatata acacaatttggctgtgtgtgtgttataggaaatttgggtgtgaaattgtggatgtgtaggaaa tatgcacaattacgtcagcgctcaggagttttataacctggcgcaacacacaaaattgtcttcgctt aacatcaccagcgaggtgtaacaaaattgaagtgaattagatcgtgtatagccagggaaccatc cctcccaacgccacatctgtgggggaagtgggataatgtgggtctatatgaattgtgtctgtagac ccacagtgaagagtgaaatagtagcttgcgggtccatttgaatgtctagcatgggtggggcggg caaccccgtaggggttccccactggccaaaagcccagggttagtcaattcaaccaagggaagct ggtaacctgtgacctgaactgagtggtgagacccccttctagagtggtgtaaacggattgttaagc attttgttcttagtatctgtgtatagcagtcattttgtataggctcaaggctgtgtgttagtag atgcccgggaaggttattactgacccgggaaccgtgactatacattaggttaaacgggttataaaacc
430	Megirivirus A LY	ttcgggacactggatggcgacttgggtgggtgcccactctatcttgcactttcgttactgactttcg gatctctgactcctctgtcttcttgcgttgggtccacggacggactaattggaattttactggctaa cctcgttctgaataacctagccaatgggtgttagtaggatcctggtgtttccattaaaccttccga

		ccatagtagctagagttatgctgtgttaggtagtgggtgaagaccgcctttttgcgtatctcccacaaga accgggattatggatgtgccgtggataaggctcgaaacctcccacactgaagggtgtgtctgtaaat attgcaagcctaggtgtgttagaggcaagtagatccctgcgcacattctgtctcggcccttttgg gttagtagtgtacctacatggacgtggggtgggaatccccacctgcataacactggttagatagac ctgcggctgggtcaagttaactatgtgtataaccagttgaaatggct
431	Megirivirus E	gcttggcaacctatatactgtactctgccgaccagctgtgggtcgtgtgtggccacacaatgggattcgtt ctgttgttagagtcacatggcattactgggtgctgacgggtggggtatccgttgcacccccaaacct tacatttactggaactgttttcttggccccggaaatgattcgtcacccgcgatggagactgttgttctta ttatggcaggattacgcgtctgtgtccgcgtgaaggactaatctatgtttatagcttactacctgttct gaacgggtggcgccacccccgcctagtaggactcctggctatcgtgttagacctctagggaccacatt agctagaggtgtaggctgctatgtgatggagtagtgacccttttgggtatcactctctaaactccgg aatgtgtcatagtacgttggaaatcctactgtttttccatgagggggaggtgtgtgtgtaaaatgtgc aagccacccctcgggttaaaacagtttgggtccgcttatgccaattaccgccccctgttagttgggtgtg tttttcgctgcgggttagtagagtagccatagtgagcgggtgttgggaatcacccgcgtgtgctgc acactcgtgtgataagctgcgggtcgaagctaattgtgtgtataaccagttgatttggcat
432	Megirivirus C	ttccgcacgggtctggcacaaccggacgggttatcctgggttagatgtctgtatgtgtgtgtgggaacgtggt ggctactgtctgccacctctggtcttcttaattgggcatctagctgggtctttgcccacaaatcatctta ctctcttaccattttctattaccagactgtgtgttaactggtaaagttagctactgtgtctgtttgag actattctggtgtgtgtgtgacactctttccacaagtagattgtatgtgaggtcatgtctgtttgaacccg gaatggcacaaccgtagtaggatctgtccctcgcataactaatctgcgcctgtgtcttttagactatg ggctgtctaaggatgacattggaaacctcttttggataatccatgtcaagtcaactgtttatctgtgtgt cgtgtggaaatcctgttccgagggtctgtctgtgaggtgtgtgtgtaaaatattgcaagccacaggcagt tccttggactgtgtccgctatcagatgtacacccctctatggcaaatgtgaaccttagtgtgacgc gtgagatgggaatccacgcccgcataagactggctgataagctcgcgctgatcgagttgcaaca gtaatcagttgatttgcact
433	Ludopivirus	tagacccccactagccctttcccgctcagtggggggttactactgggcactgtttaatctgtgc ctaactagattgacacacactcccttgggaacgttaactccacgctaactactgtgcctcagccacagac acaagggtctttctgtctatcccccgggggaagataccagatggcgacccgctgtccagccgcctagt agctactcgggtgtgagtaccacacacgggttttgacgcctgctaanaatcaagagacagaggtaggg gtgttagtgtgtgtggggaaggttccacaagcgaggcaaaagcattgtctccctcgcgtgacccgggt gcaaggtaaaattgctgtgacttccgctctacccctgtactcgcctctctggagggttogaagtga cactaggatagcgtgtgttgggaacctatgcctgtcctactactgggtgtatagcctgtggcgcg gtccgtctctgtgcttatacccggtgatttgggat
434	Livupivirus	tatctacatggggatccaggctgtatggaaatgtctgtcttaacagcactataccagaagatccac ccaangtgtgtggactgggactgtgaggtgagaaatcccgaaaccagccttctcaagcgtcggga cgatcttctgttttagtgaacacctgtccctttaaatggatgacaacaccccttcagcaaatgcgaatc tgaatcccaaaagactgttagccgaactcgtgtaatactccggagaagttaggatacgcagccc ctgttgactctgtattcaggactcaaggttagcttagagctggaacttcattggaatgacaaaggaata tatgcacattgtgtcgttctcgtggtgtgagccgtcgtgaggaatgtgtgttgggaatcgacatctt agtccagtagctgtgtatagaggtgtcgtgtggcacagttacctgagaataagtaagttgtacttaaca tgaacaaaaaaataactaccacaactaccacaatactcaatactgtaattatgtcgaattcgtgtac agtaaaaacgttccgtgggaaggacaagtattgaagtgcggttacctatccgatacgcgctggatc ctca
435	Aichivirus A FSS693	caccatacacccccacccccctttctgttaactcaagtagtgtgtcgtgaatcttgactcccacggga atggatcgtatccgtggagaacaaactgctagatccacatctcctccctcccttgggagacctcgg tctcccacatctcctccagcctgacgtatcacaggctgtgtgaagccccgcgaagctgtctc acgtggcaattgtgggtcccccttcatcaagacaccaggtcttctccttaaggctagccccgat gtgtgaattcacattgggcaactagtgtgtgtcactgtgcgtcccaatctcggccgcggagtgctgt tccccagccaaacccctgtcccttactatgtgcctgtgcaagcatatctgagaaggtgtgtccgt gtggctgcagcctgttaacagggtgcccaggtgtgtgtaaccttctcgtctcgggacgggtagtgt attgttgaagattgtgtgaaggtcatgtgccaacgcccgtgtcgggatgaaaccttactgccta ggaatgccaggcaggtaccacacctcgggtgggatctgagcctggcctaattgtctacgggttag tttcaattccaattctttatctcgaatc

[illegible]

		gggtaaccgggtcccccagtggtgtagtcgacatctccagggtcctcctggtggcggtgtccagaaa ctgcttcaggtaagtggtgggtgcccacccctacaaagggtgattcttcaccaccttaggaatgct ccggaggtaaccagcaacagctgggatctgacggaggctaattgctacgggtggtgttctt tttctttcacacaactctactgctgacaactcactgactatccactgctctctgtgcttctgctctg gttcaagttccttgattgttttgaactgttttcaactgcttttcttcacaaactctgctcagttcaaaagtc
442	Rabbit Kobuvirus	gggctataaataatgggcatctcttcccccttcccccttgaagatgagtgcgcatattctgactcgg cctggattggcccccaggcgtgaacaagcagctaggccaccatgacactgcgggtggtgctcg aaccgcgggtgcttcacgggcacctgtggtatgtaggactccaccgtggtcttcccttcccccc tcaatcttccccctggttcgactaacgggaccagtgctggaaacctgtccgggtgaacggtatagca ggccccccggcagaacaacccgggtgcttacccttaaggctagcccccttccatgaatttggtg gggcaactagtggtgtacagtggcggaacccctcgggtctaggagtgctcttggccaatcctct gtgtgtccttgcatgaggactggcaatccttcgctagtgatcggctgtgccaatgccaatcctgg cgacaggaggcccagtggtgcgaacctacgtccctctgggtgctgattgcaatcttggagta agcttggtgtgcccgaaccccagggtttacgtaccactcgtggtgtgaggaatgtgcgcagggtac cccatcctgaggtgggatctgagcggtagctaattgtctagcaccacttcttcttcttctgtgctg tcacg
443	Aalivirus	ttgaaagggggtgctcagggttagctcctgagctcttccctccaccctcttcaacgtctgcccac gatacggggccaccttcaatcttaactaatatcccttaactatttggaatttctggtttagaataattg gaacacataattgattatcttttaggattgtggataggattgttcgggatatcactccctcctgtgct aacacataattctaattccctcttcttattatctcttggagggtggtgctgaaatattgcaagccacttg agtgtatagatgaagtaggctcaagatgaatgtgtgttactcaaggcaaggtagctatcactaaga tattggtaacgtgaacgggattaccggtagtagcgtgacatctcctgtcttagtgcttagtactagag gacaacgacatggcatcacatatcttaacctccagttttggcatccgggacagaatgggctggat atccgcttcttcttgggtatgtgattgggtgattgggtgaaccacttgaccatgacgctcgata agagtgaccgcctgatcattgaacctctagtataaaattcagggtgaaatc
444	Grusopivirus A	tgcctgagtaggattgtgaatttaggtatgagagggttagccaacctctgaaccataatagatac gtcaatctgaatccatctaaatctatctcttaggcagtggtgctgaaatattgcaagctactagggata gacgtgatctgattcaagaacctatctaattgtgtgatgagaaggctagggtttatccatagtaatccct tgtctgaacaggcaatgcacatgctctagtaggactcgggctctgcgattggtcttaaacggacc aatccaggtagaggcactaagtgtaggacttgcacaaatgtattacatgctgtaccgactcactag tctggaaactccactgaaagtgaactggggggggcccatcacattgtgtactgctgtagataga gttgcggctgctcaacttgattgtgataaccagttgaa
445	Grusopivirus B	gcatccgtaggtctggtgaagggttccatcaactgttggggcgctagtgtctatgaccgcatcacg gacggatgattatagatcacccaatccgggcacaaactctttagccacttctccacattactaagg gctctctgcccagtttcaacgtctagtcacgacacggaccttctacttcttcttatttctct actaaattggtatctgtactgaagatagcggattgtgatttgtgctgtctaaactaacctattcta gggttaggtgggtaccatatactaatgggtgaacaggattacctaigtatccattagtcctatggatct ggcgaccacaaactcatgttcatagagaggctaaagctgagtgctcgcgaataagcattgcttca gggtgcgactattgtctggaaccactcagtgatagctataggggggggcccgtagcatctgcct tactgcctgataagggtggcggtgctccatgaacatgcagtaaccagttgacttgac
446	Yancheng osbecks grenadier anchovy picornavirus	ccttaggggtcgggaatgcgtccttcttgggcaacttcacaaactcaaggtaattttcaacgccagega tgagcgataatccaaagaacctttatgttttagttcgtctgtatgtttataaaatataaaattgggatta gcaacccacaaaaacattttgtattctcaccacatgagagggtgggttaacctcttctgtaacctatc ttgcttgattggctacttgggttagatttaggacaacctcttttagcaagcctaaattactcctgactt gcacctggtaacaggcgtacctgaggttacgtgtgctgactaactggacttctgtaattcgtgca agtttaaatgatccattttgaatacaagaaagtatgatagtaacttagggcgtgaagtccgcttaa cataaggcagtataagtactaagataagggtgaagacctacctaataactgtgtcttttctcatggtc tttccgtgggagccctgttaggggtaaagttaagtattctcaataatttttcttcaactcttctctc tgtttt
447	Turkey Gallivirus M176	ccactcgcacttctggatagtgctgttagatagcccgacatctgcttcagggaacaaagcccc cttttcttcccaaccagcttcgcactcaagctgtaattccatgtccgggtcttcggccttagatc atggaaatgtggtcgtgctcaaatgaaattgagttgacattgatcaatgaagttgcactgaactttg ctaaactggctagcggccacctggtgtgtgcggtgtgtctcctcacatgtaacatgtgccaacggg cccgaaaggctagtggtgcaattaccgtctcaaggagggttaccccccgaacctgaacacgg

[illegible]

		aatcatggaactccactggactttcatggccttagcttctgacctagacagactctggggccccacga ccctctgggcccctcgggggtactcaggggcatfagggttttccacaftaaagagtta
453	Hepatovirus H2	gtcatgtttctttaagaacactcaattttggccataagttagactctgtgcgaacctttcatgtcagg accatgttagggccattatcctttccctggggcattctcttggccctgtttcatcttctatcatcttctt ccgggctctcacaatgccaatggagcgaccgatgcgcacgtcggggtaacccatggattagcc atgggctgttagctgctaaaagtgtgactcctgaagcatactatcaatggttagatgtaactgaaa cactgaagcttctctgatcttgaagaagggttaggctacgggtgaaacccctcagggttaatactcat attgagagatacctttggttaggttaacgttggcggataatgttaggttaacgacaataaacattcaac gcctgtggcggaacctcaccatttcattgctttgaagtgaatgtgcgttaggggtctctatcgagatg ctatgtggatgggtccctccctggaaacagggttaggggtactcagggtgactttaggttccacatt ttaaagattttc
454	Hepatovirus I	ggctgctgtgtctcaggggtaagtactggggccgcttgaccgtcgggtacgggttatgtcttttaga ttaggatgtccgtctgtccggcactctcttttgcftaaaatggccttaafccatggaggcgtaacca tggggccctttgttacctagacatgattgcatggggccgctcttggggcttagggccccagccatttc tcttgactcgttaagagtttactccttttcttactttatfttccaggctctcagcatgccgacggct ctgaccactgcgccggtgggggttaactgcattgattagcatgcagctgttaggagttaaaagtgtctg acaggccaattctgacgttaagtcactctatattactgataagtaaggttgattgatctttgtgaga gggttaggtacgggtgaaccccttaggttaatactcatattgagagatacctccagaagggtgaag gttggcggatattgtgtgagttcttttagacaaaaaccttcaacgcctgtggggccacctcactggc acaatgtttcatcccaattgtgatgggtagtttggactgaaatcaggagtaacctgccttacgagt ttaggggttaggtcaggggtatttaggtctccacattgtagaggttatgagagtgagcc
455	Hepatovirus C	ttcaaaagccccagcgggggttcatlaccctgctgtggcttttggacttccctaggatgggggaagta aattaccatcctcgtgttgcgggtgcgttaacggctacttttcttagctgttagaagtaaaatcagca tgttttatgtttgttctgtttgtttatatacatftttacactcctacaaatgcacatgaagaacagttgt agagattaaacacgcttagctgaacctaggtgtgtaactagtagtaagataagtagagggaagct ataccttaagtgtgtgggcccctgctgtttgctctataaacaacaaagtagtagagtggtgaac agtactaaatccctgagtacagggaacctcagaggtgtgatactatgtctatgtgacctgggtgg agggtggcggtccctatgatactggagtgaggatcttttgggaacccacggtttcacactgcct gatagggtcttgcggagagactcactgtttcggcgtacttgaac
456	Fipivirus A	tgcgggtaaaccccgcatgtgtgaatgaggcgtatgtccaggaaactgcgcatcgtgtttta actacgatccgtattgttactaatgcgatatcccccaattgttgcctccatgttgttcaacgcttttg gccttgaggtttatcaaggttttagcgacatagtggaagctacggctgcgtccccattttgagtg ggcagccagtttagtgccactctgtccctgaactgcgctataatgtgaattatgtcacaaaaac ggactgatgtaactgttaactgaactaaggaaatagctcactgaagtatcaagacccggtcagcg gtgtacatatatgtagtgaacactgtctgagtcactcgaataactaatgaaggatgtcagta agcatatcatgaaccatagaaatagtggggttccggggttagaggctctctgcagcaatgtatctct aacaccatggccgaaatgagagatagagaccacgatgtttgtgtgaagtaatgtgtgtgaaag aaaattctgaatgttggatgatacagctaaagggtgagtgctcaactaagagctaccaaacattt cacagcagacaacataacgtactgagagtagttggaagggttcagaaatcagt
457	Fipivirus C	cgcgggttaaaccccgcaacctcttccagccgctctgagtagcgggttagctctgatacacagt ttcctgttgggtactgtgtcttccgggtgaatgctctgtgtgaatgttttagcgtgttaagggaagcgtt tccccgtgcgtgtgaggggtttcaccgctcttccgggtgcagctctcttctgttgcattaaagatgta tggatgcactgttgtgaaggattgtgaactggggatcgacaccccgtaggggtgccccagtgctc cataggagttgtctgagaggtgtgtgtctgtagtgaactacgtgacctggcattctaaggtgtga ccccaacctgtgaggtctggtatcgagtggtgaagtgttggagggttcaatgggttctgtagt ggatattatgtcttgacgactactgggtacgaggtatttgggggtctacatgtgtga
458	Fipivirus E	ctcttcgatcttgggggttgcgcccatgtctcatttcaactagccgtgtgtctagttaaacgcacgc ctcacccgtgctgtatcgggtcgggtcttgcgacctgttagatcgtgagcgtttctgaggatcagttc gtataagttctccgggtgtgacggtaaaatcgtcacgtccatgcaatagatgacgttaaaactcg tttgcagttacataaaggaaatgtgttacttttaattgtctgttacatttaacatcttgcagtagatg ctactgtacactacgggtgtgaggaacctttagtgtgacgtatcactcatatgttgatgggtgtccca gacctttatggaagctctcagtttagtagtgccttgaacttcatgagccctgtgaacagtggaagtc aagatgtatattgtctcaacacacttccgtgctacgaagcgttttgggaagtactcggcagggttca ttctgaatcatatgtttgtacatagtcaggagtgccgtcgtctacgacggaccttttttataatt

		acaaatctgtgtctcaagtggtgtggtgctgttttcttctgttttcattgttcatatataatagctcagagt gaaagactcgggtatatacaaaactgattccaga
459	Aquamavirus	ttcaaaagggtggcgggagagtggtccctacgcgtgttagcgtgagagctggctctcctgccccttcc cctgacggcggggtatcttggctcattccctcttttctatcctccctcattggactttacggatgacccg gcataaaacttgacaacggatgttggatttcccttggctggtgagtgagacatcctccgggtgtta gttgtgtgctgtgtgctctgctgacgcagcttcaaaagtgggtgctgaaatattgcaagcgtgctgtgctc gattaacggagtggtgtacaaactctatgaacccaagtgcattcatgcaagaagccccggaggggtgag tagcatggactcgaaatcagaagagctggagctgctgctgtgtacggcacgtagcattgtcttgcctaa agaccaaggggggtatggctatagtggtggggcctatagctgtccagtgctggttgacagactgctg ctacgcctgctgttcgagtataagtagctgcaactcact
460	Avisivirus A	ttcactgcctttccccctctctataggggcggtcttttaattctttaatttctactttactatcaaat tcttctaagtagggactgaggtcacttagccctccctctcctgggtcttccagggttatagagggttcta aagctaagccatgtgtcttggagctacacttagtacaaggtttagtaatgattgtacatgcccagtaac cttctagtgtcccattgtaaaagagtggttaacactctccatggggcccgaagggtcagtgggcata gttggcatcaaggaaaggggtccccaccaccaactgaattgctggctagaagctcacttagaaga agtgctgggtgacaacgtgtccaatcgtgaacgactgaggaacgtgtgtgagatggataltgtgg gggttctactgagtagatgccctgaaggagaaatctgatcaggggcccgtgactatacgcctaggta aacggggtataaaaaccatgaagggtggcccaaaatctcttctttattttatttctatgtgtgtgaca gtcaag
461	Avisivirus B	cacccctactgcccctaaccoccaaagttagttatagggtggtcctccactcctactccacggggta agccttaagaccgggtggaatctcaagatcagccttagcgaggactattagtaccgtcaaaccccttgg cctgttagtggcccagggggtcacagagggtgacccctcctcctggggcccaaaaggctaggtggca agacaggggtccaaagtgggggtctactctaagtagcccaagctgaacatctgtctgaagccacc cttgcaggggccagggttggattggggaaactagacaccagcttgtcctgggattggggggatalcg agtttagtcaggaggtgagtagatgccccgaaggtaccccaggcacatctgggatctgatcg ggggcccggtgactatacaataggtaaacgggttaaaaaacatgaagcgcctctcttcttact tcttttattgactgtgacaaaaatagcagt
462	Crohivirus A	gttgaaagtcatttctgtctgtcccccgatgaatcctgttaaggcctcaccggccctaagggtgaaact cgggtatccccctctgtacttgcagaagatttagtacaacactatgaatctacatcttgtgtatccggga taaccccaatcccagaaactgtgtgagggcggtcaccaccctcttatggttaacataagggtgtgctc cgcggtggcacaggacccttgggtcggtgatgttttagtaattggtgtcgaaggctcattgagctaca ggagtttccctcggccctgtgtaatgcggctaatcttatccctgagcctaagggttgcgatccagcaac ttgatgttcgtaatcgttaagtggggggcggaaccgactacttccagaaggcgtgttcttgttttg tctgttactatgtgtcatgatagataattgaatatttgatcttttgagctgtttctatctattgctacatc ctttcagggtgtggatttcatatttggtaataag
463	Kunsagivirus B	gattttctggttatcccttttggacttggtagggggcccacgtgccaccaccctctgtgtgtgtgattt ctaactcgatgccttgcagtggtggcgccaccctctccttactgtgaaacctccgggtgagtgaaagtgtca gtctacagtgaccgtgcaggatgaatgcgcacatgtgaacaaactagagtgataaccgggtgc aaactctggaacggagtcggggactctgacctgtgtgggtgagctcgaggcatcacattgatgg acgcgattcgtatctctccctagtaggacctgtgtgtgtacccctgtgtgggaatccaggggtggt cgggtgcagggtgacagcctgttctccaccctcaaccattgtaggagaatcaacccct
464	Limnipivirus A	ttcttggatataccatttaacgtgtacccatacagataaattgggggtggattctggatgcctagtctcagt gattgggttaagaactcgtttactacgtatagtattagcaagtgctcgattgacacgtaataatgatct atgtgtgttaaaaccagtagtatgttatatactagtagtgtagactgtgagtacaactcttggcgta ggagagaacaattcaccgaatccgtggcgatccatggaataaagttaacctaatgtatgttacaag gcataatgagacatttatgagatattgtttatttgaactaaacgagtgtagaggtgtgtgagctatcca acttcaagccatgcattgtgtgtgtattgataatattgaccatttttgggattgtgtacacatacaatt tgaaaattaacccctcaagaataagacatgggaccattcgtgttagataccgtgtcgcggtgcttg agatttagatgggttagactagtttgggaatgagattgccgagaaggtcccgtagacatgttttaca gtcgtgttaltccgttagacttttgcagacacatggaagggtccatgtgtgtgcaattgcagggt gacagcccaactgcagagtttcttactagaataaaaaatctgtgttcaatttt
465	Limnipivirus C	gtttctgagcactgtgaagagcttagacaaacgttttaaaatttatttctctgcaacttttgtgtgtt atttttattgttaatttggccctaagcatttgttgcgaagtatttgaattcattagataattactattgttta tttagattgttaaaagtgtggaggtatcgaacccaagcctctatctctatctctctgaacaatttt

[illegible]

		gtgcgcttcattcaagtttagtgctcactgtctgaaatattgcttctagatctagtgtagcga cctgcataatgccagcgacttcccacatggtaacatgtgctctgggccccaaaggcatgtcttt gaccgtatgcagtacaacccagtataggtcttctatggcagtatggtatcagtgatgagctat acagaatatggaagtgggtcgatagtcagcccgaaggatgccagaaggatcccgagataa ccttaagagactgtggtatctgctggggcccaccacctcggtgggttagaagctaaccatgcct tgggttaaaaaacgtctaagggtgaccagaccgggggatccgggtttccctatcttgacctact ctaate
472	Ljungan Virus 87-012	ctcattgcccacacctggttgggtccaggttcatacaataacatcaataaacttttaacatcaagat agtattatcccatactagactggacgaagccgcttggaataagctagcttatctttagtgctcctg cactgaactgtttctgtctctggagtgctctacacttcagtagggcggtacccggcggtccact cttcacaggaatctgcacaggtgggtttcacctctggacagtgcattccacacccgctccacggta gaagatgatgtgtcttctgtgtaaaagcttgtaaaatcgtgtgtaggcgtagcggctacttga gtgccagcggattaccctagtggttaacactagcctctggggccaaaaggcatgtcatttgaccac tcaggtaacacacccagtgatgcacacgcttagtaateggcttagtaacaaacattgattgatcttt gaaagctgttaggaggttttaggtatgacgggctgaaggatgcccgaaggtaaccatgtaacct taagcactatggatctgacagggggccacatgtaacacatgggttagaagctcttgcgacctggg gttaaaaaacgtctagggccgccccccacagggatgtgggtttccctataacccaattatgtat a
473	Parechovirus A2	ggcgtcgggcttacaccgactgtgagtttcttaggagagtgcccttccagccagaggtg gctgtcaacaatafcaaaacgtaactaaacatctaagataacatagccctatgctgtgtccacc agttgaaggcatctgcaataaaatgggtgattgaagacgcttaaaagcatggagtaattatctttct aactagtgatcttactgggtggcagatggcgtgcataactctattagtgaggataccacgctcgtg gatcttatgccacacagccactccttagtaagtttgaagggtgctgatgaggcggtgggaacttatt ggaaataaacttctgtgcgaagcatcctactgccagcggatcaaaccttgtaacagggtgccct ggggccaaaagccacgggttaacagacccttaggattgggttaaaacctgagtaattatggaagata cttagtacctaccaacttggtaacagtgcaaacactagttgtaaggccacgaaggatgccagaa ggtaaccgcaggttaacaagagacactgtggtatctgctggggccacctacctatctcgttgag gtgtgttaaaaaacgtctagtgggccaaaccaggggggatccctgttcttatttttagtgaatg tcatt
474	Parechovirus A3	agagtccttttccagccagaggtggctgggttaataatcctactgttaacaaaacatctaagatga acaaccacacacctgtgtctccactggccgaaggcaactagcaataaggcaggtgggttcagacg cttaaaagtgtgtgtacatattctttctaacctgtgttttacacaggggtgcagatggcgtgccataac tctaacagtgagataaccacgctgtggaccttatgtcacacagccatccttagtaagtttgaagat gtctgatgacgtgtgggaacctgttgagataacagtttctgcaaaagcatcccactgcccagcga tctacatctgtaacagatgcctctggggccaaaagccaagggttaacagacccttgggattgtgt caaacctgaactgttatggaagacatttagtacctgctgatttggtagtaatgcaaacactagtgttaa ggccacgaaggatgccagaaggatcccgtaggttaacaagtggcactatggatctgctggg ggcagctacctctatctgtgtgagttgggttaaaaaacgtctagtgggccaaaccaggggggatcc tggtttcttttaatttaagtaatacact
475	Parechovirus A8	gggcttatacccgactgtgagtttcttaggagagtgcccttccagccctgaggcggtgga taataaaggcctcacatgtaacaaacatctaagacaaaataatttgccttgacactgggtcccactag ttgaaggcatctagcaataagatgagtggaacaaggacgcttaaaagtcaatgatagttatctttct aaccactatttatagtggggtgggtggtggatggcgaccataaattctaattagtgagataccacgcttgg gaccttatgtcacacagccatcctctagtaagtttgtgagacgtctgggtgacgtgtgggaacttact ggaacaatgctttgccgttaaggcttcttagcagcggaccaccacctggtaacagggtccctct ggggccaaaagccaagggttaataagaccctaattggaatgggtcaaacctggagcatgttgaaagt acttagtaactgtgtatctggtagtaattgcaaacactagttgtacggcccacgaaggatgccaga aggtaaccgtagtaacaagtgaactatggtatctgctggggccaaactacctctatctgtgtgag ttgggttaaaaaacgtctagtgggccaaaccaggggggatccctgttcttttattttactttgtcaa t
476	Parechovirus A17	ctctattagtgagataccacgctgttggaacctatgctcacacagccatccttagtaagtttgaaga cgtctgggtgacgtgtgggaactgttgggaatcaatatttgccttaaaagcatccattagccagcgat aaaacacctggtaacaggtgcctctggggccaaaagccaagggttaacagaccctagtggaattgg tttcaaacctgaaatattgtggaacacactcagtaacctactgacgtgtgtagtaatgcaagcactagt

		gtaaggccacaggaagatgccagaaaggtacctgtagggaacaagagacacatatagatctgatct ggggctggctactctctatfttgggtgagtcagttaaaaaacgtctagtgggcccaaacccagggggga ccctgggttccattatfttacaaggcact
477	Potamipivirus A	cacatggaaagcttttgccttccatgtttacgcacacactctcttgacacctgttgatgggtgtaaa ctacaacatttgcctgtctataatcgtttatfttaccctataatgtacccaagtatttgattgctgacac acataagcatcggtaacccatactgtttatgagctactacctctgctgtctacatacatttataatgaat ggtttgagctctgctcagatcaaacatggtaacatgttccttgggtcagtagaatcttattgtataat ctaagggtgctatttagtacgtagaagttgtaacacatatggggcctgatagccgctatctctgatgg atgtaaggtaacctctttaggtctgatacattctgcacaggatccaatttccggtgcccctgtacgagt gcactctatgcacgaggacgagatagctacaaccactgcaaatftaaacccaaactttaaca
478	Potamipivirus B	tttcaacgtcgtgctgacgttaaaaagccacaattccacttacctttacctttatgtttaatgttggta gttttgatctttaacaatagatcaataattgttggtaaccaatctcggatgttccgctgcatgtg agtttatttattcattttagttgtaggtggccactacgtcctggaatcacatggttaacatgtacctcg gcgtttatccactattacgtaataaataatttaaatgaaaatgtaagtgttacggctgactttgg gcctgatagttaaatgctgcactgacagatagtaacctcctttaggatcgattctgttcatatgggac catttgggtgcccactgattcaacctcttgggtgaaaaagagttagcatactacaaatttccaaacaa aaaccttttaataactacaacttatgattttatgaattttactgctcttgaanaagafatttgacattga tcgctgtactgtttcagacattcattgcatccatttttgggtactctccacaaactcaaaacttttcca cacgagaacacctgtttattgaatttgcctttatatttggaaactgttggattattgttgccttaatta ttgacctcacacctgttttaaacactacaat
479	Beihai Conger Picornavirus	gggacaaccccacagctgttacaaccattgtgggttggtctccaccttttcaacctgtggcaactt cgggttaaagttgcaaatccccctctccctattccacctcccttactttcactccccatataatgttcca gattttattctacctctttatattttatfttagtacagtggtgggaattactccagcataaactttgctgg atcagtggtcatcaagcatactaatataatgtactgagctatactattatctggcatctcacctggat aacgggtgtgaccatatttctaggttgctcctctatgtattttagcacctgtgcatctgcacgttgg ggcgacaattgtaggtttcctggcagcggttaagaaattgtgaaagctagatgcagttaatgcaa ggcgcggttttctgctacccgacactgctaaagttttgggggggctcccttaaacattctagtatt gagtgatagctttgcggcaggctaccacaaccttactataataaacctgttgaaatctcac
480	Porcine Sapelovirus JD2011	tacgcatgtattccacactcatttccccctccaccttaagggtgtgtatcccatacttaccctcc cttccacaatggagcgacaaatggatttgacctcacggcaaacacatatgggtatgatttccgataca ccttaacggcagtagcgtggcgagctatggaaaatcgcaattgtcgatagccatgttagtgacgc gcttcggcggtctcttgggtgattcgcgactggttacaggagtaggcagtgagctatgggca aacctctacagttacttagagggaatgtgcaattgagactgacgagcgtctcctcgagatgtg ggcgatgctctggcattaccatagtgagcttccagggtgggaacctggactgggctatactacc tgataggggtcgcgctgcccctgtaactagtatagtcagttgaaaccccccc
481	Porcine Sapelovirus A2	ttgaaatgggtgtgggtacatgcgtattacgggtacgcatatattccacactcatttccccctcca cccttaagggtgtgtatccccatcttaccctcccttctaaaacagatggacaaatggatttgaact tatggcaagtgaatatggtatgactttggatacactttaacggcagtagcgtggcgagctatggaaa aatcgcaattgtcgatagccatgttagtgacgcgttcggcggtctcttgggtgattcgggcactgg ttacaggagagtaggcagtgagctatgggcaaacctctacagtattacttagagggaatgtgcaatt gagactgacgagcgtctcttagagatgtggcgatgctcttggcattaccatagtgagcttccagg ttgggaacctggactgggctatactacctgatagggtcgcggtggcgctgtaactagtata gtcagttgaaaccccc
482	Simian Sapelovirus 1	ccaaggatctgtgcataggcgtgtgtatccctaaaccttttacctaccatccaatagactgtgtatt tcggttttgattgagtaatgggactgattctatacctgttaccattcaggggaaaaatggagtttctt catgcatctgacttgatgaccaaagagtcacactttgcgtgttggccgtatggaatgcttaagggt tattcttggattatgacttcagggttggccgcccaggataaaaaggcaattgtgtgaagtatgttagt cattggtgttgaacctgctaaagacgtcctaggtctacgctgtgcgggcccgaagtaagcttagg aataacaggagtagtccattttctgctttacccaacacgacctacacgaagagctagaggca ctttggggcaagggaagcctttgcttagccgaatgttcatttgagctcttgacgaatgctgcc gtctgtcccgacggtgaggcgtatggcgcatgctcatggcattacccaatgggttatctgtgaggg ggggctctctacacttagtctagctacctgacaggccggtggctgtctgtgtgtatgtgata accagtagtaatcccccatggaattgcttaacttccccctctcccttaccagacattctcaag

506

		accatgagcagttgcatagatccatagtggtgactgtcgtacgcgtaagtgtggcggaacc gactacttgggagggcgtgttccctgttttccattgtgtgtatggtgacaagttatagatcga tctatagcgtttctgagagatttccaacatttcaagtcgtacaattctgtgttaagcagtagcgt gtaagg
489	Felipivirus 127F	gatgtcggatgacggctggccacgggggaaaaacggcaaatgtgcaccacctctgcaaccac ggcgaccacgtttaaccatggtgtagtaggagtgaccactgcagtggtgctgtgtgtgcgaca gtcagtgtagtagtagacagtcctgactgggcaatgggaccgctgtgctatccctaggtgcat cgagattcctctgctacccaccagcgttgactcctatggggggggcccataggtcaggtctatac tgctgtatagggtcgcggctgtgcaccactgactgtataaccagttgtactact
490	Boosepivirus A	ttgaagacctcggcatatctgtgtcacaacgggtatgtcgtgagatcttctcccccaccctcca attccctttcccccctcttgaacttagaagttgtacacacaggggcaataggatagctgacagcc aggacacgtgagctcaagcactctgtttcccccgtcccttcaactactacgggaatgttagtaatt gtgtgcactttagtaaggtgtatccgggalttaaccccaatcccagaaactggtgatgagcgttacc acccccccggggcgaccgggaaggttctgtcgtgtggcaccagggtctcggcaccagaaaaag gtaagcaaatgaaggcgtactgtgtcagagaagtcttccctcaggccccgtaatgctggcgtaatc ctaaccagtgatccaccgggtgcaaaacctgtactaggtgtgtcgttaacgcgcaagtcgtggcg aacgcactacttgggtgtctgtgtttccatattttatttcaatttatgggtgacaagagtaagag atacagatttgcagcc
491	Boosepivirus B	tttttccccctccccccaactaccttttccccctctgtaacgctagaagtgtgtcaaacgcctgt agggtactgcaatccagcagtgcataggctaagctttctgttaccacccccacattatactagg aggattgtgaaattgtgttagtatgggttagtagcggtagcccggttaaccccaaccagaaactc acggatgagatgaacaggacccacatgttaacgtgtgtgttctgtctgtccccgcaaggtgaggcc gtgagagcttgcacgcgaaaacctgtgaaaacccaaaagtacctgtgagctcttgcctatttgtgttc ctccaggacctggaatgcggctaaacctaacccgcgacgcagtagcaaccagctagaggtgt ggctgtaatgcgcaagtgcggcggtaccgactacttgggtgttctgtgtttccttattttatttga attttatgtgtgacaacagctagaaaaaagagtgaaac
492	Phacovirus PF-CHK1	gtgtgtcatttctccccctcccccaaaccttttccccctctaateggattgattaaaccgggtaaag atgattaatgtgtgtgagttgatagtagggcccgcaattgaatccgggaattcttaagtaattgaatt gcatccaatatgaagtgagtggtgcaagctcacaagtagtacttggctctgtcccattatttgaggga ccaactcttctgtactacaatgtgtttaaagttaaactggaccacattgtgtatccagacaactccatttg ataatgtacgtcggaaacgttttcagtgcatagggtcctaaagtgtgtgtgaaattgtcaagctcaa tgggatactgaacgtgaaaacgcggctgttatcatatgggcccctagtggtgtaattgttgctt aggcatatactgttgggaatgcagtaggtgtgttagacagggtgatagcctaccggctggcgtag ttgagttgttatagccagttgattgccat
493	HRVC3 QPM	ttaaagctggatcatgtgtttccaccatgattaccacgcgggtgcagtggtctgttattacggtac atttccataccagttttatacaccaccccgaaactcatagaagttgtacacaatgaccaataggt ggtggccatccagggtcgtaatggtcaagcacttctgtttcccgccacccctgtatagcttccacc gaggcgaaaaatgaggtgtcgttatccgcaaaagtgcctacgaaaagcctagtaacactttgaaaa cccatggtgtgtcgtcagctgtttacccaacagtagacctggcagatgaggctagacattcccca ccagcgtatgtgtgttagcctgtgtgtgtgcacacctgcgggtgtgaagccagaaagt ggacaaggtgtgaagagcctattgtgtcactttagtctcgcggccccctgaatgtggttaacctta accccgtagctgttgcagttaaccaacatgtatgcagtcgtaattgggcaactatgggatgggacc aactactttgggtgtcgtgtttcctgttttacttttcttcttcttcttcttcttcttcttcttctt ttacc
494	HRVB27	ttaaacacgggatgggtatccccaacacggaccacaggggtgtagtgcctgtgtatttgtactt gcacgcctgtttcccatgtacccctcttaaafttctcccccaagtaacgttagaagttaaggaaa caaatgtacaatagggaagcatcacatccagtggtgttatgtacaagcacttctgtttcccgagcg aggatataagtggtacccaccgcgaaagccttaacgttatccgccaatcaactacgtaattggtcta gtattaccatgtttgtactgtgtgtgtcaggtgtgttccccactagtttgtcgtatgaggttag gaaactccccacgggtgaccgtgtcctagcctgtgtgtggcgccaaccacagcttttctgtggagcc ttttacagacatgtgtgaagacctgcatgtgtgtatgtgtgagtcctccggccccgaatgtgggt aaccttaacccccggagcctgtcaacataatcaatgtgtgtgaggtcgtaatgagtaattctgggatg ggaccgactacttgggtgtcgtgtttccttttattcttataatgtcttattgtcacagcatatagcat atatactgtgac

495	HRVA73	ttaaaactggggttgggtgtgtcccaaccacacccacacgcggtgtgtgacactgttattccggtaa ccttgtacgccagttttatataccttcccccttgaacttagaagacatgcgaatcgaccaatagca ggcaatcaaccagattgtacacgggtcaagcacttctgtttcccggtctctgtgtatgtctcaaca gggcaaaaacaattggagtcgttacccgcaagatgcctacgcaaaacctagtagcatcttcgaag atttttgggtgtcgtctcagttgtctacccacagcaatagacctggcagatgaggtctagaataaccca ctgggtgacagtgttctagcctgctgggtgctgctgcacacccacacgggtgtgaaagccaagattg gacaagggtgtgaagagtcacgtgtgtctcatctttagtctcctcggtccctgaatcggtctaacctta accccgtagccattgtctgcgaatccagcagagtatatggtctgtaatgagtaatacgggtatgggacc gactactttgggtgtcctgttttacttttacttataaattgtctatggtgacaatatatagatatata tgacacc
496	EV L	acatgggcccagcccccacacccactgggtgtgtagtagtctgtgtctatggaacctttctacgcctctt ttgtcttccccccccatttctccttgcattgtctcacctgtgtatcttgcaccttagaagaaataatgaac ccgcacaatagcggggcgtgagccacagcgtcaatgtgcaagcacttctgtttccccggaattggg cccataggctgtacccacggctgaaagggaacggcccggttaccgccttggtagtgcgagaatgt tagtaactcctcgtatagctttaggcgttacgctcagccctttagcccgaaagggtatgttcgggtcg atgaggctctgctattccccactggcgacagtgtgacttgcctgctgttggcgcccggggtggggg gcaacccccatccacgcctactgaaggacaggggtgtgaaggcgtattgctgctactaaggagtc tccggccccgtgaatgcgggctaaccggaaccccgagcccacgggtgtgaaacccggccacaagtgg gtcgtaatagtaatttggggcagggaccgactactttgggtgtccgtgtttctgtttttccatagat gtctgcttattgggtgacaaccataagcaattggattggccatccgggtgttcatatttgcgaat
497	EV K	tcagctgacgcaaggtgctcctcatttggagtctctcaagccctccggggcttggagggcgccgac ccccctgctagggggagccacgacacggctgggagtcattggcacccgcagccacgattca agccagaattgaaagcgggaagcacttctgtctcccgggtgtggtatcacgctgttaccacggc gaaaagtgaagcatcgttaccggactcgttactcgaagagccagtagctgttggatctgtgca gggtatagcgtcagcgtgacccctacgtagtcttgcgtgagatggctgagagaacacccacggggcg accgtgtctctcggcgctgtgtcgaaggccgggcttcagtggtcgtggtgcttgcagagtgaag cctcgaacagcctattgagctaccgttttagcctccgccccttgaatgcggctaatcctaaccatg gagcggccggccacagtcagtggtgtagagcgtcgtaacgcgcaagtcctgtgctgggaaccgac tacttttagagtggcgtgtttccaattatctttataaagtgtcttatggtgacaccacaagagatccac gatttctgttttcttactcagagacacaagtcataatcattatcatttatttgcgggaattaaacttgggtcgt ccaaacacatcage
498	EV J 1631	cacctgtaggggccacgtggcgttagtactctggtatcaagggtacctttgtacgcctattttattccct tccccacagtaacttagaagcttatctcatagttcaacagtaggggtcactaaccaagtggctcagc gaacaagcacttctgtttccccgttctagtagctgtgaagctgtacccacggcggaaggggaaa aagatcgtttatccggccccctacttccggaagcctagtaaacacattgaagcaatcgagttgtcg ctcagcacagtaacccctgtgtgactttgtgtgatgtctgggcaactccccacttggcgacagcgg ccagcgtgcgtttggcgccgaaccgactcgggcaaccgggtcggagcgtcgtttgtgacatagg tgtgaagagctctactagctagcgttaggggtagtcctccggccccgtgaattgcggaatactcaaacccg gagcacccacactcaatccagaggtcaggaatgtcgtaacgcgtgaagtcttggagcggaaccgact actttgggtgtccgtgtttctgttttacttactttggctgttattggtgaacaatcagtagtgtgttaccatat agctattggattggccatccgggtgtttgaattgtgtgtttataactaattctttacatacacagacaacc aat
499	EV J N125	cgggtacctttgtacgectattttacccttcccttgaacttagaagcaagcaaacaggttcaata gtaagcaacacaaccagtggtgtgacgaacaagtacttctgtttccccgggagggtctgacggta agctgtlaccacggcgtgaagfatgacctaccgttaaccggctacacttctcgagaagtctagtaata ccattgaagtgtttgtggcgttacgctcaacacactacccgctgtgtgtgtttggcgtgatgagtcacgg cattccccacggggcgaccgtggccgtgtgctgctgtgcggccaaccaaggggcgcaagctccttg gacgtcacttaacagacatgggtgtgaagaacctattgagctaggtagtagtctccggccccctgaat gctggctaatcctaactccggagcacatcagtgcaaccagcatttgggtgtgtgttaatacgaagtc tggagcgggaaccgactactttgggtgtccgtgtttcctgttttacttatttggctgttattggtgaca ttgtatattgttaccatagctgttggattggccatccggattttgaaagagaccaaaaacttctct ctactcagattcaagtgcgaaggttctcttcatatattacttactaatttgaagtacaaag
500	EV I	ttagtactttctacggggatagtggtatccctccctagtaatttagaagacttgaanaaccgaccaat aggcactcgcacacagcgggtgaagggtcaagcacttctgtttccccgggtcgagtagcagatg

		actgtgcccacggtgcgaaggtgaacaacccgttatccgactttgtacttcgggaagccctagacc accaaaagattatgtcttgggggttcgctcagcagcaccctgtgtgtagataaggccgafgtgacacg cattctctcacagcgactgttggcgggtgtcgcgttgccagcctgcccagtggggcaacccaicgga cgccaagcataatgacagggtgtgaagagcctactgagctacaaagtattctccggcccccgaat gcggcctaattccaaccacggagcatttgcctaccaaacagggtagtggaaatgtcgtaacgggtaac tctgtggcgggaaccgactactttgggtgtccgtgtttccttttaatttatcattctgtatatgtgtacaact atagtgctatctcgtatttgcattactattgtgtgagattaaaactttattacattgttgcattttaccctttag tgagtttccactggaacagattaattactcatcctgtttatataataaacagcagaataacttgcgaag
501	EV F1 BEV 261	gcaatgtctgcaccagtgctactgtacgtctagctaccttttcacggagtagatgggtatcccttacc gaacctagaagattgcacacaaacggaccaatagcgcacccgcatccagccggtgcagcgggtca agcacttctgtctccccggtctgttaaagatcgttatecgcggacccactacgaaaagccctagtaac tggccaagtgaacgcgaagtgtgcctccgccacaacccagtggtgtagctctgggaagtatggggct cgcaccacccccgtgtgaacacgggtgtcctgcggcggtgtgtcttccgggttccgtctcgtgcggt cacttcaacttcacgcaaccagccaagagcctattgtgtcgtggacggttttccctcggggccgtga atgtctgctaatacccaacctccgagcgtgtgcgcaaatccagtggtgtctacgtcgttaacgcttaagt tggaggcgggaacagactactttcgtgtaccgcgtgttccctctcatttttaattatattttatgtgacaat tgttgagatttgcgctcttgcacggttgcattgaatttggcttatactatttgggttgccttttacaac ctctgataataccagttcttactgtatctgtgttttctcaatttgaagtatagactacaaatagcaaa
502	EV D94	cgtgcgggccagctactctgtgtatcaccgttacctttgtacgcctgttttatatcccttccccgcaact tagaagaaaacaaatcaagttcactaggggggtacaaaccagtaccaccacgacaacagcact tctgtttccccggtgatgtctgtatagactgttaaccacgggttgaacagattgtatccgttaccgtctt gtacttgcgaagccagctatcacttggaaatcttgcgtgttgcgtcagcactaacccagag tgttagcttaggtcgtatgtctgtgacactctcaccggcgacggtgtgtccaggctgcgttggcgg cctacctgtgtgtccaaagccacaggacgctagtgtgtgaacaagggtgtgaagagcctattgagctac aagagaatctccggccctgaatgcggtatcttaaccacggagcaagggtacacaaaccag tgtatattctgtctgaacgcgaagctctgtggcgggaaccgactacttgggtgtccgtgtttctttgt ttttatcatggctgtctatgtgtgacaatcaagattgttatcatatagctgttggatttggccatcgttaa tttatgtagatttggacatttgcgttcttctcaacaatttaccctatttcaattgcagtaactgcacaa
503	PV3	taccttgtacgcctgttttatactccctccccgcaacttagaagcatacaattcaagctcaaatagga gggggtgtcaagccagcgcctccgttggcgaagcactactgtttccccggtgaggccgcatagact gttccccaggttgaaagtggccgctccgttatccgctcatgtacttgcagaagcctagtatcgtctgt gaatcttcgacgcgttgcgtcagcactcaacccggaggtgtgactgtggccgagtgagctgtggac agtcctccactggcgacagtgtgtccaggctgcgctgtggcggccaccctgtgtggcccaagccacgg gacgctagtgtgtgaacagggtgtgaagagcctattgtgagctacatgagagctctccggccccctgaat gcggcctaattcctaaccatggagcagggcagctgcaaccagcagcgaagcctgtctgaacgcgcaa gtccgtgtgcgggaaccgactacttgggtgtccgtgttttctttattcttgaatgctgtctatgtgtgac aatcatagattgttatcataaagcgagttggaattggccatccagtggtgaatcagattaattactccctgt ttgttggatccactccgaaacgttttactccttaactattgaaattgttgaagacaggatttcagttgt caca
504	EV C102	ctttgtacgcctgttttatacctccctccccacgttaactttagaagcaattcaacaagtccaatagagg gggtacaaaccagttatcaccacgaacaagcacttctgttccccgggtgtatttatcataagctgtgcc caggctgtaaagtgaatgatccgttaccgcctcagtagtacttcgaaaagccctagtatcgtcttgggat cttcgacgcgttgcgtcagcactctaccccgaggtgtagctttaggtgtgatgtgtgtggcattcccc atcggcgacgatggcccaggctgcgttggcggccctaccatggctaacgccatgggacgtagtt tgtgaacaagggtgtgaagagcctattgtgactactcgagagctcctccggccccctgaatgcggcta cccaaccacggatcagggtgcctccaaccagggaggtggcctgtctgaacgcgcaagctgtgtggc ggaaccgactacttgggtgtccgtgttccctttatcttttaaatggctgtctatgtgtgacaatcataga ttgttatcataaagcgaattgtgattggccatccgggtgaatacaaacacattattactgtttgttggat ttactccgctcacacagcttactcctaagataattatttattgtattgtcgtgtgaaggagacactattata
505	EV 30	aagcaaggcaaacctgaccaatagtaggtgttggcaccacagccgctatttgttcaagcacttctgt ttccccggaccgagttatcaataagctgtctacgcggctgaaggagaacccgttctgttaccggacc agctacttctgagaaaactagtaacactatgaacgtttcggaggtgttctgttcagcacttcccccggt agatcagggtcgtatgagtcaccgcttctcaggggtgacggtggcggtgtgagctgcgttggcggcgt gcctacgggttgcggccctaggacgctgaatccgacatgctgtgaagagctcattgaatcagctagctgt

		gtagctctcggccctgaatcgcgctaactcgaactcgcgagcaggtgctcagaccagtgagtagcctgtcgtaacgggcaactctgcagcggaaccgactacttgggtgttttctttcttcttataitggctgttatgggtgacaataaagaattgttaccataagctattggatggccatccggtagcagcagagccattgttacctcttgggttgatcttgaccttgaaccacaaagtctgaataacatcatctcatfttaaagttcaactcagctaaaagaaa
506	SA5	agtaacttggtatccgggtaccttgtacacatttacaacccctacccttgaaccttagaagaattatttaaccgctcactagggggtgtgtctatccaagcacatcaagagcaagcacttctgtctcccgggagggggctaattgtacgtgtgtcccacggcggaatgagccctaccgttaaccggcagtgctacttcgggaagcccagtaactacattgaaactttagggcgttactcagcacaataacccaatgtgtagtctgtgtcgtatgagccttggcatcccccacaggcgactgtggccaaggctcgttggcgccagcctgaggaccaaaagtccgtaggacgcctaattgtggacatgtgtgaagagcctactgagctagaatgtagtctcctcggccctgaatgoggctaactcctaaccctggagcatccgctgtcaaccaggtacgtagggtgtcgtaatgcgttaactgtgggatggaaaccgactacttgggtgtcgtgtttctgttttcaactgggtcgttatgtgttacaactaattgttgaatcattggcagtgcgcgctgaccacggcgtatttgatattccatttgggtatccaatagtgtaactcatatacacaactttaccactgatacaagataa
507	EV A114	tgtgcgctgttttgaacccccctcccccaactcgaacgtagaagtaattgtactactgactagtagcaggcggtggcgccacagccatgtctcgtatcaagcacttctgtttcccgactgagtatcaatagactgtcacggtgtgaaggtgaaaacgtccgttaccggctaactactcagaaaaactagtagcaccatagaaactgcagagtgttctcgtcagcacttccccgtgtagatcaggtcgtatgagtcactgc aatccccacgggtgaccgtggcgagtggtgtgtgtggcggtgctgtatggggcaaccatagga cgtctaaaggtgacatgtgtgaagagctattgagctagttagtagtctccggccctgaatgc ggctaactcctaactgtggagcgcatatccccaaaccaggggagcagtgctgttaacgggcaactc cgcagcggaaaccgactacttgggtgtccgtgtttcttttattctatactggctgttatgtgtgacaa ttgagagattgttaccatagctattggattggccatccagtggtgaatagagcaatcatttaccatt tgttggatttactcatttaacccacagctctcacaacactacatttcatcttactactgaacactaga aa
508	Mobovirus A	tattctccacaaacttctgttaactctgttaagccttttaccatccatgtaatttaatttctccactaaa aggaatttcccccatggctcttcttggctcgaacaaatgctacatagggctctgttttctccccctggctc tcttggccagggttccataccccaatttcttatttccatgattttcatcatgtgttatttttactgttcttata tttctgaggtgaccaactcctaagccgactgggtcgggaagcccgactcctcgcateactagg gtgcgtagcgtatgtagcgaaaatattgtgtgtagatgcatacatatagtaattgatactacacca aactctgttcttttgaactagctatttttaagtaaggtaggctacgggtgaaaccttaccattgcag gtacgtgaaccgcaacggacatttggccgaagactgtgttaccacgtcagttataggacctcttcaacgttgggtgacggcgtatgactgatttagttaggctagtgaatttaagttcagggggtatcttttagc ttaagcgtgtatttctagtaggacttgcagagcctccccacctaggaggatctgtgtttatagccccctt tcttgttccgttagttttccacacttttacaatatttgatgattgtt
509	Burpengary Virus	ctcccccccttccccctcccgagtaggagattggcatgtatgtctcatatgccgatttctcttctgt cactctcttaaatcctgtgtggcggtctcggatfaaacatttatgtcgtatctgggatcgtcttactgtgt ggtaattcctctgttgcctagggaacctccggactcgggattaaaggtctcaacagagggcaatgt acaagggaagtcattatagctaatgaattgtatgaatgactagtgtgacagggctgaggaaactc cccccgggtaaccgggtcctcagcgtccgaagacacgtgtagatggatccacctgttatacca gacagatgtaatagtcataactctgtattgtgtaggatgtataaattgtgcaattgtaaaatttggcg tagagatgtccgaaggtaaccggttttaccgggactgtatcggaggctaaftaaccaatgcgcctca afaactcatataatttcttttttattcaaa
510	Hunnivirus A1	taacgtttggcaagaacctcactgtcaattgggaccaccactttcagtgaccacctgcgaagtag tgagagagaataagcttcttacccttcatgtgaaccttcagtcgaagccgcttggataagata ggaggaaaagtcatcttaaatggagtgaacatgtacttcagaatttctagcacgcgtgggttctc ttgcgtgtgacggcactgtcttgcggagctctccacactgacaccccacgttggacacttgggtg gcagatgacaacactgcagctggaattgagtgtctgtgtaactctgtgttaacagtgaacaaatgt gatcacttccgtgagctagtagcctgtggaccaacaactggaacagttgcctcaggggccaaaaa gccacgggttttacagcaccctactgtgttattggagcaatccaagatgtcacagagttagtaattg ccaagcagtcctgactgttatcttgacataccgtgcagtttggatagtgaaggatgcctgacgggt acccataggttaacaagtgaacatattgatatgaagcaggggtcactctacgtctgttaccgtgg

[illegible]

[illegible]

		gcgcttgggttataaaactgtcacagcttctccaaaccaagtggtcttgggtttcactttatcaaactgttcc
523	EMCV2 RD1338	aaatactggtcgaaacccgttgggataagaccggggttggtaatgtctcaatgttattccaccca attgacgtcttttgcattggaggcagtgaaaccttgccttgccttgcagaggattccagtg tcttccgctcgcacaagggaattcatgatccacaaaagtgtgaagagagcagggtccatgga agctttctgacgactgafgactgttagcgaccttgcaggcagcggacccccacctggtga cagggtgcctctgcggccaaaagccacgtgtttaacagacacctgcaaaggcggcacaacccag tgccatcaaaaagctgactgtggaatagtaaccggctttcttaagcaaatgtgtgtcgg ggctgaaggatgccgggaaggtaaccactggtgtgactgtatccggggccacagtacatgtgc tttacacatgtagctgcggttaaaaaacgtctaggccccccgaaccacggggacgtgttctctt gaaaaccacgattacaat
524	EMCV1 JZ1203	gtctgctcgatctgcaggctgggtccgtgactacccactcccccttcaactgaaggctacgata gtgccaggcggggtactgcttaagtgcacccccaaaaacaacaacccccctaactact ggccgaccccgctggataaggccgggtgtcgtttgtctatatgttattccaccacattgcctgc ttttggcaatgtgtggccccggaacctggccctgtctcttgcagcagcattcctagggtgttctcc ctctcgccaaagggaatgcaaggctgtgtgaatgtcgtgaagggaagcagttcctctggaagcttctg aagacaacaacgctgttagcgcccttgcaggcagcggacccccacctggcgacaggtg cctctcgggccaaaagccacgtgtataagatacacctgcaaaggcggcacaacccagtgccac gtgtgtgagttggatagttgtggaagagtcaaatggctctcctcaagcgtattcaacaagggtgta aggatgccagaaggtaacccattgtatggatctgactgtggcctcgggtcacatgtttacatg tgtttagtcgaggttaaaaaacgtctaggccccccgaaccacggggacgtgtgttctcttga cacgatgataat
525	EMCV1 AnrB-3741	atgtgtgctgaagccacttgggaataagaccggcgtgcgtgtgtctatatgttacttccaccacattgcc gtcttttgcattgtgaggggcccggaacctggccctgtcttctgacgaacattcctaggggacttcc ccctctcgccaaagggaatgtgaagtgtgtgaatgtcgtgaagggaagcagttcctctggaagcttct gaagacaacacagcgtctgttagcgaccttgcaggcagcgggaacccccacctggaacaggtg cctctcgggccgaagccacgtgtataagatacacctgcaaaggcggcacaacccagtgccac gttgtgcgttggatagttgtggaagagtcaaatggcttcccaagcgtattcaacaagggtgta aggatgccagaaggtaacccactggttgggatctgactgtggcctcgggtgacgtgtgtttacac ctgtgtgagtcgaggttaaaaaacgtctaggccccccgaaccacggggacgtgtgttctctaga ccactatgacaat
526	Cosavirus D1	cgtgctttacacgggttttgaaccccacacggcgtgttggcgcttgcaggacagtaggtatatttctt tcatttctctttctagccgctgaggttctatctacggggcgagtgatactccgctccttcttggac aggcgggcctccacgctcttgttgatcttaaggctgccaagtcactggtgttgaagtgaagaatgg agagacactaggggcgttctatgtgtgttgcaggggattgtagcgtatctgtgtgtgtgtcggtatt ccccctggtggcgacacagcctcagaggccaaaagccctgtccgaaggaccacacagtggtg gtgtcccgacccctccttcaaaagcttgtgttaacaactttgttttagactttcttaagcttctca catcaggccccaaagatgtcctgaaggtaaccctgtgtatctgaggtgagcaccacacactacc ggactgtgtggacgtgtccacagacgcatgtgttatccagccccctctttaggagggggct tttctcgtcagcacaggtatctcagagattcatctctgtgtctttacaccagacatggattta aaaattgccaaaggccctggcaacaacctaggggactaggtttctctattttaaaagatgtcaat
527	Cosavirus B1	cgtgctttacacgggttttgaaccccacacggcgtgttggcgcttgcaggacagcaggtttatttctt ttaactctcttcttagccacacacgactatgtgtgtgtggcgagtgatactcccgcttcttcttggga caggcggtcctccacgccccttgtgtgacttaaggctaccaagtcactgtgtgtggaagtgaagag aaaggagttccttgggaactacatgtgtgcatgacagaggtgttagcgtatctgtgtgtgtgtcggt attaccccggtggcgacacggaccccacaggccaaaagccctgtccgaaggaccacacagt ggagcaacccacgtcctcttcaatgtttgtgttagcaaccttggattatttcttcaagctcca atacaccggggcccaagatgtcctgaaggtaacccgtgtatctgaggtatgacacatcaacta cccgactgtgtcttctgagaacagacgcatgtgtgtaaccagccccgatcctaagggtcggggg ctttgtcactcagcacaggtatctgacaggagacctccccctgtctttacaggggcggggg tttaaaattgccaaaggcctgtgcaataacctaggggactaggtttcttttttttaaaagtgtca at
528	Cosavirus A SH1	ccgtgctttacacgggttttgaaccccacacggcgtgttggcgcttgcaggacagcaggtttatttctt ttatgtctttatttctagccaacagggttctatctctgttggcgagtgatactcccgcttcttcttggga

		cagattgcctccacgactcttggatctcaaggtgatcaagtcactggtaaatagagcgaaggttg aggaaacctgaggaaattccatgtgtgttggccaggagttgttagcgatgctgtgtgtgtgtcggtt tccctcatggaacatgagcctcacaggccaaaagccctgtccgaaaggacccacacagtggga gcaatcccagctcccctctacaagcttggagaatgaactcacgtttatctcttctctgtttac atcaggcccaagatgtcctgaaggtacctgtgtatctgggcatgagcaccatcaactaccgg actgtcatttcgggtgcagacacatgtgttaccagccccctctgtcttggcagagggtgtgtgtcgtc ctcagcacgagatctgatcaggagcccttcccagtggtctttacacctggcggtgggttaaaaatt ggccaaggccctggcaaaataacctagggggactagggtttccctttattaacaatgtctgtcatt
529	Malagasivirus B	ctttattttctatgtaactcttctttaagttttatttgcctactgtgagcttatgcggggaccactgtctta gacaaccccacattgtcatgagtaagtacacgcaaccattacgattacttttaaccgtctgacctttt gataacaactgaagttaggcgtgaaacatgcaittataccaaaagtgcggcgcaatttccccactacg gtgggggggtcaccctactggttggaaactgtagccattatgtgttgcctggccttcaggatctcac aacacaacagttctctacaatggaaatgggtgagattgcagtgacatgaacaagtatctagtagt acatagactcaagcctagttgcctgcggaacaacatgtggtaacacatgccccagggtccaaaag acaagggttaacagcccccttaggtgtgtgtgtgtgaagaatacttagtagtgtgttatgtatctcac ctgttagtacagaatgagtagtgcctgtggaaggatgtcctacaggtaaccattatagctgtcgt aggagaccactagtggtgtgttaccgcaggtgaggtgtttaaagcgtctagccaagccaac agcactagggtatagtccttctatatttataatttcagtgat
530	Mosavirus A2 SZAL6	ccccccctcaaatgcaacgataatgagctaatggcgagattgagatgctatatcaccctctctaagt atagacctcatctgattgataaggacgtatattgtgcgaaaccgcttggaaataagaccgatgcgct agtcatgatgatgataagatctaggaacttaccatctgtcttatgtctatgtaagtagaggggca ggcctcattgcccataattcttctaccgagtaactgtctagggtttctagcggccttgaaacttggattga gggatacaagatactactgagctgattgtcgttgggaaactttagatacttcaaaagctaccagtag cgtggactcacagccagcggactacccctcatgtaacatgagcccttggcccaagggcacg tcgcaagacctgtgagacggcaaccccagcctagcttgttggaggaaacaagcgaatacatgaca tgagagaccggaaggattctgtattgtgagcgaaggatggccttaggtacctcattttatgagat ctgaggagggtccttgcgtgtgtgtcttactgacacacagagttaaaaagcgtctaaagctcac ccggaaattgggaaattccgttatttccattttgttgcgaaagtcgttc
531	SVV	ctggccctcatgcccagctcttcttcccttccgggggggtaaacggcgtgtgttgcctagaggc acaggaggcaacatcacaactgttcttgggggaacgggtgcggctcccaattcctgcgtcgccaa agggttagcgccaccaaacggcgcatctaccaatgctattgtgtgtgtcgtgagttctagcctac tcgtttctcccctatccactcacacgcacaaaaagtgtgtgtgtaattacaagatttagccctgcac gagatgtgcgataaccgaagattgactcaagcgcggaaagcgtgtaaccacatgctgttagtc ccttcatggctgcgagatggctatccactcggatcactgaactggagctgcaccctccttagtaag ggaaaccgagaggccttctgcaacaagctccgacacagagtcacgtgattgtaccaccatgag tacctgttctcccctcgcaccaggacttcttttgaatccacggtcctgacagagggtgggg catgafcccccctagcatagcagctacagcgggaactgtagctaggccttagcgtgccttggatac tgctgtatagggcgacggcctagtcgtgtcgttctataggtagcacatacaaat
532	PTV A	actagtccttggactttgtgtgtttaaacacagaaatttaattacctggccatgaattcattggattaa cccttctgaaagactgtctgtgcgcgagctaaagcgcgaattgtcaccaggattgtcaccagtggt ggcgacagggtacagaagagcaagtactcctgaccgggcaatgggactgcattgcatacccta ggcacctattgagatttctctggggcccaccggcgtggagttcctgtatgggaatgcaggactgga cttgtgtcgtcctgacagggtgcggcgtgctgtacttctgtatagtcagttgaaactcacc
533	PTV B	cttcttttaattcgtactgataagtgatagtccttggaaagctaggattttgttacgctagtgtttgatta tcttgtgcccacattgttttgaacatatgtgtgtttaaacacagaaatctagtttcttggttatgagt ttaatggaaatcttcttgaagacttgccttggcgcggttagagcgcaattgtcaccaggattgtc accaatgggtggcgacagggtacagaagagcaagtactcctgactgggtaatgggactgcattgca tatccctaggcatctattgagatttctctggagccaccagcatggagttcctgtatgggaatgcagg actggactgtgtcgtcctgacagggtgcggcgtgctgtacttctgtatagtcagttgaaactca tt
534	Tottorivirus	cccccttacgtaactgcaacttaagagtaccctactgcatggatgtgtgtgtaaatctttacgcacac attgttagtagttagttatgttctacctaagtatgcatgcacccgtcgaaacacgctgtgtgataa gataggtgagtcattgtactaatctcattaagataaataagcaccctacaacgcacggcacgctc gtgtctccgtgcggggccgggacaacagcggcctaaattcttaggtaccaccatgcttttggg

		actatggcaccactgtggacgtgagtaacctggcagtaagctgtgaaaagatggaaaggtgtccca agctatggggcgatgcatatagcctgcggaacaaacacggcgacgtgtgtcccaaggcccaa aaggcacgtggataagatccacctatattgtttacccatagtgtaagtcactgggaagtcctagtaag gatgtctggagtaaggctcacggggtagggcgaaggatgccagaaggtaccogtaggtaacct taagagactatggatctgctggggaccggatggcgccatcaccatgacgtggagggccggtta aaaaacgtctaagcccgaacaaacactaggggactaggtttctttttattatcatgtatgacgtt
535	Posavirus 1	acatttccttgcgtgcgcacccgaaaatttattgaacttggcttgaatcataagtaafgtctttagc ggacactttgagaatataatgctgtatggattaatgtatactgttttaataacaaacctagacacgc agttgcgttatgtgtgtatcaatcacatataagtggtgaactcgtgttaactcctgacgtatattgtt gccgcctacttccaataaaatagattacatgcgcgtcatgccttggggttacctattggccttgac aaaaacaagtcgtaagattgtagcttcccggtgtaaaaagctgggcgcggctgtgctcgttagg gtggaaaggtccaccaatggctgtgtgagtgtaagctcgggtgtcctgtgtgtgcgaattccaggc gtcgtataaactatattgcatctgacttaactcttggctctactgtatctagttctgttctactaact ctaataactactgtgcttaactgaaaaactacatattgtaataatagataatccttgatcctgat atccctcacgtcactgaagttgcgcgaaaacgaatttcagatcatatcaattgaatcgaaggtgtac ttgcttgaacctactataaactcctcatgctgtttctatttctactcttggc
536	A105-675	ccaccacagcaagaatgccatcatctgtcctcacccccatttctccctcttccctgcaaccatt acgcttactcgcattgtgagtggtgcacgtgttgaaacacagctacactcacgtggggcg gggtttcccgcccttgcgtctcgcgaggccaccccttcccttctccataactacagtgctttg gtaggttaagcatctgatccccgcggaagctgtcgcgtggaactgtggggaccagacagg ttatcaaaaggcaccgggttcttccgctcaggagtagtccctgttagtaattctagtggggtctgc ttggtgccaacctccccaaatgcgcgtgcgggagtgcttctcccaactcacctagtagtctct catgtgtgtgtgtgtcagcatatctgagacgatgttccgctgtccagaccagtcacagcaatggac gggccaggtgtcgttagtcttccggcttccggcgcatgtttgtgaaccgggtggggttaaggt tgggtgtcccaacgccogtactttgtgacaactcaagaccaccagggaatgccaggagggtacc ccgctcacggcggtatcaccctgggctaattgtctacgggtgttcttctgttccatttcttctt ctgttc
537	A110-675	acctttgtgcgcctgtttataccccctcccccaactgtaacttagaagtaacacacaccgatcaaca gtcagcgttggcacaccagccacgttttgatcaagcacttctgttaccgccgactgagtagtcaataga ctgctcacgcggttgaggagaaaggttctgttatccggccaactacttggaaaacctaagtaaca ccgttggaagtgtcagaggtttcgtcagcactacccagtgtagatcaggtcgtatgagtcaccgc attccccacggcgacogtggcggtgtgctgctgttggcgccgtgccatggggaaccatggg acgtctataacacagatgtgtcgaagagtctattgagctagttgttagtctcgcggccctgaatg cggctaatacctaactgcggagcacacacctcaagccagaggcgagtggtctgaacggggcaac ctgcagcgggaaccgactactttgggtgtcgtgtttcattttattctatactggtgtcttatgtgac aatgagagatgcttaacatatagtattggattggccatccgggtgactaatagagctattatatacc cttgttgggtttataccacttagcttgaagagggtaaaacattacaattcatgttaagtgaatacac caaa
538	18-675	cccacagcaagaatgccatcatctgtcctcacccccatttctcccttctccctgcaaccattacg cttactcgcattgtgagtggtgcattgtgtgaacaaacagctacactacatgggggggggtt ttcccgccctacggcctctcgcgaggccaccccttccctcccttataactacagtgctttgtag gtaagcactcctgatccccgcggaagctgtcacgtggcaactgtggggaaccagacaggtatc aaaggcaccgggtcttccgcttcaggagtagtccctactagtgaaattctagcggggctctgtgtgt ggcaacctccccaaatgcgcgtgcgggagtgcttctcccaactcacctagtagtcttctcatgt gtgtgtgtgtcagcatatctgagacgatgttccgctgtccagaccagtcagtaattggacgggc cagtgctgttagtcttccggctgttccggggcatgtttgtgaaccgggtggggttaagggtgtgtg tgcccaacgcccgtactttgtgtacacctcaagaccaccagggaatccaggagggtaccaccac ctcacgggtgggatctgacctgggctaattgtctacgggtgttcttctgttccacttcttctctgttc acg
539	A115-675	acctttgtgcgcctgtttatacccccccccaacctcgaaacttagaagtaagcaaacccgatcaata gcaggtgcggcgccacagtcgcattgtgatcaagcacttctgttaaccccgaccagtagtataata gactgtctacgcggttgaggagaaaacgttctgttaccgggctaactacttgcgagaaccagta gcatcatgaagtgcagaggtttctcagcactacccccgtgtagatcaggccgatgagtcac cgcacttccccacggcgaccgtggcggtgtgctgtgtggcgccgtgcctatggggcaaccat

		aggacgctctaatacggacatggtgaggaagagctattgagctagtagtctcccgccctg aatgaggcctaactcgaagcagacacacacctaatacgaaggcagtgctgtaacgggta actctgacgagggaacgactacttgggtgctggttctttaaattttactgctgctattggtgac aatgagggaattgtgccaatagctattggtgccaacggtgactaacagagctattgtgtcca attgttggaattacccgctcacactcacagtcgtaagaacccctcattacgtgtttttcaactcaa gaaa
540	A73-675	ttactcattcagcttcttcggaacctgttcggagggaattaaacgggcacccatactccccaccc ccctttgtaactaagtatgtgtgctgctgactctgactcccacggaacggacccgattggtgaa caaacagctaggctccacatctccctccctggaggggcccccgcctccacatctccccc agcctgacgtatcacaggctgtgtaagcccccggaagctgctcagtggaattgtggtgctc cccttcacaaagacaccaggcttcttcccttaaggctagccccggcgtgtgaattcagctgggc aactagtgtgtgctgctgctcccaatctcgcccgaggagtgctgttcccaagccaacccct ggcccttcactatgtgctggaagcatactgagaagggtgttcgctgtggtgccaacctgtgga cagggtgccccagtggtgtaaccttcttccgtctcggacggtagtgattggttaagattgtgta ggttcattgtgccaacgccccgtgctgggatgaaaccttactgcccaggaatgccaaggcaggta ccaccccggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaatctttatgt cggagtc
541	Kobuvirus 16317	ttactcattcagcttcttcggaacctgttcggagggaattaaacgggcacccatacaccatcc ctttctgaactaagtatgtgtgctgctgactctgactcccacggaatggatcgatccgtggaga caaatgctagatccacatctccctccctggagggaacctgtgctccacatctcccccag cctgacgtaccacaggctgtgtaagcccccggaagctgctcagtggaattgtggtgctccc ccttcacaaagacaccaggcttcttcccttaaggctagccccgatgtgtgaattcattgggcaac tagtggtgtcactgtgctgctcccaatctcgcccgaggagtgctgttcccaagccaacccctggc ccttcactatgtgctggaagcatactgagaagggtgttcgctgtgctgctccagcctgtgaaca gggtgccccagtggtgtaaccttcttccgttctcggaaggtagtgattggttaagattgtgtaagg ccatgtgccaacgccccgtgctgggatgaaaccttactgcccaggaatgccaaggcaggtaaccc acccccgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaatctttatgtc gagtc
542	Aichivirus Chshc7	ttactcattcagcttcttcggaacctgttcggagggaattaaacgggcacccatacaccatcc tttctgaactaagtatgtgtgctgtaactctgactcccacggaatggatcgatccgtggaaca aactgctagatccacatctccctccctggagggaacctgtgctccacatctcccccagcc tgacgtaccacaggctgtgtaagcccccggaagctgctcagtggaattgtggtgctcccc ttcatcaagacaccaggcttcttctcctaaggctagccccgatgtgtgaattcattgggcaacta gtggtgtcactgtgctgctcccaatctcgcccgaggagtgctgttcccaagccaacccctggc cttactatgtgctggaagcatactgagaagggtgttcgctgtgctgctccagcctgtgaacagg tgccccagtggtgtaaccttcttccgtctcggacggtagtgattggttaagattgtgtaagggtc atgtgccaacgccccgtgctgggatgaaaccttactgcccaggaatgccaaggcaggtaacccac cctcgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaatctttatgtc gga
543	Aichivirus Goiania	actccattcagcttcttcggaacctgttcggagggaattaaacgggcacccatacaccatcc ttttgcaactaagtatgtgtgctgtaactctgactcccacggaatggatcgatccgtggaaca aactgctagatccacatctccctccctggagggaacctgtgctccacatctcccccag cctgacgtaccacaggctgtgtaagcccccggaagctgctcagtggaattgtggtgctccc ccttcacaaagacaccaggcttcttctcctaaggctagccccgatgtgtgaattcattgggcaac tagtggtgtcactgtgctgctcccaatctcgcccgaggagtgctgttcccaagccaacccctggc ccttcactatgtgctggaagcatactgagaagggtgttcgctgtgctgctccagcctgtgaaca gggtgccccagtggtgtaaccttcttccgtcccgacggtagtgattggttaagactgtgtaag gttcattgtgccaacgccccgtgctgggatgaaaccttactgcccaggaatgccaaggcaggta ccacctcgggtgggatctgagcctgggctaattgtctacgggtagtttcatttccaatctttatgtc ggagtc
544	Aichivirus ETHP4	ttactcattcagcttcttcggaacctgttcggagggaattaaacgggcacccatacaccatcc cttttgcaactaagtatgtgtgctgtaactctgactcccacggaatggatcgatccgtggaaca caaatgctagatccacatctccctccctggagggaacctgtgctccacatctcccccag cctgacgtaccacaggctgtgtaagcccccggaagctgctcagtggaattgtggtgctccc

		cccttcattaagacaccagggtcttctccttaaggctagtcctcgatgtgtgaattcacattgggcaac tagtggtgtcactgtgctgctcccaatctcggcggagtgctgttcccaagccaaccctggc ccttactatgtgcctggcaagcatactgagaagggttccgctgtggtccagcctggtaacag gtgccccagtggtcgtaaccttctcctgcttccggacggtagtgattggttaagattggcgttaagggtt catgtgccaaagccctgtgcgggatgaaacctctactacccctagggaatgccaggcaggtaaccca ccctcgggtgggatctgagccctgggctaattgtctacgggtagtttcatttccaattcttatgtcgg agtc
545	Aichivirus DVI2169	tactccattcagcttctcgggaacctgttcggagggaattaaacgggcacccatacacccccacccc ttttctgcaactaagtatgtgtgctgtaactctgactccacgggaatggatcgatccgctggagAAC aaactgctagatccacatcctccttccccctggaggacccggctcctccacatcctccccag cctgacgtatcacaggctgtgtgaagtcctccggaagctgctcacgtggcaattgtgggtcctcc cttcacaaagacaccaggcttcttctccttaaggctagccccgatgtgtgaattcacattgggcaact agtggtgtcactgtgcgtcctcaatctcggcggagtgctgttcccaagccaaccctggc ccttactatgtgcctggcaagcatatctgagaagggttccgctgtggtgctccagcctgttaacag gtgccccagtggtcgttaaccttcttccgtctccggacggtagtgattggttaagattgtgttaagggtt catgtgccaaagccctgtgcgggatgaaacctctactgcccctagggaatgccaggcagggtacccca ccttcgggtgggatctgagccctgggctaattgtctacgggtagtttcatttccaattctttatgtcggg gtc
546	Aichivirus DVI2321	gcttctcgggaacctgttcggagggaattaaacgggcacccatacacccccaccccctttctgcaac ttaagtatgtgtgctcgttaactctgactccacgggaatggatcgatccgctgggagaacaaactgcta gatccacatcctccttccccctggaggaccccggtcctccacatcctccccagcctgacgtta tcacaggctgtgtgaagtcctccggaagctgctcacgtggcaattgtgggtcccccttcatcaa gacaccaggcttcttctccttaaggctagccccgatgtgtgaattcacattgggcaactagtggtgtc actgtgcgtcccaatctcggcggagtgctgttcccaagccaacccttggcccttactat gtgcctggcaagcatatctgagaagggttccgctgtggtgctccagcctgtaacagggtgcccc gtgtgcgttaaccttctcgtctcggacggtagtgattggttaagattgtgttaagggtcatgtgcc aacgcccctgtgcgggatgaaacctctactgcccctagggaatgccaggcagggtacccacacctggg tgggatctgagcctgggctaattgtctacgggtagtttcatttccaattctttatgtcggagtc
547	Aichivirus rat08	tactccattcagcttctcgggaacctgttcggagggaattaaacgggcacccacttctcgtcctcctcc ctttctgtaactcaagtgtgtgctcgttaactctgactcccggaatggatgacgcctcctggtgaaca aactgctaggtcatctcctccccaccccttggcgctcctcggcggtccacacccctccccagcc tgacgtgtcacaggctgttacaagaccccggaagctgctaactgtgcaattgtgggtcccccc ttgttaaaggaaaccaggtcttctccttaaggctagacccctgtgtgaattcacagggtggcaactag tggttcactgcatgctcccgacctcggcggagtgctgttcccaagctgtaacactgaccttc acttatgtccttgcaagcatatctgagaagatgttccgctgtggtgctgccaacctgtgtaacagggtg ccccagtggtcgttagtcttctcgttctcggacggtaggtgttaggttaagatgcggcgttaagggtt caagtgccaacgcccgtgaagggtgaccttctactgcccctagggaatgcccgaggtaccccc agggttcgctgggatctgagcgggctaattgtctacgggtagtttcatttccctcttcttccactgg catc
548	Aichivirus Rt386	actccattcagcttctcgggaacctgttcggagggaattaaacgggcacccacttctcgtcctcctcc cctttctgcaactcaagtgtgtgctcgttaactctgactccacgggttgaccgccccgtgtggaac aaacagctaggtcattcctcctacccctgggcgccatttcagtggttcatactcctccccag cctgacgtgtcacaggctgtgcaagtcctccggaagctgctcacgtggcaattgtgggtcccc cctttgtgaaggaaaccaggtcttctccttaaggctagacccctgtgtgaactcacagggtggcaac tagtggttccactgcatgctcccgacctcggcggagtgctgttcccccaagctgtgacactgac ctccacttatgtgcttggcaagcatatctgagaagatgttccgctgtggtgctgccaacctgtgtaaca gggtccccagtgctgttagtcttctcgtctcggacggttaagtggtgtgtaagatgcggcgttaa gggtcaagtgccaacgcccgtgaagggtgaccttctactgcccctagggaatgcccgaggtacccc cccaggttcgctgggatctgagcgggctaattgtctacgggtagtttcatttccctcttcttccactgg ggcatc
549	Norway Rat Pestivirus	gtataagggttgggaacctgttaccagctacctctgccattcagtafttgggagtagaagtagatgt gtttacaactcacacgtgtggggcggggatagactgtgccagcgtgtgttaaccagacctac gcatacgtgtggactgcgaaccaggagagcacctaggctcgtgacaagctgtgagaacacagtagt cgtcagtgagtgcgtgtaaggatcaccacctggatactcacgtggacgaggagggttccag

		tcagaaacctacaccagaggagggtctcttgagacatggatggtctgagtaacagactacteta ctgggggtgtgtcctgacagggtctcggctgatagcctggctagcagtataaaatcagttgaatt ggcataatgagttgtgaacatctagtaaacatgaaagacaaaaacaaaaatgagcataataaaaa aattgtacaatccactactcagggtgtgtcgcagactt
550	Porcine Kobuvirus GS2	tttgaagagggtgtggggggcctcggccccctcaccctctttccgggtggccattcggccgggc caccgttactccactccactcctcgggactgtttggaggaacacaacagggtctcccatccctgt ttaccctttatccatcatctcttcccaagttaccctatccacacccactgactgactcctttggattt tgacctcagaatgectatttgacctccactcgcctctccctttcgggtgtccgggtgtcctggcg gaaaaagcacaagtgtgtgacaggctacaaactctaccggacaaaggtgctgttcggcggtgct gagtaattggataggagatgctctacaacaggctcgcctatgagtagagcatggactgcggtgca tgtgactcgggtcaccacgggcatagcattgctcaccgtgaatcaagtcatcgagattctctgacc tctgaagtgcactgtgtgtcgtggctgggaatccacgcttgaccatgtactgctgtatagagtcgc ggctggccgactcatgggttaagtcatgtgacaagacac
551	Kobuvirus SZAL6	cctaccaagggttacatgggaccatactcctcctccctgtaactttaagttttgtgcccgtattcttg actccaggcggatgtgtgtcggcctcctgtgaacaacagctagacactttcctccctcctct gggctgctccggcagtcactcctcctcccccagcgtaacatgccccgtggagtgatgcacctgg aagtctgtggacgtgggttagtaacttcgggtgaaaaacccactataatgacaactgttgacccccac actcaaggactcgagctttctcccttaaggctagcccgccacatgaatttcagctggcaacta gtgagtcacacatgtccgcaacctcggctggagtgctgttccccaaagcgtatgcttctctg taagagtgcctggcaagcacatctgagaagtcgttcgctgctgccaacctggcgacag gtgacccagtgtgctgtagacttctcggattcgtccggctcttcttaggaacatgctgttaaggt tcatgtgccaagccctgcgcgggtgttcttactgcccaggaaatgtgcgcgaggtaccctac ttcggtagggatctgagcggtagctaattgtctacgggtagtgttcatttccatcttctctcaggtcgac atc
552	Kobuvirus sheep TB3	gacctctggtacttctcgcctgggtcacaaaagcgaagaacctgacctctaacgcccagacgag cggcattaaactgaactctggcactctcactctccttttccctgtcccttcccactgcgtctc aaggctcgcgaatcctggactagcccagttttaaaggctcgtggcaccctttgcccctctagggcc ttaaggtaggaactgacctgtgtgctgctcgggtgcgggagtgctaccacgtatgcatcgtgaagc ctcgtttctggttctgcctggcaaggctacagagtaccgtgttccgctgtgtggatgcatccgggtg accggaccccagtggtgtgtagcgggtatgttcacgggtccgctgttcacagatcctgacactgg ctttgctagaatgtgtgtgcccataccctgtgaccagtatcaattacatcacctaggaaatgctagg aagggtacccagctcgtgagctggatctgactaggttaattgtctacgggtgatgctccttttattt ttacaactgctattgactgtctgattgtgctcttctgtctcttctgctcgtcattctcaaggg ttctctttgtccaagatccttgggtctcctgttccacttgcactgccaacgctgtc
553	Pronghorn antelope pestivirus	gtatacgcagtttagtctatcctgtgtatatacagattggagactctaaaaaacagattcggaaatagg gcccgcggcgaaagaccgaagacaggctaacatgccgttagtagggctagcaccanaaacgcg ggaaactagacacttaggagagtgctgtggtactctaaagggtgagtagacacttaaccgtcaagg gttctactcctcagttgagtagatagatgccctgtggacgggggcatgcccaagatgtagcttag ccggggcggggtgttgcgggtgaaagtagcaatatgaccacactgcctgataggggcgaggca ggccccctaggtatgctagataaaatgtctgctgtacatggcac
554	Porcine pestivirus isolate Bungowannah	tacgcgggtataacgacagttagttcaagtgtcgttatgcatcattggccataacaaattatctaatt ggaataggacgtcgacctgtacgaaggcggagcgtcggtagccattcgcactagtaggacta gtacaaatagggtcaactggtgagcagggtgagtgctgcagcggctaagcggtagtacaccgt attcgtcaacagggtgctactggaaagatcacccactagcgatgcctgtgtgacgaggacatgtc caagccaatgttatcagtagcgggggtcgttactgagaagctgcccagaatgggtaggtgacat acagtctgataggatgccggcgatgccctgtattttgaccagtataaatatccgtgttaagcat
555	Porcine pestivirus 1	gcagatatcgggtgtggacctgggggtgggtcaccgtgcccttcatgggtgagacctcactg cttgatagagtcggcggtgcctcaggtaagagtataaaatccgtgttactaac
556	Pestivirus giraffe- 1	gtatacagatttagctcaatcctgtatacaatattggcgctaccaaatatagatttggcataggca acacccgatgcgaaggcggaaaagggtacacatgcccttagtaggactagcaaaaatcggg gactagcccagtggtgtgacttctggtgacgaagccctgagtagcaggcagtcgtcaacagt tcaacacgcagaatagggttgcgtctgtatgtgtgtggacgaggcgatgccacggtacatctt aacctatccgggggtcggataggcgaaagtcagttatggactgggagtagcagcctgatagggtg ttgcagagacccatctgataggctagtataaaaactctgctgtacatggcac

557	Classical swine fever virus	gtatcacgagggttagtgcattctctgtagtcatgattggacaaatfaaaatttcaatttggatcaggggctc cctccagcgacggccgaactgggctagccatgcccacagtaggactagcaaacggaggggacta gcccgtagtgccgagctccctgggtggtctaaagtctgagtagcagacagtcgtcagtagtgcac gtgagcagaagcccacctcgatatgctatgtggacgaggggcaigcccacagacacacttaacctt agcggggctgctaggggtgaaatcacaccagctggtggagtagcacctgatatgggtgctgcag aggcccactattaggctagatataaaaatctctgctgtacatggcac
558	Human pegivirus isolate JD2B11	tcagggttggtagtgcgttaaattccgggtcaccttggtagccactatagggtgggtctfaagagaaggt taagattcctcttctgctcgcggcgagaccgcgcacgtccacaggtgttggccctaccgggtggg aataaaggcccgacgtcaggtcgtcgttaaacggagcccgtagccaccctgggcaaacgacg ccacgtacgggtccacgtcgccttca
559	Human pegivirus isolate GBV-C-ZJ	cccgacactgggtgcgaagcccagaaaccgacgcctatttaaacagacgttatgaacgggccc gacccggcgaccggccaaaagggtggtggatgggtgatgccagggttggtagtgcgtaaatccg gtcatcttggtagccactatagggtgggtcttaagggttggtaaggctccctctggcgctgttggcga gaagcgacgggtccacaggtgttggccctaccgggtgtgaataaggcccgacgtcaggctcgt cgttaaacggagcccatfaccacacttgggcaaacacgcgccacgtacgggtccacgtcgcctaca atgtctctcttgaccaataggcttggcggcgagttgacaaggaccagtgggggctggggcgacgg gggtcgtatagggaagaaaaatgccaccgcctcaccggaaggttctgggtaccgccggctgcga ggccggccgggagctggggtagcccaagaaccttcgggtgaggggcggggtgacattttctccta taccgac
560	Human pegivirus isolate JD2B8C	tggtcaccttggtagccactatagggtgggtctfaagagaaggttaagattcctcttctgctcggcg agaccgcgcacgtccacaggtgttggccctaccgggtgtgaataaaggcccgacgtcaggctcgt tcgttaaacggagcccatttccgcctggggcaaacgacgccacgtacgggtccacgtcgcctttt aatgtctctcttgaccaatagggttcacccggcgagttgacaaggaccagtggggggcggggggtca cagggtatggaccctgggcccctgccccttccggcggggtggggaagcatggggccaccagct ccggcgcgcgctgcagccgggtagcccaagaaccttcgggtgaggggcggggtggcattttctt cctataccgac
561	Hepatitis GB virus A	ggcgggttgggaaggcccggaaccgccccaccacctcaactagggttggtaagggtacgtctatcggt ccggctgcccgaaggcggtggatcctgtgtgttaggggtcgtagggttgaataccagcacag gttgtaatcgctatagggcaggttatcccggtgacggcttccttggatcctggagcgggctgtgg cggcacggtccacaggaagtggggcctccggtgtgaataagccctgtctggagcatcagacgtt aaactgagacgtcccgaagagatcggaacgacgccccacgtatggcaacgcggcttaaaacct tcggggacagctatcgggttgacaatgccagtgggggcgggggccactattgttgggtcc gagttcctctagggtggccgaaaggcagccatggggccaccaggcgggcgccgtgtcacagg cggcaagggggaaaaatccttcgggtgaccccggttgcatcctcctccttagcagcatgagtg gtggtagctgcaacc
562	Simian pegivirus	ggggaatctcacccccgtccgggtccggaagaatcggaaacggacacccctgaccaatcattctt gatcatagagtgatgttagtgaaagccagacgaaagccggcgatgggttggtagcagggttgg taggtcgttaaatccggccaccctgtgaccctgtataagttgggcggaagctgactgaagctcgt tgctctttctgtgcgttcttgggtgcacggtccacaggtgacgcctataccgggtgtgaataataggcc gactcgagcggagctgttaactgagaacctcctacaggtgacgcaacttggcttgcgtacgggga cgccgctaaagtcacagtggtgttaagtcggcggttgacaacccagcaaggcgagggggtc ctattgttgactctgacagttccgggtggaggtaggcatgggttggccagctccggcgcgcc tacagccgggtagcccaaatccgaagggtgaggggcgggccacatgtccgaatttagtcaag c
563	Pegivirus I	agaatgggtctaagtggttggccacgtgggtccgaaggggaggaggacctacgctgccagggttgg cagggtcgttaaatccgggtgttaggagatccctccttcttaggactcgtgtgtagctgggggtcgtg gacccctggggcaaccgccaacccggacgacccgggtggcggtcctcatgttggcagcgtccac agggtgtgaacctacccgtgtgaataagggttgggtgttgggtccaccttaaacgtagtatgcatt gggcttggtaaacacgcgtctgtagtacggaacggcgcctttaagacacagtaggcgttagccg gctgggtgacaatccataccgggggtgggggtgtgtgtcatggatgtccacaccaccttcatcg gcccctaaagcaagccataccgggggggaggcgcgcgccacgcacgtccggggcaagggggaa gaaccttcgggtgaccccccccccaaccacgtccgatcaatgttaattgttgggttaggcgtgaca ccggcaca

564	Pegivirus K	agaatggtgtgatccgtccgctccagcggaaagcggcggggatctagtgttaggtgttctg cgtaaatccacactagtgtgtacgtctgtataacgtgggagcagccggTggggtcgaccccc Aacctggcggtgctgtagcaccggacgaagcgggggggtgaacgctaaccggcgcccg ctccaacgttaggcacgtctgtgctggaagacgttaaacacagggccccctcaacctgatc cgaggccagagaccaagtagcggcccttttaaaggcgttactcgtccaataggatctctcgg cgggtgtcaaaccttctggtccctgtgtgtgtacgggaggggtggggcggggagtagaag ccccgccccgcatgggggtaccaagctcggcagcccagcagcgtggtgtaggggaaaaat ccttcgggtgacccctgtgtaccataaagtaattaacatgagcatgcccgtagggtgtgtctttctcc ttccttgggaaggcgggtggcacc
565	Theiler's disease- associated virus	tgataccgtgtcccggtacgacctcgcgctcccaagctcgccttgaggggggagcgttaagg cgcgtagtgggttagcccccaaacggagccaccctagttagtgaactttagaatgttagggaga ctaccgcttctgctgtttggggacctaatgatccgctgcccagggttcttgggtaaatcccggcgc gggtgtttgggttcagggtcagtaggggcagacgggcccagcagtcgtgttcttgggtaccaccac cctatccggacgacctccctcaggaaggtcggccacggctgtgtgctgacgacgacctataattca gtccgaggggcgcagccctcgttaacttaggcaagggttctcgtccattgatttggccagggtgtt aagtgaacgccgcccctttaattgtttaatagggttctttccggcggtgtgacaaacacttccctggg ctcttctgtggcctcgttcttctgtctcgcacccatgagcgacaggggggggacccgtcgca cagtcgcccaagaggaaaatccttcgggtgacctcgtgcgcaaccaatcccttcttctccatg gcgtgtctgtgtgcatgctgtg
566	Rodent pegivirus	ggacttcggtccctgttactctgcgagccaccgagagccagggtgtgtacgcccgaggtgtt agaccccgccgaaagctcctaaccatgggttagtaggacgtgtgtaaatgccactgaggggtt ggagagctgttagagcgagtagtgcggcgttaaggcccgagtagggcctccagccccgggtca gcctaaccctgctgtgataccgggtcagtgagggcggtgtcccaacgctcgtatgctgtagggt gggtccctgcagttgggtgtgtgtacctgtctgtactgtctgtatagagtcgccggcgagggacc agctctcgtcagtcctgagggtgcac
567	Human pegivirus 2	aactgtgtgttagcaatgcgcatattgctacttcggtacgcctaattgtaggcgccggccgacc ggccccgcaaggccctagtaggacgtgtgacaatgccatgaggatcatgacactgggtgag cggaggcagcaccgaagtcgggtgaactgcactccagtcgacaccctggtctgtgtgtctatg gagggtcatgcccacgggaacgtgctgtgcaagggtatgggtccctgcactggtgccaatgcg cggcaccactcgtacagcctgafagggtggcgggcgggccccccaggtgtgacgtcgtgag cgcaac
568	GB virus C/Hepatitis G virus	cccccgactgggtgcagccccagaaacggacgectatctaatgtagacgcaatgactcggcg ccgactcggcgaccggccaaaagggtgtgtgatgggtgtgtgacagggtgtgttaggtcgttaatcc cggtcacctgttagccactataggtgggtcttaagagaagggttaagattcctctgtgtcctcggc gagaccgcgcaggttcacaggtgtgtggcctaccggtgggaataaggggcccgacgtcaggtc cgtcgttaaacaggagcccggttaaccacgtgggcaaacgacgcccacgtacggtccacgtgcctt tcaatgtctcttgaacaaataggcgtagccggcgaggtgacaaggaccaggtggggccggggg cttggagagggtacccaagtcgcccttcccggtggcggggaaatgc
569	Equine Pegivirus 1	agaatggggagttactcctggcactggcccgaaagcatgaactgacgcgggtggcagggttctt gggtaaatcccgcccggtgtgtgtgtgttagggcagggtgacagtcggcagggtcgaccccc tgcttcaggaccactgtcttctggacgacctgtgtgtaaaaaggcccgccacgggtctgtgtcgt ccgacgttcttaattcaggccggaggaccacgctcgttaatcgagcccaagtaicnaaacccag cacccttgggtcagccctacgcggcccttttaacgcttcggctaatagggtctatccggcggggt gacaaaggcgcgagggttacctgttactacgagcttgggtgtccctggagtaatccagggtgc c
570	Culex theileri flavivirus	atataaatccagtttgggttaaacctattcaaggcttaagttgtttattttatcgccgctcgtgactat aaagttgcttagcggagagagataaagaaggaaggttcaaggctcaggccagggtcgcaagtt ccctgtgctcctaggccgctcgcagggaaggaggtgaagaagaagaagagaaggagaggac caccgcccgaagaaggcaggtgtcctcacaagaggccaaccagcgtgttggaccagttggcca acgcccggacggcgtgtgtgctgtgggacgcctggggattggatggaggtccttctacagga agacatcgttcaagccatc
571	Bussuquara virus	agtatttcttctgctgtagaccattgcgacagttcgtaccgggtgagtttgaacttaacgagtgagaa aagtttctgaggaaagacgagaagcgaattctctga

572	Zika Virus	agttgtgatctgtgtgagtcagactgcgacagttcagagtcgaagcgagagctaacaacagtatca acagggttaatttggatttggaaacgagagtttctggtc
573	Yokose virus	agtaaaatttgcgtgctagtcgctgagcgtcagaccgcaaagtgagtttttagtgatctaaagtgagg agttattcttactgtcatcaaacactacaataaacacggttgaattatttccggaagaacaactgtcc ggaatcaaagacg
574	Wesselsbron virus	agtatatctgctgctaatacgttcgacgttagtcctggagtgagcttctattagagtcgttaacacg tttgataatttctactgaaaggagtagaagaaggagattcattccca
575	Equine hepacivirus	acctccgtgctatgcacgggtgcgttgcagcgttttgcgttcgcatgcgctacacgcgtctcaac gaggaggattctccacattaccatgtgtcactcctccctatggagggttccaccgccccacacg gaaataggttaaccatacctatagtagcgggtgagcgggtcctcctagggcccccggcaggtc agggagctgaattcgtgaatcgtgagtagacggaaatcgggctgtaacgtacatagtgacctt cggagcggaatttggcgtgccccacgaaggagggcggggcggtgttggcgccgcccc ctttatcccacggctgataaggatgcttcgaggggcacctgcccgtctctagaccataggac
576	Hepacivirus B	accacaacactccagtttgcacactccgtaggaatgctctggagcaccctccctagcagggc gtgggggatttcccctgcctgtcgcagaagggtggagccaaccaccttagtatgtaggcggcg gactcatgacgctcgtgtagacaagcgcaagcctgacttgatggccctgtagggccttcatg ggttcgtggtgtggtggcgttttaggcagcctccacgcccaccctccagatagagcggcggc actgtagggaagacgggggaccggtcactaccaagacgcagaccttttttagtatcacgct cgggaagtgttgggcaagcccacatatgtgttgggatggttggggttagccatccataccgta ctgctgatagggtccttcgaggggatctgggagtcctgtagaccgtagcac
577	Hepacivirus I	cagggttccgacctggcccggatacctatgccttacgcggaaggtaacgagtaggagtcggg tccccaggcccctaccgcccacgaagccaggtggggaggtatgggagccgggggggtgcagctg gtagctccatgggggacgcccgtgagcggatgctgcatcgataccgggttagctctctgggaga ggggcacttgacaccacgaatccgggaacggacaatcgcggcggtgggacgcgttgcctccg tgcccgagcaatttggcatgcccgtggtgaagagtgtgtggggggggggcccccttcagta ccgtactgctgatagggtcttgcctcaagcccagagagtcgaggctgaaaaccgccaac
578	Hepacivirus J	gccgtcccgaaaggagtcggcgctcatccactccgaggagtggggtggcgctcccggtg tgccggggaaccatgaagcctaaggcctacacatttgaatgaactgaagcttcgttcgtg ccggaaagtctgggtcccatggccagggttcgaggtgggttaaatccggtgggtggtccatc caggatatacggcaggcgggctgtagtcggcggttcggacgagctgtgggtgcctacgggtg attgttcacaggatgggcactccgggtgtgaataggcccgtcagggtgcgtgacgttaaacacag gccttgcctgtgttggggaggattgcagggccacgcgcctcctaaggccggtatggcacagta cttcttggggcgggttgaaggccctccaacgcgacaccagtgctcggcaggcatggggcca ccagctcggcgctccgcacacagacggcgtaggggaaaatcagcaatgtgacccgggtggc atttctctctctacttccatgcatgaaccgcaatc
579	Hepacivirus K	gggaacaatggctcgtccgcgggaacgactctagccatgagtcctagtagtgctgcccacat tagcacaaaaccactgactgagccacacccctccggaaatcctgagtaggacattcgtcgg acgacgcatgagcctccatgcgagaaaattgggtataccacgggtaagggtggccaccacg cgggaatctggggcgtgctcactgactatggtacagcctgatagggtgctgcccgcagcgtcagtg gtatgcggctgttcatggaac
580	Icavirus	cgaagtttaagctaagcaccctcggcggttccggattatgtatcacataaatttgatggctgtca ccacgcaacgcctggagatactcttacttttcttaagatccccggctattgacgctgtaggat gatagggttattttccactataaatacttctactcttgatgttctatatacgaagacgggaggacct ccccgtaccccttagaggtgagatgccaagaacaggcccttctgttctctcgacaaatggcatcata ggcaacaagcatcacaccaagattgctaagtttggtaagagttcttcaagctataggggtgctgtag cgacctctgagtcctgcggatttccacggagcgatccgtgcccacaggggccaaaagccacg gctaacgcccacagagcggcacctacccgtgccccacccctgaacttgaatgttcacactgg cttctctggcttctgaactgtctgctgttggggccccgaaggatgccctggaggtagcccatftta tgggatctgaccaggggacacctcagctcttaagttgctggtgtttaaanaacgttaaggggccc ccaccccttaggtggaggatccaccttcttatttttaaaactctttatggtcacaattgttt
581	Antarctic penguin virus A	ctaggagactacgcagtgggataagatgactatgatgtctgacgggcagaagccagtagcagtcga agtcgagaccgacgtcagggttgcactgtgctgacctagtgccatc

582	Forest pouched giant rat arterivirus	tattggatccgcctccgggcaagggttactttctgtacctcggttagccacaggggtgacccctgtt acgtaggggcccgacgttagcactggtctgacaacaccttctcggaatttcaacttctgcccgtct tccggcggtgtgtcaagaagcagcagtgctcttcttcttctctgagttcaccgagccctacg gggggtagggtg
583	Avisivirus Pf-CHK1	cattccctttccccagccatgggttaaatggccctcaccagggtcggtgctgtctaggtcttcaggt aaagaagtcaaccgagcattgaacaaaacctcagtggtatggttagtaaccccgctcactggac aactttgtctctttaaagtggatcaatccaccccaactcccccttagccacctgagccatgggtg atagcagtgacgaaactaggggaccccaatactctagtccaagagaattccccctcgcgagag gtgctcttggcccgaaaggctagtgtggcagggtgaagtgaaggagctgtagcgtggcaacc ttaagcgttagccgaagctgaccttagaggtaaccctagtgaccactggtatgaagctgtggag gtgggtgtaggaagtgtgccactgtgtgtagatgccagaaggcataaggctgacttggggc cagtgactataccgttccggtaaacctgtgataaaaacatgaagcaagtggggttaaaatttctt aattcttcatttcagtagtgataactggcaga
584	Avian paramyxovirus penguin	accaaaacaggactagataaccacgtgaccgttaactggaaaataagatgtgtaggggcgacc tagttggaattcgaccccgctccgaaacctctaattgtgttattggcagctagtctacttctaaccg
585	Newcastle disease virus	accaaacagagaatctgtgaggtacgataaaaggcgaagaagcaatcgagatcgtaggggtaga agggtgtaacccccgagcgagggccgaagctcgaaacctgagggaaccttctacccat
586	Bat Hp-betacoronavirus	ftaagcttggcttgtgcataggacggaaaggtaactatctacctaactctttagttagacttctta aacgaactttaaaactggtgtgtctctcagtagtctgtatggccattggaggcacaccggtaattac aaatactaagaagattcatagtagatcctgtctagcttttgggtggcagtgagcctacgggttctcc gtgtcgtcacaattatccacacagtaggttctgctcgctgtgtgagttgtagtccgtgtgttct gtcagccatctacaactcgacacc
587	Basella alba endomavirus	ggaaataggctaatcggcttattctaatcaaacgcaaaagacttatgacacagaccggacctgaac gaggtgataaaacacctctggttcagggttcaaacgtagaagattcattctccgattagaaatacaact acgtctaaagcagtagagatggtatcaagatcggttttagaccacagaaatcgccaaatgaaa gtacaagtgggtgtgttaattaccaagaacagtaacattcaagaacaacggcaacccgttgtta ctcatttctgtaattgttagaagtaacaagataaattatttaattgtgggaagaaccttaagtacag taccagccagaagtagtgaattgacagaaatgttatgttcatgtccacgctagaggggccaattgtc aatccaagatcgagatccaaaaataatcaataagctatatacatgtagagggtta
588	Ball python nidovirus	ccccctacccatagcactaggagaacaggataacccctaaccgggcatcctgtgaccttt cagattcgtagttagatattctacagactctgctaggcttctgaccagtccttccaaagtcctgt taccgcccagtagcgttagcgcgaaggggacgggaagtacctccagtaagcgaagctgaa gtaagggaataacggcaagactaactgttagtcttacagtgtggataacctgttagttatcccgga caagagacctgactcgtatgtgaaaacatccaactagggttggtctcaactagctacaggcaggata tccccaacggggctccaagggaatcgagaaccaacctcattccacgtctgtatgaagcaaaaa caggggcatcttcacgcacacctctcaccacagagcacaccaacctctgtgaagccaatttctc gtccaaggacagggtattgagggtcaactttcttcgaccagaagaagggaatttctacaaaaga aaaaccaaatccaccaacaccacaagggtaaaacaacaacttgtgaagccaatacttagtcaaga ctaactattgagggtcaactttcttcaatagagaagggaatttctgtgtaaaacaataacaacaac taacatcagcaact
589	Bat sapelovirus	gacagggtgtttggaggcggtgacgatatctggctggccaccagggaataacggcaaatgtct gatcatacgggtcacaaagtctaccggcgatagtggtcaacaccatgtgtagcagggtatttctgct atgtgaaggcgacagtgc
590	Bat Picornavirus 3	taagcggaaagcattctgtcccccggtcagtaacctataggctgttccacggctgaagggtga acatccgttaccgctcagtaattcgagaacacttagtacgctgtagtccaaattggtatgaccc ggtaacccagaccagaaactgtggtgggtggtcaccattctagtatggcaacatacagggtgt ccccgctgtgtcacaggcccttacgggtgccatttcggatgagctgtgccgaagagctattgag ctactgtgatacctccggccccctgaatcggtctaattcaaccccgagccactgggtgtgaa ccaaccactgtgtgtgtaafgagcaattctgggacggaaccgactacttgggtgtccgtgttt ctttgttcatattaaactgtttatggtcacaacacaacttggtagcattgtgatttactgctcactt gtcacagtaaatatacacaatc
591	Bat Picornavirus 2	gggtttacgaacccgtafacaccagaccttttctccctccccctccaccttcttccccctctt ggaccgaaacaaggacagtaagtggaacggcgaatttatatgtgtgtggccaccacgggaataac

		ggcaattgtctacatgtgggaagtgcacacctccctagccgataacccctgaccgggtgtgtaggat aggaaagtgcccactgtggcgacaggttatggtagtgatgacacccctgacggggcaatggg actgcttgcataccctaatgaagtattgagatttctctgctcattaccgggtgafgggtgtggggg gggcccatacactagatccatactgcctgatagggtcggcggcggcgaccataacctgtatagt cagttgaattcagccaag
592	Bat Picornavirus 1	gaaacccgtatacaccggaccttttctccctccctcctcacttaccttttccctcttggcatgaaa caaggattattcaagtggaaacgcgatttaatatgcggctggccaccgcgggaataacggcaattgt gtatctgctggaaagccaagcctgcttagccgtagcccttgaccgggtgtgtaggatacccagg aaccagcaatacgcgacaggttatgtagtagtagatccctagccaggggcaatgggactgcattg cataccctaataaaccattgagatttctctgctcattaccgggtgtaggtactagagggggcctc tagtactagatctatactgcctgtaggggtcggcggcggcgaccatgacctgtatagtcagttgatt tgagcaat
593	Bat Iflavivirus	acgaatcgggtatacgtctcgggtacattgttgcaagttcgttccctattttcgattgctgcccgaatt tgactcaaaacattgtgacatactatgtctctgttgaagcactacacgagtttgcgccacgcatgta tgtttcaagtcctttgtataagctgcctctatagtggtttcttgaccttaagcctgtcaaccatcct atafctgcatcgagacgttgatgtcaatctgcctctactacgcaaatgtctagtaattagttataagggt ttactatttccctcatttccaattttagttgttagtgtagtatcattcttactccgactgttaagaga aaccaattttatagtcgttaaatatgataaatggaatgaatgtagtgcattttaaaaacactcttctcta taggcgtaagcattctcgtcttagagtcgtaagaaagaaatgcctgtctatcatgtagttatgcga tttatttttgcacgcgctttagtgcaatctagttgacatacagacattgcctaccactcgcgagggg tcgaccggtagtgtaaggagtaagtgatgattccgcttattctgtaccctttgcctggtgaggacag atcctgactaattttaataataaataaacactagcttccaag
594	Bat dicibavirus	gtatagcaccggaaatgtattttactactccaagtatacgtactaggaggttaaacctgttaattacag gggatttagtgactttatccgtaaaagtcgattggacgttaatcggtaacgaggccaagttaccgtg aaccaatttaaaaacgtattttctcatgtggtagaaccaacttggaaatagcatggtcatataggtgttt aggg
595	Betacoronavirus HKU24	gataaagtgtgaatcgttccgtagcatgcaccctcgatctctgttagatctaatactaaactt tataaaaaactaggtccctctgacccctatgcctgagggtttaggcgtgcatactagtgcttagga atttgactgataacacttccctgctaaccggcgtgttgactctcagttcaagcctccaccatagga gggtatc
596	Betacoronavirus England 1	atttaagtgaatagcttggctatctcacttccctcgttctctgcagaactttgatttaacgaacttaa taaaagccctgtgttttagcgattgttgacttctgtgtgggattgtggcatttaatttgcctgtcctc taggcagtggaacatagctcaaacctgggtataattctaattgaatactattttcagttagagcgctgt gtctctgtacgtctcggtcacaatacacgggttctcgggtgcgtggcaattcggggcacatc
597	Boone cardiovirus 1	tacgacgtctgtacattccactactgccaattagctcccttcccggtgctccctctataaggaga gacctctctgcaaggtgaagccttcacccccggcgaaagccgcttggaaataagacagggttattt tctctctcctcggcgcttgcctcttctaagctgaatagggttctatctattcaggcggatggtctgtcc gttcccttctggacagagtggtatctgggttttccggatctcgaccacacactcaccagagctcagg agtgattaaagcaaggcccgatctgcggcgaaaaggaaatgaagtattttgcagctgtagcgacct ctcaaggccagcggaattccccacctgggtgacaggtgcctctggggccaaaagccacgtgtta agcacccttgagagcggtgtgacccaccacctgcaaatatgggatttgacttagtaactaaaaga ttgacttggcatacctcaacctgagcggcgctaaggatgcctgaaggtaccggtgtgaaatcg cttcggcgaccatggtatgacagggggccctgcctggagtggttctatccacacagcgtagggt taaaaaacgttaaccgccccacaagaccggcgaggatgcgggttcttctttaccatttcttg acact
598	Breda virus	atcaectagtacttacaagcgggtcaaacgccctccggaacgggtcataacccctcccgaaact ggccttgacgtgactggctttcagtcctagctttctgagaaatctcgggggtgtaaccacatttt gaccttgggtcagtttggtaacactccaaccaacagcatctgaccacctccagctgtcgcaggc catttggacCaaacgggttcagatacttgggttaaacctctgaccacctccagtttactgcagg ccttttggactaaacgggtAcagactcttgggttagtttttaactaccaactgttagccgccaacc tgattttattgtacaaaatttgggttacacatttttctacgggttggcagtttgggtgttggcaca gttttctgatacctaatttactgtcttttgggttttcggctaaggctgttttccatacttagttgct tgagtaaccttcacaacatctgtttgttttgggttctaaggtaagagtttcaggaaaaacatagg

		cgcccatcttgggtgtctagtttaataatctggcaacaagtatcaagtcactcccttggag tgagacttaagagtagccaattcgctattttggccatccatataaaa
599	Bovine viral diarrhea virus 3	gtatacgcagcttagttcaggtgacgtgtacgattgggtatccaaattaataatttgggttaggga ctaaatccctggcggaagccgaacaggttaaccatacctttagtaggacgagcataatggggg actagtggtgagcagtcctcgatcaccgaagccccgagtaggggtagtcgtaatggtt cgacgcatcaaggaatgctcgagatccatgtggacgaggcggtcccacggtgtatcttaacc caggcggggggccgttgggtgaatagggtgttatacaagccttggggagtacagcctgatagg gtgttcagagacctgctacaccactagtataaaaactctgctgtacatggcac
600	Bovine rhinitis A virus	ttttgggctctgcgcgcgttcgggttttacctgttttcacagagcaaacaggaccttagtttctgtg cttaaacgagatcatgctcgaactagaactataacgctggtcactggaccgtgcccgcgttgcg gatctttgcgggaatggtgctagtggtgctgtggaagtacttaaccacacgcccccaagtgtg ggaaaacacgaactggttagcgacgacgataaggccttggacaccctctccagtgtatggagac ccaagggggccaaagccacgcctgtgcccgtctgttcacaaccacagtcagttcgtgccagta cctgcttttgggaagtgtgctttggacagctgaaaacagtcctagtgggagactaaggatgccag gaggtaccgggaggtacaagtgacactctggtatctgacttgggagagcggtgtctgtttacag acgccactcttfaaaaaacttctatgtctgtcaggcaccggaggccggccttttcttfaaaaca tacacttt
601	Bovine picornavirus isolate TCH6	ttggatctgagcagggggcccccttgggttgccttacaactcaactgggggttaaaaaacgctaac ccgacacgcagaggatctggtttctttattcttttactcaccactgcatgcagattgacgataa acgtgttgtttgtactattgactgctgtctctacgggttactttactgttatacttctgttgttt gggttctactgtactttgtctccttctacattcaca
602	Bovine nidovirus TCH5	caccaatagattagtcgaagctgtctataggcataaactaaccccccaacccattacccccggggcca gggtggcgccgcgttcggggcaaacccgtgctggtataatcaagggttcacagccagattcact ggcggttagctagtggtggcggtagcctggcaaacccgaagaggttggaagggaacttcagg gtagttatcctaggctagcgttagctacagttcgggtcaagataaccgctcgtggtctagggtctagta gagacagtggttaactggacaagggtccagggccactttaggggaataccctacggaaggctagg tccgtaagggaagacccccgcagttgtccggttgagcagagctcgtgctagacaaaaggcaa aaagtggattacattgcctgcaggaaggcaaacgctgttgagtcggagcctaagtactgga cgattgataccacgcctgctgcggtagataaa
603	Bovine hepacivirus	ccatcgacactccaggctcacggalttaagttagggttccgcgaagcgggctaaccaggcccctag taggaggcgctatcccggtgagcccttcccccacggattgagtgagctggagctgggaaggac cgagtacgggccaatcgagaagaacctgatgaacattccaggccctctcggtagatttggatata ccaccagtgaaggcggtgctgtgggtacaggccccctagtcacacagcctgatagggtcctgc cgcaggatccgtgggtgcggctgtacatgtacc
604	Botrytis cinerea mitovirus 4 RdRp	gccccgcgaccttctatttttctaaatagggaaggtccgactagtcggataattcggtttaacg aattatggttagatctattaaagttaaaatagattgaatttcttctatcctctattctatacttggga gtaaatgacaaatgtctatctcctcaaacggaaatggcctaagtgatgaattgaagaaaggtaggtt taagaatataaggcatcaatataatataccttgaatgtaagtaccacggcggtgacgattagggt atcttaggatagacagccatcaacgcgacagcagtggaatcagcttagcatctcaagatcatgta taatatatacataacctfacaattataaaacaaacaaacacactataattttatataaattataga gtatcggaactggagcggtacctactatttaactgatatagccaatgcagggaagctc
605	Botrytis cinerea mitovirus 2 RdRp	ggaaactttcagttccagaagtggttttaagccttcaaaagtftacactttgaacctgcgattaatcc ccatagtgactctgttactgtgaattagggaatagttgttagttcaacttctaatgaggtgaacaataa taactcatcttattaaacctatacgttagacaattgtccaaagagacagttggaaattctgccaatctgga atgtttggtacgcgtagaagataaagaagacctctattccctgectcatgactaagtcattggccc gggtgttaatagagatacttttatataattatacaatc
606	Canine picodistovirus strain 209	cgctctttatacaaatctgtcaaccccttgtataactctaagccgaacaattatagctaggcttttattat ataacattaattaggcattagcgttgctgccaatctctgttaactctaaggataaccttctgttgacta agatgaagcgccctcggttaccgatcccggtgtccacgaagccatgctgtgctggcgctcccc cacctctcccaacttggactccatgttttcagtaggtgtaattgattatattgattctgtcgttcaat gtgtttatcttcaagatctgggacccaacacatgcttactcatgttttaaatgtgttccctcatttga agacaccaaacatagagtgcgagaatgaggatttctacttccattctgttaacagaatgaattc ctgcgtgtgtctgtaaatggaatctttaaagaactcagataaatcgaaacataactaatacaagttg ttttctaccaacatgttaattcggttaactgaccgtggagctgtgaagccctcaaaccccgagttt

525

		ctactgccccaccgcgcaaccttttgccttgcactcgggttgagaggcatggggccccgttcat atccccagtcagttggtagaccttccccccccgctcggtagatgggtccagagggtttgccccgac gacctctatgatgttgccttcttctcctcctcagcgcgagcatgcaactgtcagagccacggaaa cacagcctagcttttgccttctcaccctgcgtacctggcgcccttcgctcagattcgtttgttc gacacctggcggtccccccaccgctacgtgatttactcgtggcatacaccgacctggcgttcagt acattccactgctaatttgggtggcctcctcaatctcccgacccccattgcgacgctcatcac cgccgcccgaacgcgacccggcggttctcactggcactgtccctcgtccgccccgggttcca ctcatggttggcttttacttattctggttactggtgttcacactactcatgatggctgccaatgacat gac
615	Chicken orivirus 1	aaacctcacgagtgcttggtaggtccaggccaatattctcgttaaggcttgggttccaattttcca ccactcgtgttgggttgcctatgttaccagaggggcggttgggggaatttaactccccctcc cctgtggtcctataccaccacacacctgttgggttcttcttactatcttctgttttcgacttttaaca ctaggcagggcgccctagtcatacaccgccccggcgtgcttcttccagctcgttggggcggttgcgcg ctggtccatcgtgccagcgacatagcacttctggacacctccgaacgccccctcctgtatggg gtggtgccaggggtttcagtggtgacacactccctggggccccgaaggctagtgtgcaacag gtgaggtacagccagctcccccggtgctcggaggacccaagcttgtgaagcacacctcaccttct tgggggtgggctagtaaggtgtaagcatagtgtccgtgtcgtggccaacactttgggtcaagt ccagccactcagtgagtagatgccaggaggtagccctagtggatctgacttggggcctgttactt aatgcagggttaaaaactatgaagctgagtagttagccggcgtggtgttcttcttcttattcattc tatttt
616	Chicken gallivirus 1	gggttaactgtttaaccaaggcttcggtgcagggttgcacgttaggagttcatgttgattcatgctcc aatgccaaaactttgtgtgttgtttatgtctttccaaagtftccccaaggttccggtactcaaacct taattcctagtcccatcttttgggctttagtatctaggaaatgtaccgtgccttgacgaacgttaagaa agctgtctttatgaacggttctaataaagataactggctcgcgccacctggtgtgtgcccgtt ggaaatcccccatggttaacatgggtccaacgggccccgaaggctagtggcaatcgtcctcaca ggaaagggttcccaccggacctgaacaggattgatgaagctcactcccagggtcctaacccc aagggaagttacttatagtaattagaatttagtatgaattgtctggcaatcttctagtagtcaggaaacg ttatgaccaaafgagtagaccccagaaggtagcccaattatagggatctgacttgggcccatact gtgtgtctccccacatagagggttaaaacatgaaggtttgggtccaaaataattcttcttattctttct ttagtgtgacgccattatatacagcagtttgcgtg
617	Chicken calicivirus	gggtcatcatcactgaacacctcggcgagagatgcaagggtgggaagtcactcctgccccctgg caacatgcagggtgcccgatcccaagcttagttctgacttccctcctcgtgggtggtgcaacactccaag gttgatgaacaaacctggagggaacctctgggcaacctgtctctcagagatctccggcgcatctcc acagactacctatgctcccggaacctgagaagaacactgatcctatgatggctggctgatggctc ggcgagatgtgacctttgcccgtctcattggctgggagg
618	Carp picornavirus 1	agctacaggaaagagagataatcacagcacataatacaactacagaagagagcctttccctg agcactatftacagcaaaccacgctgggaaaagtggtagcatgaccacttacgggtacttagtat aggattttaatactcgtgattcttttcttttcttaactaagtctgcccggactttatccgtgtttatftta gtttaagcaaaatgagtaactaagttttaacctgccaaatgggtgagaagtaactctgtgaaaatacc atttgtcatgaaattgtctgaaaactcttaggcttttgggggtccactcgtgtttgaggactatt gacagactctattgttagtgtagtgaactaatgataacgatttgcgtattacgcaatgggctgtacc cgtagatttagtatgccggggggagggtccactggattgcactatgtaacctgacaggggcgtc tgccgacgcactacaatgaggataagatcggctgttttattt
619	Falcon picornavirus	cgcttgggaataagttgagaggaaattatgcatgctagtgtgtttgtttacaaataattgttcaatecaa gtgaagctcttcgcttggggcgccagacactgcccgaattcttctaccgtccctccacacctgtg gatgaaggggccggatgtgtggcctctggctaaccctctctctgggggtgagctactggatgtttta ctcctagaccaaatcacatgaactccttctgacactcgggtgggctatgagcctgaggtaata gctggcgacagctacccaggggccaaaagccacggtgttagcagcaccctcatagtctgatgc ccaagggtgatgttgggagctagtagtgtgtctggcctatgtttaggacttctggccaagcgca gaggagtggggtgaaaggtacccgaagggtacccgtaggtaaccttaagagactatggatctg atctggggccccctacgttggctttaccacgtgttgggggttaaaaacgtctaggcccaccagc ccacgggagtggttcttcttaaaaagcccaacaattattatggtgacaattcactgtttcttcttgc aattttgtattcttggactccttaattttgttgcctgaatttagtgacattcagattcaatafcaaa

[illegible]

		gtgctcatcacgftaacagagtaacctgtgcaatagatgagcttgtttaaaccgtcttgatacgcac acc
627	Guereza hepacivirus	cactcaatactacactcgcgcafttggggagaagcgctggcgttcgcggaaaccgcgftaacatac gcgtagtacgagtgcgacagaccccggtgctactgggtgtagcgagacacgagccgaagtctgt ggggggaaactccacttagaggcgatgcccggcgtaggcttctgagttgggatgggcccact tgcccctgagtgggggggtgttacgacctgatagggtgcgggctggcgctaccactttccagt cgtacatgagtc
628	Grapevine associated navavirus-1	gccccggggggcagtcctgtgaaagggtctgcaccatactatatagttatattgattacatccaaa aggcgacttcgttcaggttttaaatctgacgtaggtccagtaataagcatgtcaaaacatgtaagtt atcctgtaatctactctcataagatgagataagatgatatgtcagttcccatgtaaataaactcattatg aattcattcatataaggtagaagtgtaactatggtgaaacattatataaaacggcattttgcatga acgtcattaaaggaaactggcataccaatgtctatttagtgactatgatatattagagtatccctattataat taacaattattccttttagcatacatccgacacaaattttaaaagaagaataattactcattaaaa
629	Goat torovirus	gtacttacaagcggggttaaacccgctcggaaacggttacaccccctccgaacgtgcgcttga cgtgactggttctcagtcgtcttctgagaatactccagggtgtatccaccacttgcacctctgg tcatttggtaaacaccataaccaaaactctacacacctaacccactccagcttgcgtcagggcc tttggactaaacgggttgggtgtttgtgaccaactcgtctacccactccagtttactgcaggcct tttggactaaacggCtttagacttttgggttagttttaactaccactgttttagccgaacctgtatt ttcattgttgaataatttgggtttatatactatttcataggttggcagttgtttgtttgacagttt ttgtgataccaattttactgtgcttttagtgattctgtaaggctgtattttacatacttagtttgggtga agcaattttacaacatttatattgttttattcgaaggtaagagcttaggaaacaccatagacgcc attctgtgtgtctagttccaactaatctggcaaacagtaaccaagtcattgactcactttggagtg gacttacgagtagcaatttgcctatttgcggacatccatataaag
630	Foot-and-mouth disease virus O isolate	acaagcttgacacgcctgtcccggtttaaagggaagtaaccacaagcttacaacgcctacc cggtgttaatgggatgtaaccacaagatacacccttaccgggaagtaaacggcaatttcacaca gttttggccgtttttcatgagaaacgggacgtctgcgcacgaaacgcgcgttcgtgaggaggac ttgtacaaacacagatcaaacagggtttcccaactgacatacaccgtgcaattgaaacccgcctg gtctttccagggtctagagggttaaacattttagtctgtcttgaactccacgctcgttccactgagcag gttagtaacagcactgtgtctctgtagcggagcagtggtggcgtgggaactcctccttgtaacaa ggacccacggggcgaaagccacgtctctgacggacccacatgtgtgcaacccacgacggc aacttttctgtgaaactcacttaagggtgacactgatactggtattcaagtactggtgacaggctaag gatgcccttcagggtacccgaggtaacacgcgacactcgggatctgagaaggggactggggtt ctgtaaaagcgccagtttaaaagcttctatgcttgataggtgacggaggcgccgctttcc attataactactgacttt
631	Feline infectious peritonitis virus	actttaaagtaagtgagtgtagcgtggtatataactcttctttactttaactagccttgtgtagatttg tcttcggacaccaactcgaaactaaacgaataattgtctctctatgaaacatagaaagacagcggtg attatttcaccagtttggcaatcactcctaggaacgggggttagagaacggcgacacagggttccg tcctgtttggtaagtcgtctagattagctgcggcggttcggccgctgtagttgggttagaccgggt tcctgctgtgactcctcgcggcgccaggaga
632	Farmington virus	acgacgcataagcagagaacataagagactatgttcatagtcaccctgtattcattattgactttat gacctattattagacccttcacgggtaaatccttctccttgacgttctcgccaagtacctccaagta gaacg
633	Avian infectious bronchitis virus	acttaagatagataatataatatactattacactagccttgcgctagatttttaacttaacaaacggac ttaaatactacagctgtgctcctcataggtgttcattgcagtgcactttagtgccctggatggcacctg gccacctgtcaggtttttgttaaaacttattgttgcgtatcactgcttggcgtgtctcacttt atacatctgttgcgtggctactagtgctccagcgtcctacggcgctgctggttgcgagtgca ggaacctctgttcatctagcggtagcggtgtgtggaagtagcacttcagacgtaccgggtctgt tgttgaaatacgggtcaccctccccacataccttaagggttttgagcctagcgttgggtctac gttctcgataaggtcggctatacgacgtttgtaggggtagtgccaaacacccctgaggtgaca gggtctgtgtgttttagtgagcagacatacataatagacagtgacaac
634	Human rhinovirus 1	ttaaaactgggtgtggtgttcccacccacaccaccaatgggtgtgtactctgtttattccggtaac ttgtacgccagttttccctccctcccatccttttacgtaacttagaagttttaatacaagaccaat agtaggcaactctccagggtgttaaggtaagcactctgtttcccggttgatgttatgtctcca acagggcaaaaacaacagataccgttatccgcaagtgctacacagacgttagtaggttctga

		aagatcttttggttgctgctacagctgcataccagcagtagaccttcagatgaggctggacattcc ccacttgtaacagctgggtccagccctgctgctcctgcgcacacctctcatgagctgtgaagccaaa ctctggagcagggtgtgaagagccgctgtgctcacttttagtgcctcggccccctgaatgcggcta accttaaacctgcagccatggctcataagccaaatgagtttatggctgaacagagtaattgcgggafg ggaccgactactttgggtgctcctgtttcacttttctttatattgcttatggtgacaatatatattg atatatattggcatc
635	EV22	ccftataaccgactgctgagcttctataggaaaaaacctttccagccttgggggtgctggtcaa taaaaacccccatagtaaccaacacctaagacaattgatcaacctatgctggtgccactatc gaaggcaacttgcaataagaagagtggacaaggatgcttaagcatagtgtaaattatctttcta acctgtattatgtacagggtggcagatggcgtgccataaatctattagtgggataaccacgctgtgg acctatgcccacacagccatcctctagtaagtgttgaataatgtctggtgagatgtgggaacttattg aaacaacaatttgcftaafagcatcctagtgccagcgggaacaacatctgtgaacagatgcctctggg gccaanaagccaaggttgacagaccattaggtgttcaaaacctgaattgtgtggaagatatt cagtacctatcaatctgtgtgtgcaaacactagtgtgaaggccacgaaggatgccagaag gtaccgcgaggtacaagagacactgtgtgatctgtggggccaactaccttatcaggtgagtt agttaaaaaacgtctagtggccaaaccaggggggatccctgtgttctttattgttaatttgaaca tt
636	Human TMEV-like coronavirus	tccgacgtggttggaaattaacatctttccgacgaagtgctattatgctctcccgattgtgtgctt tctgcccctgctggcgagcgtctcctgggttgagaaaccttgaactcttttcccttggagccttggctc ccccggtctaagccgttggaataatgacagggtatttttccaaactctttatttctactttcatgggttct atccatgaaaaagggtatgtgttgccttctcttcttggagaatctgcgcggcggtcttccgtctc aacaggcgtgtggtgcaacatgcgggaacggtgaagaaaacagtttctgtgaaattagatg gacatcgaacacagctgtagcgacctacagtagcagcggatcccccttggcgacaagagcctc tgccggccaaaagccccgtggataagatccactgctgtgagcgggtgcaaccccgacacctgtgtc gatggccattctctatggaaacagaaaatggttttctcaagccctccggtagagaagccaagaatgt cctgaaggtacccccgcgcgggatctgatcaggagaccaattggcagtgctttacgctgccactt tggtttaaaaactgtcacagcttctccaaaccaagtggttgggttttccaatttgtgtgactgacaat
637	Human coronavirus 229E	acftaagtaccttatctatctacagatagaaaagtgtcttttagactttgtgtctactttctcaactaac gaaatttttgcctatggccggcatcttgtatgctggagtcgtatgtaatgaaatttcatttgggtgca acagtttgggaagcaagtgtctgtgtctcctagtctaaagggttctgtgtccgtcacgagattccattcta caaacgccttactcagaggttccgtctcgtgtttgtgtggaagcaagttctgtcttttgggaaccagt aactgttctcta
638	Hubei zhaovirus-like virus 1	gtgcaggatggcctttccactcttaagtgtttgtaggatttcgtgggtccatacccccgatttctg gtacgtattccatgcacggagaaacagacaaaactctatttcaaaaaatattatttttactctgtgg gctgagtgccgaccaccagttccagcttagcaacctgggaagttgttggagattatggaaccaaatt acacatgcgtggaggtgcgcgccactccgtatctgttctactcattacgcgattaaagtcttcgcacgaga cgaggcgaa
639	Hubei tombus-like virus 9	ggaccatccaggcagggtgtaggctagtaccctcacctgacctgtcgcgatgttggcttgtgagg cttgtgggaggatcccttggcccatgcatgtctgtcatcgtgaaaaatgtgtgtatgctgcacct ggcgtggaggaaaacggcatttgacggatgctaaaggagattgaaggtctcaactccatgcctt attacaagtcccccttctcaagcttcgagcccacgtctgtgacgagcagaggtgaggttgagtc aagaggaggtctattcaagctgttcgcagcaagaaccataaaaacttcttctcaatgggcaaga gcccgtaaaggctcgttctcgtttgcagacagtgtagcggagccctgtaaatgtcgcggccctt catcgggtggttgcacggagtttaagaaaattggcatgaactgtctccaggtttctgtacgtgatcgat gagggtgttgaaacttagtttgaaaccaacgtatgagcgtatgggtgtcgaanaactaagaagcaattcc gtcatcgggcacgtatggaafactacaacgagaaaaaatgcttgaanaagatccactagggaacgc
640	Hubei tombus-like virus 32	ttcgggtttaccgcgtaagcggccacactgactggtgtcgggtgttgaaattgtataccagatgagg agacgttaccacgctctcggcagtgctacgtctgggaaaggactgtgatagtggacagttctacc gtcttttagatacattgcgtgtgtatagccccgggagggaattaaactaatagcaacgcaatgcacacgg cgggttcggatttgcctgactgaggaagacatctaagttctaattgaacc
641	Hubei sobemo-like virus 3	gagatgtttgcgtggccgttgcgtgcggggcgccacctccctacggggacggtgagacacc gctgggggaagccccacccccggccaaggggatcctgcgagaggcgaggagaagaagggcc cagccctctggggcgcttttaggggTgccttgggagggaagtagccagcggggcgccgctc ggatccgctgtgtcatttgcaggctaaccttaaggaaagccctgagctccctcgcgttgcggaa

		agggaagacgaaggcacttatcaaggcttfcagggaagcaggagcgcaattacagcgcgccgtg gcggccctgggtaaccgcatttggccg
642	Hubei picorna- like virus 2	agcaactcttctgaaaactagctagagttcgcacgaictctctggctaafgacaataaccaatcaa aaagtcaaatgttcagtataatataatatttagtagtgacctttatttagaaaaacttttagatgatttatcgtc aagttgccctttgtgaagcgatcagcttttatacgtttcattttgtatagcttaattgacgttgaag tacgttttgcatacctcacattgaggatagatcgtttcctgactaagaagttaaactagctaccata gcaaccatagatagatagctttgttttaacaaggtttaactgacatctatgctcttgccttgcgtgtttta tgtatagaaaaGttttataaaaaactacataattgtcaaaagaaaaagcgtagctactgacgcattta tgttcacagtgtagacacaaacctctatattttgattgtaaaataggtttgcctgaccatttatcaata caaaactgttcaaacgcctctccgagccatattggccgacgcgaatcgacataacagggtagtt acagctgcagttgcagccgaggatccactattttgagagagaattactttaacgaaatttgaagtc actctacacagttgaagctgagtagggcgttatccaaacgtaaa
643	Hepacivirus P	acatgggggggggctgacagtgtagtactgtgccaagcaggtgctacgctatgcttaggtgtc gctgtaggccaagacatgtcccagtcacccaggtgaggggggggttccctcaccgctgcc actgctgtaggggtcctgcccggagggtctcgggtgctccggctgtac
644	Harrier picornavirus 1	gatgtgtgacgggtgaattactttccggatcccactttctattataacttttccccaaggttaggg aaagaacctgggctcggtaccacacagccctccgacgctagtggaactctccggagataacggt acccctttgttagtcacgtgtgtgtgaagaaccacctagattgtcttgggtgctgccccttagct tccatttcttggagcactgtgcaatgagggttcccaactgtgtaacaagtgccctaggtcccga ggatactgtgggtgtgtgtgaccgcaggagctgtctccacggctcctctaatgttacgctat ccacaggccagtgctgtcatcgatcccgatgacagagctagtattgcgaacccccaaagtaaga aaagtggctagtaacctgatactgtgtgaagagggtgggtcagttgagtagatgccctagaggta cccgaaaggatctgactaggggaccgtgactatacattaggtaaacgggtatataaaacccatgaa aaactgaccactttctttaacctctcttctttttatgtgtgtaagtggtttgagttggactgtaccttg cccgctttcttggattttctctatcgtctcttctacacctactgttatcaaggcactcttttagagata
645	Kunsagivirus 1	tttcaaatcgactccggtagtataccggagcccgggttggacgcttggccgggttaacagccc cccaccccttcccactgactgatttctcggttgactcatttgcaactctgatttctggatttc ccggttatgtctgctggcggaagtcacgtactctgacgttgatctatggcgtttgttccagg ggggagggtggcggcagaaacgccccgcgtaaaactcggcgccgacgcctgtcaagccact ccctggggccgagcgccgtgaggtgatacagagagataagcacactggcgctgacaacgccc ggacctcagtgagaagagcagtagggccgtgtttatgggactccattggatataccccgctgtcg gaactcacggctactccgggttgggaagcccgcgactgtgtactgtactgggtgatagccgtgtc cttccctctactgtgtatgaaggctgaaaaacccct
646	Kagoshima-2-24- KoV	tatagtttgcctgttctgcaccgttaccgctcgttgggaatgtgaaactggcaccctctctccc ctaccaccccttctctctgccccattcataatttacaacgcgcacacagcgccggccgccaag ggctagcctggcggtataaaaggaaactgggtcttctcttcaagcctaaaggtaggttccctg tgtccctgaatgtcctggtaggaatgctgcaccgtaacgctttgtgaagtgttgcaggtctggccc ggcaagcctacagagtgctgtgatccgctcgggacgccaatcctgtaacaggacccccagtg cgcaacagtagttacagactcgggtttgtcacttgccttcatggaccattgcgcgaagtgcgtgc ccatatecctgtacttcagggtgtcttctctggaccctagggaatgctgcgaaggtaccctgttggc gggatctgacgcaggctaatgtctatgggtcagtttcttttcttactccacaattgactgttaa ctgactctggatctgtgttccactgcttctctctctcaaacggcttacttaccactcaccct tcgaccaacacacatttacacactaactttttcgaactcttctgactcctggttggtaagac
647	Kashmir bee virus	tacgtacaattttgacgcttctgctcatgcaacaatgaactacatgtggcgctcggtagtaaccagag ggcgctcattccccctatggagtgagataataagctaccgactcgagctgtagaataattcagc aacttataacgaacacgaatttttagtcgacgaacacattttagcctttaattctatgatttataatg ataaacagctcatgtaactgtctaaactacataatacaactggattacgaaccttaagtaactatctt agatgaagtctagtagtctcccaatatttccgtgaaaagaatgaggacgagatagctattttaa gacgtgaggctttaaattctgataaatacattacctgagaattctctctttaggagttagaattgaaaa gatttagctatacttaataagaatttaaatatgtaataatgctgaagacaagtatttcgatttataacct ctataattctataaagtatctgtgagctcagtggaatcacagtaaggctgcagttaaactgttaac tttcatctcgtgtcggagcagtggaatggagccggacgatttcgccccaac
648	Jingmen picorna- like virus	tgtgtttttgcaagataattgttctgtgataaacagtgattgtgtgtgtaafgcgacgcagtc gtagttttgatgaagtgtatgagagagtggaacactatctcataagaagattgaagagtgtaga

		tcaggctattgacgagcttctaagcttcgtgattacaagcttaattgttcacaatgctcccgacggg aatcatctgacccctcttatttcgccccattcgtgttatcgcttgggtatctaagtttgctgttga gcagcatcacagtttgccttcagttgagctctgaagtgtctgtctgtct
649	Mumps virus	accaaggggaaaatgaagatgggatgttggtagaacaaatagttaagaacagtaagcccgga agtgggttttgcgatttcgagggcggtcgcgaccccttcattgtcgataggggacatttgac actacctggaaa
650	Mouse Mosavirus	gaagtgtatcatgaacttggttattggtggaacgcacatgaactcccaacaatgatctgaagacac agcgtggttaacaattaccatgcccctgtggctgcccagacattgatggctattgggtgatttatgat gac
651	Miniopterus schreibersii picornavirus 1	gtgtgacacccgttgccttggttaagcatgaggtggattccccgattatgtctacccgttactatggc gggcggctcgttcagggtgttctactgaggaactgcaccaagtcttattcttattctcagatctccgg ctgtttgacgctttaggacagcaggactatttctttaaattcttctaccactagggtctctatccta gtggaggggaggtgcccccttctctctttagagagtgcgctggcggtcttccgtctctgga aaggagcacatggcatgctacaattggcacaagaaaacaagcttgcggaattcttctttagtaga ggaagctgtagcgacctgtatggcgagcgactccccctctcggcgacgagagccctctcgggcc aaaagccaagtgttaatagcacccatacaggcgcgagtaaccactgcccttctcaacatacaat gactgatgaaccactgttgggttttctgacacctttagtaggattccaaaggaatgctctgaaggtac cctgttagcttacgcgcaggatctgatcaggagctcttctacagtgtgtacactgtgcaagggatta aaaattgttggaggaatccccgagatagtggtctatctcttcttatttcttctacagacacg
652	Linda virus	gtatagcagcagtagctcaaggctgtctatacattggacataccaaattccaattggtgttagggac cacctagggtgaaggccgacgacaggtagccattcctgttagtaggacgaaccgttatgttggtgact gggtgtcaggtgagcaggctgcaatgcgttaagtgtgtagtacaccacagccgtcaagggtgcc actgtaagatcaccactggcgatgcttggtagggggcggtgcccacgcgaatgttagcg gtggcgggcgctccatcgtgaaagctaggcttggatggacctgttgcctgtacagtctgatagg atcccgccggaatgcccgtgacagccagtataaagaataaccgtgtgattgcac
653	Lesavirus 2	tccttcttatttcttatgttaactcttctttaaagtttatttgcctactgtgagcttatgcgggaccactg tcttagacaaccccacattgtcatgtagtaagtacacgcaaccattacgattacttttaaccgctga ccttttgataaacaactgaagttaggcgtgaaacatgcattataccaaagtagccccgatttccca ctacgggtgggggggtacccctactggcttgggaactgtagecattatgtgttgcctgtgcttcaggat ctcacacacaacagttctctcacaatgggaatatgggtgagattgcagtgacatgaacaagtatcta gtagtacatagactcaagcctagttgcctgcgggaacaacatgtggttaacacatgccccaggggtcca aaagacaagggttaacagcccccttctaggtgtctgtgtgtgaagaatactttagtagtgtttatgat ctcaactgttagtacagaatgagtatgcttgggtgaaggatgtctacaggtaeccattatatgatac tgagtaggagaccactagtgttggcttaccgccaggtagtggtttaaagcgcttagccaagc caacagcactagggtatagcttctctatatttataatttcagtgatatagggtgacaa
654	Lesavirus 1	gtaactaataagcaagtttactgcctgcaaaactgctcaatgggaccaccgcttcggcgacccctt gttaggtttgtatgttttaagtaactttgcaaccatacgaattatttagccgctctctataatgatctgt tatagttgggacgtgaacaattgttttctcacacacgtccggtcaccggcggtgtgttctccgta agtcctatccacataaccatggtggtagccagcatgttgcacaggctgtgacagttgggtggg cttccacctctcaacaacacactgaattgcaatgcactacggaggaaatgacaatttgggtatagt tttgaactgtgctagtaatttttcaattaaagccatgttgcctgcggaatcacatgtggttaacacatg ctcaggggcccaaaagcacgggttagcagccccctcatgtgtgttagaagtgaanaacacatag tatgagctataatactttgtgtcttctctgtagtgtacccgccaagtgaaggatgccagcaggt accattttatggatctgagctggggattgatgttatctataaafgcactgatcaatttaaaaagcg tctaagtttggcacaacactggggacaggtgttttcttattcttatttattgatta
655	Phopivirus strain NewEngland	gggagtaaacctcaccacgttgcggtgtttacggctacctattttggatgtaaatatttactctg cagggtcaggtctctgaattatgtccacgctagtggcactctcttaccataagtgcagccttagcg gaacctttctacactgtatgttgggttaggggtacattatttccctgggcttcttggcccttttccctg cactatcattcttctccgggtctcagcatgccaatgttccgaccggtgcgcccgcggggttaa ctccatggttagcatgtgagctgtagccctaaaagtgtgacactggaaactggactattgaagcat acactgttaactgaacatgtaactccaatcgatcttctacaaggggtaggctacgggtgaacccc ttagggttaatactatattgagagatacttctgatagggttaaggttctggataatgttgagtttaacga caaaaacattcaacagctgttgggccaacctatcaggtagatgcttttggagccaagtgctgtagg

		gggtgtgtgtgaaatgcttcagtggaagggtgccctccgaaaggctgtaggggtaatcaggggca gttaggtttccacaattacaatttga
656	Pestivirus strain Aydin	gtatacgagattagctcatactcgtgtacaaatggacgtagcaaatftaaaaaattcgatagggtcc ccatccagcgacggccgaacggggttaaccatacctctagtaggactagcagacggatggacta ggcacagtggtagctccctggagggtctaagtctgagtagacagaacagtcgtcagtagttcaacg ctggttaacccagccttgagatgctacgtggacgagggtcatgcccaagacacaccttaacctgg acgggggtgtccaggtgaaagtacccatcttgggtgctgggtagtacagcctgataggggcgctg cagaggccactgacaggctagtataaaaatctctgctgtacatggaaac
657	Quail picornavirus QPVI	tttgcacagttcgtccctccctcaccataccttttccctcttttaggactgafacttggttatgatga gcagaggatttcgaagttatgcttctgataaaaagtaattcacgaatcatgggattatagccttga agtgaacactcatgtggcaagtggttagtagctctccatgttccatgtgcagtggtgactgacaaca gtgagttcggggtgtgtagtaaaagggaagtattacttaccgcacctgtatagctgtgtacgtta ggatacagagtttagtgcttagcaactttaactgtgtctgaatatgtcaaggctcactgaagtgtt gaacgcgaacgtccgccactgccatgtatagcgtgcaatgcataaatgggtcactacatgatacg agggaatgggaacccctccatggccgaatgcagggtgacagcctgcggcggtgactgctgtgtt agtataatccgttggttgccac
658	Porcine sapelovirus 1	acactcattccccctccacccttaagggtgtgtatccctacaccctaccctccctccacatagg acgaataaacggacttgagattaggaagcaagtagataaggtatgggttttggatcaccttaaatggca gtagcgtggcgagctatggaaaaatcgcaattgtcgatagccatgttagcgacgcgttcggcggtg ctccttgggtgattcggcgactgggttacaggagagtagacagtgagctatggcgaacccctacag tattacttaggggaatgtgcaattgagacttgacgagcgtctctttagatgtggcgcatgctcttgg cattaccatagtgagcttcaggttgggaacccctggactgggtctatactgcctgatagggtgcgcg gtggcgccgtgtaactagtatagtcagttgaaaaccccc
659	Porcine reproductive and respiratory syndrome virus 2	atgacgtataggtgttggctctatgcttggcattgttattgtcaggagctgtgaccattggcacagc ccaaaactgtgtgcacagaacacccctctgtgatagccttccctcaggggagcttaggttgtccct agcacctgtcttcggagttgcactgctttacggctctccacccctttaacc
660	Porcine enterovirus 9	Gaaccttagaagttacacaaacaaagaccaataggagtcacacccagttggattgcggtcaa gcactctgtttcccgacactagtagtgataggctgtaccacggcggaagatgaacccgtccgtt atccggctacactacttcgggaagcctagttaacattctgaagctctgagggcgttctcgtcagcacga ccccgggttagatcgggctgaggtgtctccgcataccccacggcgacccgtggcgaggccgc gttggcgcccgccatgtgcgaagccataggacgcctcttagatgacaggggtgtgaagacct actgagctgggttagtgctcctccggccctgaatcggtctaatectaaccagagcgtccacca gcaatccagctggcaggcgctgtaacgggcaactctgtggcggaaccgactacttgggtgtcc gtgttctttgatcctattttggctgtatgtgtgacaacgataagttgtatcataaagcttgggtt ggccacctggaaaaagttatcagttgtgafattgttcggctctcacgcctaccaataagacaagcc ctataattacttgtgcatfttactgtcagaagaaatcacagagtatcatttggatttgttactcacatta aggacaag
661	Pigeon picornavirus B	tttatttagctgttaagttttattgtgccgagAccecatagtaggafcttgggttccacattaaagctt Cccgaccacacatccaaacgataggcggtgttaagggtccctggctaagtgtttactcattgctag ggaagtgttgcgacccgttacagtaggaatacaggaggtcttagtgcttaaccagataaagtgg tgctgaaatattgcaagctaatgtctggcgaaacgacggactaccgtgaactattgttaacgcccg cgtgttaggcaacacacgggttagtaggtcacttacattgacatccgtgccgggaaagcggtatct gagctatcgattgctgatagggtgccggcgggcgcggtacgtgtgtgtatagtcctgtctcttggg gtatggcgtctactcgggtttgtgtgtgttggaaatgtccatgtcggagatgtgttaccgggtgtt cttacttgtgcacgaataagaaaacagagtttggagtttggaaagttaattggataacatcaagtttc attgcgttctctgtgcattttaggccagaaaatcgaccacaaaatcgaga
662	Picornavirus HK21	aaacgggaggatcggcttggcttctcttaaatagctacaaaactgggctgactggtggggga gctactagatccgggagaggactagttaccccgcgtaacttccctttgtcactccctccactacc ccttccctgtcccttagactcttaagtaggtgtgcaggcctggccccccggaatgggcaaatcgg acgtttgcgttagggctgtgtgaaagggtggggccactccaccgtagtaggacttctgtttcac gtttaaacctctccggcgaggatcgtctagacactaggctgtataggatggcacgacttaggttcc gcaccttactagtcgcgaagatccgcttttgagctcaagtacatgggttcttgaagatgtctaacc agaggagggtgtgtgaaatattgcaagccactctggcgaaactgactgcttgcctggaacag

		gctctctagcgccctccactgtagcgcggtggtgggttagtaggatacctatatggacaggggat gcgggaataccctcactagctagtgactggttgatcgactggcgcggaatccagtgatacttgca taatccgcagactggggag
663	Picornavirales Tottori-HG1	agccatgaactagtgtgcgattccacagtggtgctgaagccccgactgtctactcgcatattt gactccagtccttcgcagccagcggttaggttctggttaaagtgcattatgtgcacggcgccacg cagaactgccagaaatggtgtaagctgcgcccacgccaacggtttgtgtatcccgtagtcacagt ttacagctgttcgttaacagtaggggtttgtcgacccggacccttaatgcgcgcgaattcaacgc gcctagtcggcaatggtatcatttaaccatgcactacgggagaaattgagaccaagaattct gagggccactgctgtcttaagtgaatgcctcgaggagacttctgtaggagcctagcggtttca accgcgacagctaactcctgcgggatgttgggttcatactactggtcctcacaacgctaagt ggatgtgtccacaggtaggcaaacaccgagccccacattcaggagacctgtatgaacgatcctat cagcattagagttggaattgggtgtgctaactccgcataagtgcacccggtgtaacgtgggaa actatccagcgcacgtactgtcctcaatgtctagggaaggaccgccctaagcgtacaacggggc catgtgtcgagc
664	Rodent hepatovirus	ttcaagagggtccgggatttctggtcccccttttgggcacccttggctcggggtgtgaata ccgtgctgcgttttgcgtgtaacggcttcatfagtgtttgtctgttttattatgttGgtttgtct gtttgtatgttggtattgttcgtgttaattgtatgaccacattacactccagccaatgaagaacagt gtgcggttattgtctgcggaattcctaactgcctggatccgttgtacgcatacaaaacaatttga gagagagtgtgtaaacggctgggaatccctgagtacagggaatcacactgatagctcatcttg gctgtttcagtcattgacctaagcagtgtaatttgggtgtacccccatagcttaggaggaattt tgtcttggcactagagtgaggacgtgatgcctccgtgtctaggtatggtcgaaggacagaatggg tgctctgtatgccatactactgtataggggtcctcacggcctctgcatcttagtgagaatttcaatt t
665	Rinderpest virus	acaaaacaaagtgggtaaggatcgggtctatcaatgattatgatttagcacattaggattcaagatc ctatcgactggagcaggcttaaggtaagggttctttaa
666	Rabovirus A	ctacggatatttgcagaccogcttctatogcccaacaatcccccttgaaccacaagcttactca ggctagcagcccgactagctgttgaagaaaaggctagggcacacaccaaacaccgaccc cactggtcgaaagccgcttggcaataagactggtggaacagggtcgccgttagttgttgaacat tctttctaagcttctgtagcggtgctactcacaccgtaactcttctaccctatccccacgcttggga actaggagggtgagtgattcaagttaagtactgtcagaatgtgaaaatgatctgattctgaaacg ctatggatccatcgaagatggggctacacgctgcggaaacacacatgttaacatgtgccccag gggcccgaagccacggtgtagatgacccgtgtgtgttagatcataatgattcatagctagt aagatgattgaaatcaactgagctgatggctaactgtgtcttattgcggcctaaggtatgctgtc aggtaaccttagataaccttaagagactattgatctgagcaggagccaaagtgtgtcttccagctt ggtaaaaaacgtctaagccgaggcagggggcgaggcccccttctcccaaaacttaatt gattgt
667	Shingleback nidovirus 1	ctgtgagtaccgacaggtcgaagtctattataggcgctcgaacagaaaacctgtaaactccg gtttcatctactgctcgtaagaggcagaagaggacgaccaggtgtccagatcactgttatct gtttcagtcaggaaagtaacttttcgacgaagtgcgaccattcatcaccgctgaaaagcgtagaa gtcgatgaagatgagctcacaagagccatcgccgatttcgacaacaaagtactgttccactcct tcgagctcgtgaatttcgagctgaacaacaataacgagaaacgagtggtacggttattataatta cgacaacaaaactgcaaagttcagttgccagtcacatgtcgaatcgaggacgtaacctgggac aggtttacgtg
668	Seneca valley virus	tttgaatgggggctgggcccgtgatgccagtccttcttcccccttcgggggggttaaccggctg tgtttgctagaggcacagaggggcaacatccaacctgttttgcggggaaacggtgcggtccgatt cctgcgtcgcaaaagggtgttagcgacccaaacggcgacctaccaatgtattgtgtgtgtgtc gagttctagcctactgtttctccccgaccattcactcaccacgaaaagtgtgtgtaaccataag atttaacccccgacgggatgtgcgataaccgtaagactggctcaagcgcggaagcgctgtaac cacatgctgttagtccctttatgtctgcaagatggctacccacctcggatcactgaactggagctg acctccttagtaagggaaccgagaggccttcgtgcaacaagctccgacacagagtcacgtga ctgtaaccacatgagtagcatgttctccccctcgcaccaggacttcttttgaataatccaggtc atccagagggtgggcatgacccctagcatagcgagctacagcggaactgtagctagccctta gcgtgcttggatactgctgataggcgacggcctagtcgtgtcgggttataggtagcacatac aaat

669	Sclerotinia sclerotiorum dsRNA mycovirus-L	ttgaattaatcttttacgtttacgcgcataaaatcaggacacatctctgtatatacttttagtatataatgatgtttttgttttatgcgtaataatcgttaagagaacttcttccatccgcctgtatggcggggataataagttcaccgaccttggcgaggcgcaaaacttgtatgtgcaaaagtgagctatgtctcgaaatagtcgtaactaacacacagccactacctgtagagctctattgatccggaatccttttagtgggaatgcagagctcacaacggacctgcgggtatcttcggcggttagggacttctgttccagccttgaatcatttacctttataccttctctgaggcgccgtggcgccggcgatattaagtacaagcaaggacatcgcgggtagtgctctaatcagccgctagctcgtcgtggagagttccaacttagttgggtgtgggtgcatactagctggatagagtaggtatgtattgtctaactgtatgcccggaggctatccgtcctcggtagaacgtgccggaggtagtctctgcagaccccggaacgcgtgggtcttactaaatgtaggcgggaggagcgctcgtagggtggaaagactgctcccagtcgaatgcaagatttgcacgcggaccagctcgcgggcaattcccgggtg
670	Yak enterovirus	ctccggcacagcgcaccagtgacgtgtagctagctacacctttcacgggtagtcgggtatcccccccgttaacttagaagcatgtaacaaaccgaccaataggtgcgcggcagccagctgcgttgcggtcaagcacttctgtctccccgtccgcaaggatcgttaccgcgccactccactacgaggagcctagtaactggccaagtgtatgctggagttgcgttcagccacaacccagtgtagctctggaagtggggctgcacatccccgttggttaacacgggttgcgttgcgcgtgtgtcttcgggttcagctcgcgactgttcactcaacatcacgcaaccagccaagagccgattgtgctggagtgtcttcccccggggcggtgaatgctgctaactcctaacctccgagcgtgtgcgcacaatccagtggtgctacgtcgtaacgcgtaagtggaggcggaacagactacttccggcaccocgtgttcttatttattcttatttattgtgacaattgcagagattgtgataatgcgacttaccgttaacatagcactgcattacctgtgtgcatccacaaaacttcagagattcctagttctacattgacctactgtttatttgaatcttaatacaaaacttgagcaagtga
671	Wobbly possum disease virus	cggctgtgagtgcttagcatatgctagagtactacagccgggtgtggagtcataatgactggttgcctgtataatagtcgggactgtctgacctacattatcttggaggtgcttatcacgacaattctcggaagtgtctgtcgacagcttaccoccgattcgacaaggccccgtgtccaccgcagacctatcgatttcaacgagaacactatcagaggtttaaatttaaaactcaccaaca
672	Avian orthoreovirus segment S1	gcttttcaatcccttgtgtgatgttccgtatgtcatctgtgtcatgtaacgggtgcaacttctatttttgtaacgttcaactgtcaggcagcgcaaaactcggcggtgtgtatctcaagcgacttctctctgttgcctattggccctacttggctgcggcggtgtgtctcgttgttattatataattgttggcgctgtttgtgtgcaaggctaaaggtgatgagcgagcagccgaaacgttttctaccgagagctgttctgcactaattcgggtaaaaggtgatgcagggaactccgatttaccaggttttagtgcagcagatttgagttttacacttctgtcttagaggagtgcaactactccttccagactataactaataccgatccggctctctactttaacattgagtttccgtcaagtcacgtctctcccccttattccagaactgttctcagcctgttaccgttcacgttctattgattcggagattcgtctctgtgcaaccttatctagtattgtgaatacagactgtgcgctactgccaatcatcaacgctattacgacgatccctacaccaggtgcgtcatctctctgattgttcatggg
673	Caprine Kobuvirus d10	ggggcctcggccccctcaccctctttccgggtggccacgccccgggcccaccgatacttcccttactccttcgggactgttggggagggaacacaacagggtccccctgttttccatcttcccccttttcccaaccccaaccgcgtatctgtgtgcgcgcaagacacacgggtcttccctctaagcacaattgtgtgtgtgtcccagggtcctctgctacgtgctgggagtgctccaccaccaactgttgaagcctgtccaaacgctgctctgcaagactatgacgtcgtatgtccgtcgtcggtatgcgacccgggtaaccgggtccccagtggtgttagtgatctccaggtcctcctgtgttggcgtgtgtccagaaactgcttcagggtaaagtgggtgtgcccatacctacaagggtgattcttccaccacttaggaatgctccggaggtatcccagcaacagctgggatctgaccggagggttaattgtctacgggtgtgttcttcttcttccacacaaactctactgctgacaactcactgactatccacttgcctctgtgtccttctgtcgtgttcaagttcttgattgttttactgcttcttactgcttcttctcacaactctgtcagttcaaaagtc
674	Caprine Kobuvirus d20	ccccctaccccttcttccgggtggccacgcgcggggccaccgatacttcccttactccttcgggactgttggggagggaacacaacagggtccccgttttccattcttcccccttttcccaaccccaaccgcgtatctgttggcggaagacacacgggtcttcccttaagcacaattgtgtgtgtgtgtcccagggtcctcctgctacgtgctgggagtgctccaccaactgttgaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgtatgttccgtcgtcggtatgcgacccgggtaaccgggtccccagtggtgtagtgcgacttccaggtcctcctgttggcgtgtgtccagaaactgcttcaggtaagtgggtgtgcccatacctacaagggtgattcttccaccacttaggaatgctccggaggtacccagcaacagctgggatctgaccggagggttaattgtctacgggtgtgttcttcttcttccacacaaactct

		gacaactcactgactatccacttgcctctcttgcctttctgctctggttcaagtctcttgattgttttga ctgcttttcaactgcttttctctcacaatccttgcctcagttcaaatgc
675	Caprine Kobuvirus d30	ctctttccgggtggccacgcccgggcccacgatacttcccttcaactcttcgggactgttggggagg aacacaacagggtccccctgttttccattcttcccccttttcccaaccccaaccccggtatctggt ggcggcaagacacacgggtcttcccttaaaagcacaattgtgtgtgttccaggctctctgctg acgggtgcgggagtgctcccccaactgttgaagcctgtccaacgcgtcgtcctggcaagactat gacgtgcgatgttccgtgcggatgcgacccgggtaaccgggttccccagtggtgtgtgtagtgcgatct tccaggctctctggttggcggtgttcagaaactgttcaggtaagtggggtgtgcccatacctac aaaggttgattcttccacccttaggaatgctcggagggtacccagcaacagctgggagctgac cggaggctaattgtctacgggtgtgttcttttcttttccacaactctactgctgacaactcactg actatccactgtctcttgtgcttctgtctcgttcaagtctctgtgattgttttgaactgcttttcaactgc tttcttctcacaatccttgcctcagttcaaatgc
676	Caprine Kobuvirus d40	gtggccacgcccgggcccacgatacttcccttcaactcttcgggactgttggggagggaacacaac agggtctccccctgttttccattcttcccccttttcccaaccccaaccccggtatctggtggcgga gacacacgggtcttcccttaaaagcacaattgtgtgtgtgttccagggtcctctgctgacgggtgcg ggagtgcttccaccacaactgttgaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgc atgttccgctgcggatgcgacccgggtaaccgggttccccagtggtgtgtgtagtgcgatcttccaggctc ctcctggttggcggtgttcagaaactgttcaggtaagtggggtgtgcccatacctcacaagggtg attcttccacccttaggaatgctcggagggtacccagcaacagctgggagctgaccggaggct aattgtctacgggtgtgttcttttcttttccacaactctactgctgacaactcactgactatccact tgctctctgtgcttctgtctgttcaagtcttctgattgttttgaactgcttttcaactgcttttcttca caatccttgcctcagttcaaatgc
677	Caprine Kobuvirus d50	ccgggccacgatacttcccttcaactcttcgggactgttggggagggaacacaacagggtctcccc gttttccattcttcccccttttcccaaccccaaccccggtatctggtggcggaagacacacgg gtcttcccttaaaagcacaattgtgtgtgtgttccagggtcctctgctgacgggtgcgggagtgctcc caccacaactgttgaagcctgtccaacgcgtcgtcctggcaagactatgacgtcgcgatgttccgctg cggatgcgacccgggtaaccgggttccccagtggtgtgtagtgcgatcttccagggtcctctggttg cgttgcagaaactgttcaggtaagtggggtgtgcccatacctcacaagggtgattcttccacca ccttaggaatgctcggagggtacccagcaacagctgggagctgaccggaggctaaattgtctacg gggtgtgttcttcttcttccacaactctactgctgacaactcactgactatccactgtctctctgt gcttctgtctggttcaagtcttctgattgttttgaactgcttttcaactgcttttcttccacaatccttgc cagttcaaatgc
678	Picornavirales sp. isolate RtMruf- PicoV	ttgtctcagcgttaacttctccgggttacgtggagacaaaaggctacggagactcgggtacggc cctggagcacctaggtgctcctaaagacgttagaagtgttacaactcgtcccaatagggtccccc aaccagggggttagcgggcaagcacttctgttccccgggtatgatctcatagctgttaccacgg ctgaaagagagattatcgttaccgccctcactacttcgagaagcccagtaattggttcagtaagtgtat ctcgttgaccgggtgttccccacaccagaaactgtgafgggggtgtgcatccgggtatggcg acatgacggacctccccgcggcgacagggtccttctggaggacgagtgacatggattcaacc gtgaagagcctattgagctagtgttgatctcctcgcccccggtgaatgcgggctaaccacacccgga gcaggcgggccccaaaccagggtctggcctgtcgtaacgcgaaagtctggagcgggaaccgacta cttccgggaaggcgtgttcttcttcttcttataagtttatgttgacaactcctggtagacgtttat tgcgtttattgagagatttccacaattgaacagactagaaccactgttttatcaaacctcagagaa taagataaca
679	Apodemus agrarius picornavirus strain Longquan- Aal18	ttactcagcgttaactactccgggttacgtgatgaagaagggtacggagactctcgggtacggc cctggagccactcgggtccttaagatttagaagtgtgacacacccgccactagggtccccc tccaggggggcaacgggcaagcacttctgttccccgggtatgatctgataggctgtaaccacggct gaaacagagattatcgttaccgttcaactctcagagaagcctagtaattggtggaattgaatc cgttgatccgggtgtctccccacaccagaaactcatgatagggttccatccgggtacggcga cgtagcgggcatccctgcgtgcatgaggcctcttaggaggacggatgatgtgacgtgtgtg aagagcctattgagctagtgtcgtactcctcgcgcccggtgaatgcggctaactcctaaccggagc agggtgggtccaatccagggtcctgcgtgtgaatgcgtaagtctgggacgggaaccgactacttct gggaaggcgtgttccattgttcatattgtgtgtttatgggtgacaactcgtggtaaacgttctattgc gttattgagagatttccacaattgaacaaacgagaactacgtgtttattaaattacacagagaag aattaca

680	Niviventer confucianus picomavirus	ccctttcataacccccctttaacccaaccttctgaacctgacgttctactgcctttgggtatag cggcccaatgtgtgaagaaaaggatacgtataaggggccaacgggtgggtgcccttaagacc acccaacctagaagctgtacactcgggcaatagtgaggccacatccagtggggtcaagcccaaa gcattctgttccccgggtatgtatcataagctgtacccacggctgaagagtgatlatcgttatccca ctcagtacttcggagagcctagtagaccacttgaaatggaagctgtatccgggggtgaccctg aaccaccagaaactcatgatgaggctaaccttccgaacacggcgacgtgtggttagcctgcgtg gcatgaggcccttctttaggcagactgaaatggaagggtgacgaagagccgactgagctactgtt tattcctccggccccctgaatgcggctaatcctaactctggtcagtagtgaacccaacagggtg gtcgtgcgaatgcgaagccgggagcggaaacgactactttggggcgctcggtgtttcaatatta ttcatttctagcttatgggtgacaattatgatgagagattgtgctgtatttgggtgagagaagaagt aacaat
681	Bat picomavirus isolate BtRs-PicoV	tttcaaaaggccctgggcatagggcggtattcgtaacgtcgtatgtccaggggcggtgagcatcaggc caaggccctgatgtaccacgtgtggactaaaccacacactcttctgtgacacgtgtgtcacctatc cctttcttgtaacttagaagctgtgtacacttacgcagtaggtgccccacatccagtggggttgtg caaagcaatctgttccccgggtaaacctgtaggctgttaaccacggcggaacaaagggttgcgtt acccgactcactactacacaaaagcctagtaaaagttaagtgaagtgcgacgcgtgatccgggtcaaa accccccttgaccagaaacacatgatgaggggtcaccacccccactggcgacagtgtgtgttccct ggcgttggcatgttggcctcgtagaggcggtgcaatctggattgtcgcgaagagccccgtgtgtagt gtttatactcggccccctgaatgcggctaactcctaaccgccgacgtgtacacacaagccagtgt gtagcatgtcgaatgagcaatttggggatggaaccgactactttaggggtgtcgtgtttctcattatt cttggttgatgttt
682	Rhinolophus picomavirus strain Guizhou-Rr100	ttttttctcaggcggtgtagcatccagccaaggccctgatgtacacacgtgtgactaaaccacactct ctttttgtgatactgtgtacacatccctttcttggttaacttagaagctgtgacaccacgcacgtag gtacccacatccagtggggttgtgcaaagcattcttgggtcccggttaaaccctgataggctgtaa ccacggctgaacaagggttgcgttaccgactcactactacgcaaaagcctagtaagttcaatga aagtgcgcagcgtgtatccgggtcaaaaccccccttgaccagaaacacatgatgaggggtcaccac ccccctggcgacagtgtgtgttccctgcgttggcagtgtggtcctatagaggcggtgcaatctggatt tgctccgaagagccccgtgtgtagtgtttatactccggccccctgaatgcggctaactcctaacc ccgagcatgtacacacacgcagcgtgtgcagcatgtcgttaatgagcaatttggggatggaaaccgac tactttagggtgtccgtgtttctcattattctttgtttgatgttttatgtgtgacaaca
683	Rhinolophus picomavirus strain Henan-Rf265	cggaaacgtgtatgtctcaggcggttaggcaccaccacgggtgtgtgctacacgtgtggactaaac cacacactctttcagcacttagtgcgtctatctcttttgaacttagaagttgtacacaatgcgttag ggccacacatccagtggtgtatgcgaagcacttctgttccccgggtgctagtaggaggggtggtg ctccacggccacttgcgaacccatcggtacccgactcattacttgcgaaggcctagtaaccaggt gaagcaagccccggcggtgttccgtcaggaaaaacccccctggccagaaacatgtgtatgaggggt gggctatccccactgggtgacagtgtgagccctccctgcgttggcacatggccccatctgggcgtggt ctgtggatgctgcgaagagccccgtgagctagtgtttataccgccccgtcgtgaatgcggccta accctaaccgggagcagaggctactgaagccacagtagtgcgtgtcgaacgagtaattctggg atgggaccgactacttccaggtgtccgtgtttcctttattctttattgtgtttatgtgtgacaaac
684	Human enterovirus C105	ccctaggatccactggatgtcagtagactgggtatcgtgtacgttgtacgcctgtttataccccctt ccccgcaactttagaagcatcaaaaagcacgcctcaatagtcaccacacccccagtggtgttcgag caagcacttctgtttcccggtgcgtcccatatgctgtgcaaacggcaaaaaggacataatcgtta ccgctgtatfactacgggaaacctagtaccaaccattgattgtgttgagaggtgcgctcatcaccttc cccggtgtagctcaggccgatgaggtcagaatccccacagggtgactgtgtctgagcctgcgtt ggcggtcctgcctcgttattggcgtgggacgcttgatacatgacatggtgcgaagagtctactgt gctatgcaagagtctccggccccctgaatgtggctaactcctaaccactgacccacgcacgcaaac cagtgtgtagtggtcgtaacgcgcaagtgcgtggcgaaccaactacttgggtgacogtgtttc ctttacttattgaatgtttatgtgtgacaattgtttgattcagttgttgcattctctacattcattacca gcatcaaccaattgaactgttaca
685	Human poliovirus 1 strain NIE1116623	agtctggacatccctcaccggcgacgggtgtgtcaggctgcgtggtggcgccctactgtgtgtccaaa gccacgggacgtacatgtgaacaagggtgaagagcctattgagctacaaaagagtcctccgg ccctgaatgcggctaactccaaaccggatcaagggtgcacaaccaggtgtacacctgtcgtga acggcgcaagtctgtggcggaaccgactacttgggtgtccgtgttcttttaattttagtgctgtt atgggtgacaatcatagattgttatcataaagctaattggattggccatccgggtgagagtgaaatatt

		gtttacctccctgttgggttactctaactacttctccattataaacttgcacacagtttaataatta gaagtgcagtttaca
686	Human enterovirus 109	tttaaacacagcctgggggtgttccacccccagggccactgggctgttagtactctggatcgcg gtaccttagtatgctgtttatgtctcttccccgcaactttagaagtaatacaagttatggctcaaca gtgcacacccccagtggttccgagcaagcacttctgttccccgttgcgtcttatatgctgtgt gaaacggcagaaagggacaatatcgttatccgctcaactactacgggaagccttagtaccaccatgg attgacctgaaagtgtgcttcagcgcacccccagcgcagctcagggcagatgaggtcgcgaatacc ccacggcgacccgtgtcggagcctgctgttggcggcctgcccacgttgcaaacgtgggacgctc attcatgacatggtgcgaagagcctactgtgctagttgagagtcctcggccccctgaatgttgata atcctaaccactgaacctacggcgcaaacagcgtctgttagggcgttaacgcgaagtcgggtg gcggaaaccaactacttgggtgtccgtgtttctttatcttttgaatgttatgtgacaattgtgtgta cagttgttaccatagtttgcattcagaataaaccataacacttccaattattgttaca
687	Human poliovirus 2 strain NIE0811460	ttgtgcgcctgtttatattccccccgcaacttagaagcacgaaaccaagttaatagaaggggg tacaaccagtagtaccacgaacaagcacttctgttccccgtgacattgcatagactgctcacgc ggttgaagtgatcgatccgttaccgctgtgtacttgcgaaagcctagttatgcttggaaattcteg acgcgttgcgtcagcaccgcacccgggggtgtagcttaggctgatgagctctggacattctcacc ggtagcgtgtgctcaggctgctgttggcggcctacctatgctaatgccaatagacgctagatgtg aacaaggtgtgaagagcctattgagctacataagagtcctccggccccctgaatgcggctaactca accacggagcagggcgtgcgaaccagtgactagctgtgtaacgcgcaagtctgtgcgga ccgactacttgggtgtcgtgtttctgttattttattatggctgttattgtgacaatcagagattgtt atcataaagcgaattggattggccatccggtgagtggtgtgtcagggtatatacaactgtttgttgaa actgtgttagtttaacctctcttcaaccaattagtaaaaaacaatacgaagatagaacaacaacta ca
688	Bovine picornavirus	ttttccccctccccccaactaccttttcccccttgttaacgctagaagtgtgtgcaaacgcctgt agggtactgcaatccagcagtgcataggctaaagcttttctgttaccacccccacattatactgagg aggattgtgaaattgtgttagtatgggttagtagcgggtgacccgggtaacccccacccagaaactc acggatgagatgaacaggacccacatggtaacgtgtgtgttcgtctgccccgcaaggtgaggcc gtgagagctttgcacgcgaaaacctgaaaaaccaaagtaccttgagctcttcgctattttgttttc ctccaggaccctgaatgcggctaaacctaacccgcgctcgcacgtagcaaccagcctagagtg gttcgtaatgcgcaagtgtcggcggtaccgactacttgggtgttctctgtgttcttattttatgtga attttatggtgacaacagctagaataaagagtgaaac
689	Human poliovirus 1 strain EQG1419328	accctgtacgcctgttttatactccccctccgtaacttagaagaacaaaaataagttaaataggag ggggtacaaaccagtagtaccacgaacaagcacttctgttccccgtgacattgcatagactgtc cccacgggtgaaagcaattgatccgttaccgctctgttacttcgagaagcctagtaccatcttgaa tcacgatgctgtgcgtccacactcagtcacagagtgtagcttaggctgatgagcttgacattcct caccggcgacgggtgtccaggctgctgttggcggcctacctgttggcccaagccacaggacgct agatgtgaacaagggtgtgaagagcctattgagctataagagagtcctccggccccctgaatgcggc taatcccaaccacggatcaagggtgcacgaaccagtgataccctgtcgtaacgcgcaagtccgtg gcggaaaccgactacttgggtgacgctgttctttatttcaatgctgtctatggtgacaatcatt gattgttatcataaagcgaattggactggccatccgggtgaaagtgaacataattgttgcctcctgtt gggtctacttcaaccaatcttacttacaatcttaccactacagtttgcgtgttagaagtgtgtttcacg
690	Human poliovirus 2 isolate IS_061	ttgtgcgcctgttttatactccccctccgcaacttagaagcacgaaaccaagttaatagaaggggg tacaaccagtagtaccactacgaacaagcacttctgttccccgtgacattgcatagactgctcacgc ggttgaagtgatcgatccgttaccgctgtgtgtacttcgaaaagcctagtaacgcttggaaattctc gacgcgttgcgtcagcaccgcacccgggggtgtagcttaggctgatgagcttgacattctca ccggtagcgggtgtcaggctgctgttggcggcctacctatggctaacgccaatagacgttagatg tgaaacaggtgtgaagagcctattgagctacataagagtcctccggccccctgaatgcggctaacc taaccacggagcagggcgtgcgaaccagtgactgctgtgtcgaacgcgcaagtctgttggcgg aacgactacttgggtgtcgtgttctgttattttatcatggctgttattgtgacaatcagagatt gttatcataaagcgaattggattggccatccgggtgagtggtgtgtcagggtatacaactgtttgttgaa ccactgtgttagcttcttctcaatttaaccaatatacaaaaaacaatacagaggataaaacaacaatac taca
691	Coxsackievirus B5	cccttgtgcgcctgtttatgccccctcccccaattgaaacttagaaggtacacacaccgatcaacag cgggcgtggcataccagccggctgtgtatcaagcactcctgttccccggaccgagtagtaataaga

		ctgctcacgcgggtgaaggagaaaacgttcgttaccgggctaactacttcgagaacctagtagca tcatgaaagtgcgaagcggttcgctcagcacatcccagtgtagatcagggtcgatgagtcacgc atccccacgggagaccgtggcggtggctgcgttggcgccctgcctacggggcaacccgtagg acgctcaatacagacatgggtgcgaagagtcgattgagctagtagtctccggccctgaatc cggctaactcctaactgcggagcacataccctcaaccagggggcatgtgtcgtaacgggtaact ctgcagcgaaccgactacttgggtgctcgtttctttattcttataatgctgcttatggtgaca attgaaagattgtaccatatagtattggattggccatccgggtgcttaacagagctattatatacctct ttgttggattgtaccacttgatctaaagggaagcaagacactacaattcatcataaattgaacacag caaa
692	Coxsackievirus A10	tttgtgcgcctgtttacaaccttcccacttgaacgtagaagtaatacacactactgatcaatag caggcatggcgccagtcagtctcgtatcaagcacttctgttccccggactgagtatcaataga ctgctcacgcgggtgaaggagaaaacgttcgttaccgggctaactacttcgagaacctagtagca ccatagaagctgcagagtggttcgtcagcacttccccgtgtatcaggctgatgagtcactgca atccccacgggtagaccgtggcagtggtgcgttggcgccctgcctatggggcaacccataggac gctctaattgtggacatgggtgcgaagagctattgagctagtagtctccggccctgaatgcg gctaactcctaactgcggagcacatgccttcaaccagaaggtagtggtcgtaacgggcaactctg cagcggaaacgactacttgggtgctcgtttcttttattctatattgctgcttatgggtgacaatca cggaaattgttccatagctattggattggccatccgggtgcttaatagagctattgtgtacctattgtt ggatttactcgcgtatcacataaattctgaacactttgtgcttatattgaactaaacacccgaaa

[0680] In some embodiments, an IRES of the invention is an IRES having a sequence as listed in Table 17 (SEQ ID NOs: 1-72 and 348-389). In some embodiments, an IRES is a Salivirus IRES. In some embodiments, an IRES is a Salivirus SZ1 IRES. In some embodiments, an IRES is a AP1.0 (SEQ ID NO:348). In some embodiments, an IRES is a CK1.0 (SEQ ID NO:349). In some embodiments, an IRES is a PV1.0 (SEQ ID NO:350). In some embodiments, an IRES is a SV1.0 (SEQ ID NO:351).

Table 18. Anabaena permutation site 5' intron fragment sequences.

SEQ ID NO	Permutation site	Sequence
73	L2-1	GAAGAAATTCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCT TAAATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAG TAGTAATTAGTAAGTTAACAATAGATGACTTACAACCTAATCG GAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGAC GAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGT AGCGAAAGCTGCAAGAGAATGAAAATCCGT
74	L2-2	AAGAAATTCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCT AAATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGT AGTAATTAGTAAGTTAACAATAGATGACTTACAACCTAATCGG AAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACG AGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTA GCGAAAGCTGCAAGAGAATGAAAATCCGT
75	L2-3	AGAAATTCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTA AATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTA GTAATTAGTAAGTTAACAATAGATGACTTACAACCTAATCGGA AGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGA GGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAG CGAAAGCTGCAAGAGAATGAAAATCCGT
76	L5-1	GTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATT

		AGTAAGTTAACAATAGATGACTTACAATAATCGGAAGGTGC AGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAA AGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAG CTGCAAGAGAATGAAAATCCGT
77	L5-2	TTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTA GTAAGTTAACAATAGATGACTTACAATAATCGGAAGGTGCA GAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAA GAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGC TGCAAGAGAATGAAAATCCGT
78	L5-3	TATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAG TAAGTTAACAATAGATGACTTACAATAATCGGAAGGTGCAG AGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAG AGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCT GCAAGAGAATGAAAATCCGT
79	L5-4	ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAATAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTG CAAGAGAATGAAAATCCGT
80	L5-5	TAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTA AGTTAACAATAGATGACTTACAATAATCGGAAGGTGCAGAG ACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGT
81	L6-1	ACAATAGATGACTTACAATAATCGGAAGGTGCAGAGACTCG ACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAG TCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGA GAATGAAAATCCGT
82	L6-2	CAATAGATGACTTACAATAATCGGAAGGTGCAGAGACTCGA CGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTC CAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGA ATGAAAATCCGT
83	L6-3	AATAGATGACTTACAATAATCGGAAGGTGCAGAGACTCGAC GGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCC AATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAA TGAAAATCCGT
84	L6-4	ATAGATGACTTACAATAATCGGAAGGTGCAGAGACTCGACG GGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCA ATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAAT GAAAATCCGT
85	L6-5	TAGATGACTTACAATAATCGGAAGGTGCAGAGACTCGACGG GAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAA TTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATG AAAATCCGT
86	L6-6	AGATGACTTACAATAATCGGAAGGTGCAGAGACTCGACGGG AGCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAAT TCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGT
87	L6-7	GATGACTTACAATAATCGGAAGGTGCAGAGACTCGACGGGA GCTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATT CTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGT
88	L6-8	ATGACTTACAATAATCGGAAGGTGCAGAGACTCGACGGGAG CTACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCT

		TCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAA AATCCGT
89	L6-9	TGACTTACAACATAATCGGAAGGTGCAGAGACTCGACGGGAGC TACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCT CAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAA ATCCGT
90	L8-1	CAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATA GGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
91	L8-2	AAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAG GCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
92	L8-3	AGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
93	L8-4	GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
94	L8-5	ACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCA GTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
95	L9a-1	AATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
96	L9a-2	ATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
97	L9a-3	TAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
98	L9a-4	AGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
99	L9a-5	GGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
100	L9-1	GAAAGCTGCAAGAGAATGAAAATCCGT
101	L9-2	AAAGCTGCAAGAGAATGAAAATCCGT
102	L9-3	AAGCTGCAAGAGAATGAAAATCCGT
103	L9-4	AGCTGCAAGAGAATGAAAATCCGT
104	L9-5	GCTGCAAGAGAATGAAAATCCGT
105	L9-6	CTGCAAGAGAATGAAAATCCGT
106	L9-7	AAGAGAATGAAAATCCGT
107	L9-8	AGAGAATGAAAATCCGT
108	L9-9	GAGAATGAAAATCCGT
109	L9a-6	GCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
110	L9a-7	AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
111	L9a-8	GTAGCGAAAGCTGCAAGAGAATGAAAATCCGT

[0681] In some embodiments, a 5' intron fragment is a fragment having a sequence listed in Table 18. Typically, a construct containing a 5' intron fragment listed in Table 18 will contain a corresponding 3' intron fragment as listed in Table 19 (e.g., both representing fragments with the L9a-8 permutation site).

Table 19. Anabaena permutation site 3' intron fragment sequences.

SEQ ID NO	Permutation site	Sequence
112	L2-1	ACGGACTTAAATAATTGAGCCTTAAA
113	L2-2	ACGGACTTAAATAATTGAGCCTTAAAG
114	L2-3	ACGGACTTAAATAATTGAGCCTTAAAGA
115	L5-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTA
116	L5-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAG
117	L5-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGT
118	L5-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT

119	L5-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTA
120	L6-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT A
121	L6-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AA
122	L6-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AAC
123	L6-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACA
124	L6-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAA
125	L6-6	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAAT
126	L6-7	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATA
127	L6-8	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAG
128	L6-9	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGA
129	L8-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAATAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGT
130	L8-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAATAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTC
131	L8-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAATAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCA
132	L8-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG

		TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAA
133	L8-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAG
134	L9a-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCC
135	L9a-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCA
136	L9a-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAA
137	L9a-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAAT
138	L9a-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATA
139	L9-1	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGC
140	L9-2	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCG
141	L9-3	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT

		AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGA
142	L9-4	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAA
143	L9-5	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAA
144	L9-6	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAA
145	L9-7	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC
146	L9-8	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCA
147	L9-9	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCA A
148	L9a-6	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAG
149	L9a-7	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGC
150	L9a-8	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA

		CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAATAATCGGAAGGTGCAGAGACT CGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCA
--	--	---

[0682] In some embodiments, a 3' intron fragment is a fragment having a sequence listed in Table 19. In some embodiments, a construct containing a 3' intron fragment listed in Table 19 will contain a corresponding 5' intron fragment as listed in Table 18 (e.g., both representing fragments with the L9a-8 permutation site).

Table 20. Non-anabaena permutation site 5' intron fragment sequences.

SEQ ID NO	Intron	Sequence
151	Azop1	tgcgccgatgaagggtgtagagactagacggcaccacctaaggcaaacgctatggtgaaggcatagtcca gggagtgccgaaagtcacacaaccggatccgt
152	Azop2	ccggcgctatggcaacgcccagccaagcttcggcgctcgccgatgaagggtgtagagactagacggc accacctaaggcaaacgctatggtgaaggcatagtccaggagtgccgaaagtcacacaaccggat ccgt
153	Azop3	acggcaccacctaaggcaaacgctatggtgaaggcatagtccaggagtgccgaaagtcacacaacc ggatccgt
154	Azop4	acgctatggtgaaggcatagtccaggagtgccgaaagtcacacaaccggatccgt
155	S795p1	attaaagttagaattatcagagaatgatatagtccaagccttatgtaacatgagggcacttgaccctggt g
156	Twortp1	aagatgtaggcaatcctgagctaaagctcttagtaataagagaagtgcaacgactattccgataggaa gtaggtcaagtgactcgaaatggggattacccttctagggtagtgtatagctgaacatataatggaacatata aaggataggagtaacgaacctattcgtacataattgaacttttagttat
157	Twortp2	taataagagaagtgcaacgactattccgataggaaagtaggggtcaagtgactcgaaatggggattaccct tagggtagtgatagctgaacatataatggaacatataagagataggagtaacgaacctattcgtacat aattgaacttttagttat
158	Twortp3	taggaagtaggggtcaagtgactcgaaatggggattacccttctagggtagtgtatagctgaacatata aaacatataagagataggagtaacgaacctattcgtacataattgaacttttagttat
159	Twortp4	ctagggtagtgtatagctgaacatataatggaacatataagagataggagtaacgaacctattcgtac taattgaacttttagttat
160	LSUp1	agttaataaagatgatgaatagctgaaccattttgagaaaagtggaataaaaagaaatctttatgataac ataaatgaacaggctaa
161	Phip1	caaagactgatgataatagtcgacactcctagtaataaggagaatacagaaggatgaaatcc
162	Nostoc	agtcgagggtaaaggagagtgcaattctcaaagcctattgagcagtagcgaaagctgcgggagaatgaa atccgt
163	Nostoc	agccgagggtaaaggagagtgcaattctcaaagccaataggcagtagcgaaagctgcgggagaatgaa aatccgt
164	Nodularia	agccgagggtaaaggagagtgcaattctcaaagccgaaggttaftaaaacctggcagcagtgaaagctg cgggagaatgaaatccgt
165	Pleurocaps a	agctgagggtaaaggagagtgcaattctcaaagccagcagatggcagtagcgaaagctgcgggagaat gaaatccgt
166	Planktothri x	agccgagggtaaaggagagtgcaattctcaaagccaatggtagtagcgaaagctacgggagaatgaa atccgt

[0683] In some embodiments, a 5' intron fragment is a fragment having a sequence listed in Table 20. A construct containing a 5' intron fragment listed in Table 20 will contain a

corresponding 3' intron fragment in Table 21 (e.g., both representing fragments with the Azop1 intron).

Table 21. Non-anabaena permutation site 3' intron fragment sequences.

SEQ ID NO	Intron	Sequence
167	Azop1	gcggactcataattcgaatgtgccttgcgcgggaaaccacgcaagggaatgtgtcaaatcggcgaac ctaagcgcccgcccggtatggcaacgcccagccaagcttcggcgcc
168	Azop2	gcggactcataattcgaatgtgccttgcgcgggaaaccacgcaagggaatgtgtcaaatcggcgaac ctaagcgcccgcc
169	Azop3	gcggactcataattcgaatgtgccttgcgcgggaaaccacgcaagggaatgtgtcaaatcggcgaac ctaagcgcccgcccggtatggcaacgcccagccaagcttcggcgccctgcgcgatgaagggtga gagactag
170	Azop4	gcggactcataattcgaatgtgccttgcgcgggaaaccacgcaagggaatgtgtcaaatcggcgaac ctaagcgcccgcccggtatggcaacgcccagccaagcttcggcgccctgcgcgatgaagggtga gagactagacggcaccacctaaggcaa
171	S795p1	aggattagataactaactaagtgtccccagactgggtgacagtctgggtgcatccagctatatcggtgaa acccattggggtaataaccgagggaagctatattatataataataatagcccgtagagactatgta ggtaaggagatagaagatgataaaatcaaatcatc
172	Twortp1	actactgaaagcataaataaftgtgcctttatacagtaagtatatcgaaaaatcctctaattcagggaacac ctaaacaaact
173	Twortp2	actactgaaagcataaataaftgtgcctttatacagtaagtatatcgaaaaatcctctaattcagggaacac ctaaacaaactaagatgtaggcaatcctgagctaagctcttag
174	Twortp3	actactgaaagcataaataaftgtgcctttatacagtaagtatatcgaaaaatcctctaattcagggaacac ctaaacaaactaagatgtaggcaatcctgagctaagctcttagtaataagagaagtgcaacgactattcc ga
175	Twortp4	actactgaaagcataaataaftgtgcctttatacagtaagtatatcgaaaaatcctctaattcagggaacac ctaaacaaactaagatgtaggcaatcctgagctaagctcttagtaataagagaagtgcaacgactattcc gatagggaagtagggtaagtgactgaaatgggattaccctt
176	LSUp1	cgctagggattataactgtgagtcctcaatattataaaatgttgtaatatattgggtaaatccaagaca acttttccacgacaggaatagtgatttgaagcgaaactatttttagcagtgaagcaataaggaac gttcaacgactaaaaggtagtattgctaacaataatccttttttaattgccaacatcttttaact
177	Phip1	gtgggtgcataaactatttcattgtgcacattaaatctggtaactcggtaaacccctaatggggcaatacc gagccaagccataggaggatattgagaggcaagaagttaattcttgaggccactgagactggctgta tcacccctacgtcacacaaactaatgccgatgttatttcagaaagaaaaccaatggcgtcttagagatgt atcacagaacgggtgtggaaggagcataacggagacatacctgattgcttcgagatagaccataaggtgc

		gcaatagggtctgctgtaatatagagcattfacagatgcttgagggtacagcccacactgttaagaccaat cgtgaacgctacgcagacagaaaggaaacagctagggaatactggctggagactggatgtaccggcc tagcactcgggtgagaagtttggtgtgtctgtctctctctgctgtaagtggalttagagaatggaaggcgtaga gactatccgaaaggagtagggccgagggtgagactccctcgtaacccgaagcgccagacagtaact
178	Nostoc	acggacttaagtaattgagccttaagaagaattctttaagtggcagctctcaaaactcagggaaacctaa atctgttcacagacaaggcaatcctgagccaagccgaagagtcagtgctgagtagtgagtaaaat aaaagctcacaactcagaggtgttaactctaagctagtcgggaaggtgcagagactcgacgggagctac cctaacgtaa
179	Nostoc	acggacttaaaactgaattgagccttagagaagaattctttaagtgtcagctctcaaaactcagggaaacct aaatctgttgacagacaaggcaatcctgagccaagccgagaactctaagttattcgaaggtgcagaga ctcgacgggagctaccctaacgtca
180	Nodularia	acggacttagaaaactgagccttgatcgagaaatctttcaagtggagctctcaaaactcagggaaacctaa atctgtttacagatatggcaatcctgagccaagccgaacaagtcctgagtgtaaagctcataactc gggaaggtgcagagactcgacgggagctaccctaacgtta
181	Pleurocapsa	acggacttaaaaaaattgagccttgagcagagaaatctgtcatgcaacgctctcaaaactcagggaaacct aagtcctggcaacagatatggcaatcctgagccaagcccttaacaggaacaaacattttacctttacctt gaaaggaaggtgcagagactcaacgggagctaccctaacagggtca
182	Planktothrix	acggacttaagataaattgagccttgagggcagagaaatctctcaagtgtgaagctgtcaaaactcagggaaa cctaatactgtaaattcagacaaggcaatcctgagccaagcctagggtattagaaatgagggagtttcc ccaatctaagatcaatacctaggaaggtgcagagactcgacgggagctaccctaacgtta

[0684] In some embodiments, a 3' intron fragment is a fragment having a sequence listed in Table 21. A construct containing a 3' intron fragment listed in Table 21 will contain the corresponding 5' intron fragment as listed in Table 20 (*e.g.*, both representing fragments with the Azop1 intron).

Table 22. Spacer and Anabaena 5' intron fragment sequences.

SEQ ID NO	Spacer	Sequence
183	T25 L10	agtatataagaacaaaccacTAGATGACTTACAACCTAATCGGAAGGTGC AGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAA AGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAG CTGCAAGAGAATGAAAATCCGTggctcgagc
184	T25 L20	ctgaaattatacttataactcaacaaaccacTAGATGACTTACAACCTAATCGGAA GGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGC GAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
185	T25 L30 (180-10) [Control]	ctgaaattatacttataactcagtatatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC

		AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
186	T25 L40	catcaacaatatgaaattatacttatactcagttatgacaaacaaccacTAGATGACTTAC AACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCCTAA CGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCA ATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggct cgagc
187	T25 L50	catcaacaatatgaaactatacttatactcagttatgagcattatcgcaacaaccacTAGATG ACTTACAATAATCGGAAGGTGCAGAGACTCGACGGGAGCTA CCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCA AAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAAT CCGTggctcgagc
188	T50 L10	tagcgtcagcaacaacaacaaTAGATGACTTACAATAATCGGAAGGTGC AGAGACTCGACGGGAGCTACCCCTAACGTCAAGACGAGGGTAA AGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAAG CTGCAAGAGAATGAAAATCCGTggctcgagc
189	T50 L20	atactcatactagcgtcagcaacaacaacaaTAGATGACTTACAATAATCGGA AGGTGCAGAGACTCGACGGGAGCTACCCCTAACGTCAAGACGA GGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAG CGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
190	T50 L30	gtgtgaagctatactcatactagcgtcagcaacaacaacaaTAGATGACTTACAATA ATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCCTAACGTCA AGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
191	T50 L40	cctcacctgagtgtaagctatactcatactagcgtcagcaacaacaacaaTAGATGACTTA CACTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCCTA ACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCC AATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgg ctcgagc
192	T50 L50	cggaatgatgectcacctgagtgtaagctatactcatactagcgtcagcaacaacaacaaTAGAT GACTTACAATAATCGGAAGGTGCAGAGACTCGACGGGAGCT ACCCTAACGTCAAGACGAGGGTAAAGAGAGAGTCCAATTCTC AAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAAT TCCGTggctcgagc
193	T75 L10	cgggtcgagcaacaacaacaaTAGATGACTTACAATAATCGGAAGGTG CAGAGACTCGACGGGAGCTACCCCTAACGTCAAGACGAGGGTA AAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAA GCTGCAAGAGAATGAAAATCCGTggctcgagc
194	T75 L20	cgctcgcagccagtcgagcaacaacaacaaTAGATGACTTACAATAATCGG AAGGTGCAGAGACTCGACGGGAGCTACCCCTAACGTCAAGACG AGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTA GCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
195	T25 L30 1MM	ctgaaattatactAatactcagttatgacaaacaaccacTAGATGACTTACAATA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
196	T25 L30 3MM	ctgaaaAtatactAatactcaCtatatgacaaacaaccacTAGATGACTTACAATA ATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCCTAACGTCA AGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGG CAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
197	T25 L30 5MM	ctgaTaAtataGtAatactcaCtatatgacaaacaaccacTAGATGACTTACAATA AATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCCTAACGTC AAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAG GCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
198	T25 L30	ctgaTaAtaAaGtAatacAcaCtataAgacaacaaccacTAGATGACTTACAA

	8MM	CTAATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACG TCAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAAT AGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgc agc
199	T25 L30 OffTarget 10	ctgaaattatacttatactctctaagttacaacaaaccacTAGATGACTTACAACCTAAT CGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAAG ACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCA GTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
200	T25 L30 OffTarget 20	ctgaaattatgtgtgttacAtctaagttacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
201	T25 L30 OffTarget 30	gttgatcggtgtgtgttacAtctaagttacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagc
202	T25 L30 I25- 10	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattaaacag
203	T25 L30 I25- 20	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattcacaatataaa ttacg
204	T25 L30 I50- 10	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcatagc
205	T25 L30 I50- 20	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcgcagcataa tatccg
206	T25 L30 I80- 20	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagcgcg cctaccg
207	T25 L30 I80- 20x2	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgcagcgcg cctaccgaaagccggcgtcgacgttagcgc
208	T25 L30 I50- 20x2	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcgcagcataa tatccgaaacgaggatacaagtgcacatgc
209	T25 L30 I25- 20x2	ctgaaattatacttatactcagttatgacaacaaaccacTAGATGACTTACAACCTAA TCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGC AGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattcacaatctaaa ttacgaaacgataaatgataactctaac

210	T0 L0	aaacaaaccacTAGATGACTTACAACCTAATCGGAAGGTGCAGAGAC TCGACGGGAGCTACCCCTAACGTCAAGACGAGGGTAAAGAGAG AGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAA GAGAATGAAAATCCGTggctcgagc
211	T100 L5	cgggcaacaacaaaTAGATGACTTACAACCTAATCGGAAGGTGCAG AGACTCGACGGGAGCTACCCCTAACGTCAAGACGAGGGTAAAG AGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCT GCAAGAGAATGAAAATCCGTggctcgagc
212	T75 L30	cgtccgacgagcttcggccagtcgagcaacaacaaaTAGATGACTTACAAC AATCGGAAGGTGCAGAGACTCGACGGGAGCTACCCCTAACGTC AAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAG GCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
213	T0 L0a	aaacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCC GTggctcgagc
214	T25 L10a	agtataaagaacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgagc
215	T25 L20a	ctgaaattatacttatactcaacaaccacGGCAGTAGCGAAAGCTGCAAGAG AATGAAAATCCGTggctcgagc
216	T25 L30a (I80-10) [Control]	ctgaaattatacttatactcagtatatgacaacaaccacGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggctcgagc
217	T50 L10a	tagcgtcagcaacaacaaaGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgagc
218	T50 L20a	atactatactagcgtcagcaacaacaaaGGCAGTAGCGAAAGCTGCAAGAG AATGAAAATCCGTggctcgagc
219	T50 L30a	gtgtgaagctatactatactagcgtcagcaacaacaaaGGCAGTAGCGAAAGCTG CAAGAGAATGAAAATCCGTggctcgagc
220	T75 L10a	cgggtcgagcaacaacaaaGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgagc
221	T75 L20a	cgtccgacccagtcgagcaacaacaaaGGCAGTAGCGAAAGCTGCAAGA GAATGAAAATCCGTggctcgagc
222	T75 L30a	cgtccgacgagcttcggccagtcgagcaacaacaaaGGCAGTAGCGAAAGCT GCAAGAGAATGAAAATCCGTggctcgagc
223	T0 L0b	aaacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAA GCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCC GTggctcgagc
224	T25 L10b	agtataaagaacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATT CTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAA AATCCGTggctcgagc
225	T25 L20b	ctgaaattatacttatactcaacaaccacAAGACGAGGGTAAAGAGAGAGTC CAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGA ATGAAAATCCGTggctcgagc
226	T25 L30b (I80-10) [Control]	ctgaaattatacttatactcagtatatgacaacaaccacAAGACGAGGGTAAAGAG AGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGC AAGAGAATGAAAATCCGTggctcgagc
227	T50 L10b	tagcgtcagcaacaacaaaAAGACGAGGGTAAAGAGAGAGTCCAATT CTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgagc
228	T50 L20b	atactatactagcgtcagcaacaacaaaAAGACGAGGGTAAAGAGAGAGT CCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAG AATGAAAATCCGTggctcgagc
229	T50 L30b	gtgtgaagctatactatactagcgtcagcaacaacaaaAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTG CAAGAGAATGAAAATCCGTggctcgagc

230	T75 L10b	cggtgcgagcaaacacacaaaAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGA AAATCCGTggctcgagc
231	T75 L20b	cgctccgacccagtcgagcaaacacacaaaAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
232	T75 L30b	cgctccgacgagctccggccagtcgagcaaacacacaaaAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagc
233	T25 L30 I0-0	ctgaaattatacttatactcagtatatgacaacacaaaccacTAGATGACTTACAATACTCGGAAGGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAA GACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
234	T25 L30a I0-0	ctgaaattatacttatactcagtatatgacaacacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
235	T25 L30a I25-10	ctgaaattatacttatactcagtatatgacaacacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattaaacag
236	T25 L30a I25-20	ctgaaattatacttatactcagtatatgacaacacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattcacaataataaattacg
237	T25 L30a I50-10	ctgaaattatacttatactcagtatatgacaacacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcatagc
238	T25 L30a I50-20	ctgaaattatacttatactcagtatatgacaacacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcgagcataatatccg
239	T25 L30a I80-20	ctgaaattatacttatactcagtatatgacaacacaaaccacGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagcgcgcctaccg
240	T25 L30b I0-0	ctgaaattatacttatactcagtatatgacaacacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGT
241	T25 L30b I25-10	ctgaaattatacttatactcagtatatgacaacacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattaaacag
242	T25 L30b I25-20	ctgaaattatacttatactcagtatatgacaacacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTgattcacaataataaattacg
243	T25 L30b I50-10	ctgaaattatacttatactcagtatatgacaacacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcatagc
244	T25 L30b I50-20	ctgaaattatacttatactcagtatatgacaacacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggatcgagcataatatccg
245	T25 L30b I80-20	ctgaaattatacttatactcagtatatgacaacacaaaccacAAGACGAGGGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAGGCAGTAGCGAAAGCTGCAAGAGAATGAAAATCCGTggctcgagcgcgcctaccg

[0685] In some embodiments, a spacer and 5' intron fragment are spacers and fragments having sequences as listed in Table 22.

Table 23. Spacer and Anabaena 3' intron fragment sequences.

SEQ ID NO	Spacer	Sequence
246	T25 L10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC

		TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacaaactatataact
247	T25 L20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacaaagagtataagtataaatttcag
248	T25 L30 (180-10) [Control]	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacaaagtcataactagatataagtataaatttcag
249	T25 L40	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacaaagtcataactagatataagtataaatttcataatttgcag
250	T25 L50	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacaaagcgataatgcttcataactagatataagtataagtttcataattg ttgatg
251	T50 L10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacacaaagctgacgcta
252	T50 L20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacacaaagctgacgctagatagagtat
253	T50 L30	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacacaaagctgacgctagatagagtatagcttcacac
254	T50 L40	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacacaaagctgacgctagatagagtatagcttcacactcaggtgagg
255	T50 L50	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacacaaagctgacgctagatagagtatagcttcacactcaggtgagggc atcattcgg
256	T75 L10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacacaaagctgcacgg
257	T75 L20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacacaaagctgcacgctggctcagggcgg
258	T25 L30 1MM	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacaaagtcataactagatataagtataaatttcag
259	T25 L30	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC

	3MM	TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
260	T25 L30 5MM	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
261	T25 L30 8MM	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
262	T25 L30 OffTarget 10	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtaacttagagagtataagtataaatttcag
263	T25 L30 OffTarget 20	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtaacttagaTgtaacacacataaatttcag
264	T25 L30 OffTarget 30	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtaacttagaTgtaacacacacagatcaac
265	T25 L30 I25- 10	ctgtttaatcACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCT TTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
266	T25 L30 I25- 20	cgtaatttatattgtaacACGGACTTAAATAATTGAGCCTTAAAGAAGA AATTCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATC TAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAA TTAGTAAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
267	T25 L30 I50- 10	gctatgatccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
268	T25 L30 I50- 20	cggatattatgctgcgatccACGGACTTAAATAATTGAGCCTTAAAGAAG AAATTCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAA TCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGT AATTAGTAAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
269	T25 L30 I80- 20	cggtaggcgcgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAA GAAATTCCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAA ATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAG TAATTAGTAAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
270	T25 L30 I80- 20x2	gcgctaactgcgacgccggaacggtaggcgcgctgcgagccACGGACTTAAATAA TTGAGCCTTAAAGAAGAAATTCCTTTAAGTGGATGCTCTCAAAC TCAGGGAAACCTAAATCTAGTTATAGACAAGGCAATCCTGAG CCAAGCCGAAGTAGTAATTAGTAAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
271	T25 L30 I50- 20x2	gcattgtcactgtatcctcgaaacggatattatgctgcgatccACGGACTTAAATAATTG AGCCTTAAAGAAGAAATTCCTTTAAGTGGATGCTCTCAAACCTC AGGGAAACCTAAATCTAGTTATAGACAAGGCAATCCTGAGCC AAGCCGAAGTAGTAATTAGTAAGTTAACAAcacaacacacagtcatactgagtataagtataaatttcag
272	T25 L30 I25-	gtagaggtatcattatcgaaacgtaatttagattgtgaacACGGACTTAAATAATTGA

	20x2	GCCTTAAAGAAGAAATTCTTTAAGTGGATGCTCTCAAACCTCA GGGAAACCTAAATCTAGTTATAGACAAGGCAATCCTGAGCCA AGCCGAAGTAGTAATTAGTAAGTTAACAAcacaacacaagtcatactg agtataagtataatttcag
273	T0 L0	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAcacaacacaa
274	T100 L5	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacaaagcccg
275	T75 L30	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAaacaacaaagctcgcaactggccggaagctcgtagcgagcg
276	T0 L0a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaacacaa
277	T25 L10a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaacacaacttatatact
278	T25 L20a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaacacaagagtataagtataatttc ag
279	T25 L30a (I80-10) [Control]	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaacacaagtcatactgagtataa gtataatttcag
280	T50 L10a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaacaaagctgacgcta
281	T50 L20a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaacaaagctgacgctagtatga

		gtat
282	T50 L30a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctgacgctagtatga gtatagcttcacac
283	T75 L10a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctgcgacccg
284	T75 L20a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctgcgactgggtcgg agcg
285	T75 L30a	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAaacaaaaacaagctgcgactggcgga agctcgtcgagcg
286	T0 L0b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCCacaaacacaa
287	T25 L10b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCCacaaacacaacttatatact
288	T25 L20b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCCacaaacacaagagtataagtataatt tcag
289	T25 L30b (I80-10) [Control]	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCCacaaacacaagtcatactagtat aagtataatttcag
290	T50 L10b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA

		GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgacgcta
291	T50 L20b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgacgctagtatg agtat
292	T50 L30b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgacgctagtatg agtatagcttcacac
293	T75 L10b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgcgacgg
294	T75 L20b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgcgactgggtcg gagcg
295	T75 L30b	gctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCaacaaaaacaagctgcgactggcg gaagctcgctggagcg
296	T25 L30 I0-0	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAAcacaacacaaagtcataactgagtataagtataatttcag
297	T25 L30a I0-0	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACTC GACGGGAGCTACCCTAACGTCAGACGAGGGTAAAGAGAGA GTCCAATTCTCAAAGCCAATAcacaacacaaagtcataactgagtataat ttcag
298	T25 L30a I25-10	ctgtttaacACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaacacaaagtcataactgagtataa gtataatttcag
299	T25 L30a I25-20	cgtaatttatattgtaacACGGACTTAAATAATTGAGCCTTAAAGAAGA AATTCCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATC TAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAA TTAGTAAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGT GCAGAGACTCGACGGGAGCTACCCTAACGTCAGACGAGGGT AAAGAGAGAGTCCAATTCTCAAAGCCAATAcacaacacaaagtcataac

		tgagtataagtataaatttcag
300	T25 L30a I50-10	gctatgatccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTCAAGACGAGGGTAAAGA GAGAGTCCAATTCTCAAAGCCAATAcacaacacaagtcatactaggtataa gtataaatttcag
301	T25 L30a I50-20	cggatattatgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAG AAATTCCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAA TCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGT AATTAGTAAGTTAAACAATAGATGACTTACAACCTAATCGGAAG GTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAGG GTAAAGAGAGAGTCCAATTCTCAAAGCCAATAcacaacacaagtc atactgagtataagtataaatttcag
302	T25 L30a I80-20	cggtaggcgcgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAG GAAATTCCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAA ATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAG TAATTAGTAAGTTAAACAATAGATGACTTACAACCTAATCGGAAG GGTGCAGAGACTCGACGGGAGCTACCCTAACGTCAAGACGAG GGTAAAGAGAGAGTCCAATTCTCAAAGCCAATAcacaacacaagtc atactgagtataagtataaatttcag
303	T25 L30b I0- 0	ACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTCCTTTAAG TGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTTATAGA CAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGTAAGTT AACAATAGATGACTTACAACCTAATCGGAAGGTGCAGAGACTC GACGGGAGCTACCCTAACGTcacaacacaagtcatactgagtataagtataat ttcag
304	T25 L30b I25-10	ctgtttaatcACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTcacaacacaagtcatactgagtat aagtataaatttcag
305	T25 L30b I25-20	cgtaatttatattgtaatcACGGACTTAAATAATTGAGCCTTAAAGAAGAA AATTCCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATC TAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAAT TTAGTAAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGT GCAGAGACTCGACGGGAGCTACCCTAACGTcacaacacaagtcata actgagtataagtataaatttcag
306	T25 L30b I50-10	gctatgatccACGGACTTAAATAATTGAGCCTTAAAGAAGAAATTC TTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAATCTAGTT ATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGTAATTAGT AAGTTAAACAATAGATGACTTACAACCTAATCGGAAGGTGCAGA GACTCGACGGGAGCTACCCTAACGTcacaacacaagtcatactgagtat aagtataaatttcag
307	T25 L30b I50-20	cggatattatgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAG AAATTCCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAAA TCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAGT AATTAGTAAGTTAAACAATAGATGACTTACAACCTAATCGGAAG GTGCAGAGACTCGACGGGAGCTACCCTAACGTcacaacacaagtc atactgagtataagtataaatttcag
308	T25 L30b I80-20	cggtaggcgcgctgcgagccACGGACTTAAATAATTGAGCCTTAAAGAAG GAAATTCCTTTAAGTGGATGCTCTCAAACCTCAGGGAAACCTAA

		ATCTAGTTATAGACAAGGCAATCCTGAGCCAAGCCGAAGTAG TAATTAGTAAGTTAACAATAGATGACTTACAATAATCGGAA GGTGCAGAGACTCGACGGGAGCTACCCTAACGTCcacaacacaagt catatactgagtataagtataatttcag
--	--	---

[0686] In some embodiments, a spacer and 3' intron fragment is a spacer and intron fragment having sequences as listed in Table 23.

Table 24. CAR sequences

SEQ ID NO	CAR	Sequence
309	FMC63-4-1BB	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCTCTG GATCAGACAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGT GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACTCCAAGAGCCAGGT GTTCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGCTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG CACCTGTGTCTTCTCCACTGTTCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTCAAGAGA GGCCGGAAGAACTTCTTTATATATTCAAGCAGCCCTTTATGC GACCCGTTTACAGACTACCCAAGAGGAAGATGGATGCAGTTGCC GCTTTCCAGAAGAGGAGGAGGGCGGGTGCGAACTGtaa
310	FMC63-CD28	ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCTCTG GATCAGACAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGT

		GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACTCCAAGAGCCAGGT GTTCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGCTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG CACCTGTGTCTTCTCCACTGTTCCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTCCGAAGC AAGCGGAGCCGGCTGCTGCACTCCGACTACATGAACATGACC CCTAGACGGCCCGGACCAACCAGAAAGCACTACCAGCCTTAC GCTCCTCCTAGAGACTTCGCCCGCCTACCGGTCTaa
311	FMC63- CD28-zeta	ATGCTGCTGCTGGTACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCCTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCAGCTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCCTG GATCAGACAGCCTCTCGGAAAGGCCTGGAATGGTGCGGAGT GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACTCCAAGAGCCAGGT GTTCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGCTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG CACCTGTGTCTTCTCCACTGTTCCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTCCGAAGC AAGCGGAGCCGGCTGCTGCACTCCGACTACATGAACATGACC CCTAGACGGCCCGGACCAACCAGAAAGCACTACCAGCCTTAC GCTCCTCCTAGAGACTTCGCCCGCCTACCGGTCCAGAGTGAAG TTCAGCAGATCCGCGGATGCTCCCGCCTATCAGCAGGGCCAA AACCAGCTGTACAACGAGCTGAACCTGGGGAGAAGAGAAGA GTACGACGTGCTGGACAAGCGGAGAGGCAGAGATCTGAAA TGGGCGGCAAGCCAGACGGAAGAATCCTCAAGAGGGCCTG TATAATGAGCTGCAGAAAGACAAGATGGCCGAGGCCTACAG CGAGATCGGAATGAAGGGCGAGCGCAGAAGAGGCAAGGGAC ACGATGGACTGTACCAGGGACTGAGCACCGCCACCAAGGATA CCTATGACGCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
312	FMC63-zeta	ATGCTGCTGCTGGTACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTG GTATCAGCAAAAGCCCGACGGCACCCTGAAGCTGCTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC

		<p> TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTGG CGAGGGATCTACCAAGGGCGAAGTGAAACTGCAAGAGTCTG GCCCTGGACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTG TACAGTCAGCGGAGTGTCCCTGCCTGATTACGGCGTGTCTTG GATCAGACAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGT GATCTGGGGCAGCGAGACAACCTACTACAACAGCGCCCTGAA GTCCCGGCTGACCATCATCAAGGACAACCTCCAAGAGCCAGGT GTTCTGAAGATGAACAGCCTGCAGACCGACGACACCGCCAT CTACTATTGCGCCAAGCACTACTACTACGGCGGCAGTACGC CATGGATTATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCT GCCGCCGCTATCGAAGTGATGTACCCTCCTCCTTACCTGGACA ACGAGAAGTCCAACGGCACCATCATCCACGTGAAGGGCAAG CACCTGTGTCTTCTCCACTGTTCCCGGACCTAGCAAGCCTT TCTGGGTGCTCGTTGTTGTTGGCGGCGTGCTGGCCTGTTACAG CCTGCTGGTTACCGTGGCCTTCATCATCTTTTGGGTGAGAGTG AAGTTCAGCAGATCCGCCGATGCTCCCGCCTATCAGCAGGGC CAAAACCAGCTGTACAACGAGCTGAACCTGGGGAGAAGAGA AGAGTACGACGTGCTGGACAAGCGGAGAGGCAGAGATCCTG AAATGGGCGGCAAGCCCAGACGGAAGAATCCTCAAGAGGGC CTGTATAATGAGCTGCAGAAAGACAAGATGGCCGAGGCCTAC AGCGAGATCGGAATGAAGGGCGAGCGCAGAAGAGGCAAGGG ACACGATGGACTGTACCAGGGACTGAGCACCGCCACCAAGG ATACCTATGACGCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa </p>
313	CircKymriah - Q388	<p> ATGGCTCTCCCGGTCACAGCCCTTCTCCTGCCCTGGCACTCT TGCTGCATGCGGCACGACCCGACATCCAGATGACCCAGACCA CAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCATCA GCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTGGT ATCAGCAAAAAGCCCGACGGCACCGTGAAAGCTGCTGATCTACC ACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTCTG GCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAACC TGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGGCA ACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGAAA TCACCGGTGGAGGTGGTTCTGGCGGAGGGGGATCTGGTGGAG GCGGTTCAAGAGTGAAACTGCAAGAGTCTGGCCCTGGACTGG TGGCCCCATCTCAGTCTCTGAGCGTGACCTGTACAGTCAGCG GAGTGTCCCTGCCTGATTACGGCGTGTCTGGATCAGACAGC CTCCTCGGAAAAGGCTTGAATGGCTGGGAGTGATCTGGGGCA GCGAGACAACCTACTACAACAGCGCCCTGAAGTCCCGGCTGA CCATCATCAAGGACAACCTCCAAGAGCCAGGTGTTCTGAAGA TGAACAGCCTGCAGACCGACGACACCGCCATCTACTATTGCG CCAAGCACTACTACTACGGCGGCAGCTACGCCATGGATTATT GGGGCCAGGGCACCAGCGTGACCGTTTCTTCTACCACAACGC CCGCCCCGCGACCGCCTACTCCCGCTCCCACAATTGCATCACA ACCCCTGTCTTTGAGACCCGAAGCTTGTCGACCAGCTGCCGGT GGCGCGGTTACACGCGGGGGCTCGATTTCGCCTGTGATATA TATATATGGGCCCCATTGGCTGGAACATGCGGAGTATTGCTTC TGAGCCTGGTGATTACCTCTACTGTAAAGAGAGGCGGGAAGA AACTTCTTTATATATTCAAGCAGCCCTTTATGCGACCCGTTCA GACTACCCAAGAGGAAGATGGATGCAGTTGCCGCTTTCAGA AGAGGAGGAGGGCGGGTGCGAACTGAGAGTGAAAGTTACGA GATCCGCGGATGCTCCCGCTATCAGCAGGGGCAAAACCAGC TGTACAACGAGCTGAACCTGGGGAGAAGAGAAGAGTACGAC </p>

		GTGCTGGACAAGCGGAGAGGCAGAGATCCTGAAATGGGCGG CAAGCCCAGACGGAAGAATCCTCAAGAGGGCCTGTATAATGA GCTGCAGAAAGACAAGATGGCCGAGGCCTACAGCGAGATCG GAATGAAGGGCGAGCGCAGAAGAGGCAAGGGACACGATGGA CTGTACCAGGGACTGAGCACCGCCACCAAGGATACCTATGAC GCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
314	CircKymriah - K388	ATGGCTCTCCCGGTACAGCCCTTCTCCTGCCCTGGCACTCT TGCTGCATGCGGCACGACCCGACATCCAGATGACCCAGACCA CAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCATCA GCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACTGGT ATCAGCAAAAAGCCCGACGGCACCGTGAAGCTGCTGATCTACC ACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTCTG GCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAACC TGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGGCA ACACCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGAAA TCACCGGTGGAGGTGGTTCTGGCGGAGGGGGATCTGGTGGAG GCGGTTTCAAGAAGTGAAGTCAAGAGTCTGGCCCTGGACTGG TGGCCCCATCTCAGTCTCTGAGCGTGACCTGTACAGTCAGCG GAGTGTCCCTGCCTGATTACGGCGTGTCTGGATCAGACAGC CTCCTCGGAAAGGCCTGGAATGGCTGGGAGTGATCTGGGGCA GCGAGACAACCTACTACAACAGCGCCCTGAAGTCCCGGCTGA CCATCATCAAGGACAACCTCCAAGAGCCAGGTGTTCTTGAAGA TGAACAGCCTGCAGACCGACGACACCGCCATCTACTATTGCG CCAAGCACTACTACTACGGCGGCAGCTACGCCATGGATTATT GGGGCCAGGGCACCAGCGTGACCGTTTCTTCTACCACAACGC CCGCCCCGCGACCGCCTACTCCCGCTCCCACAATTGCATCACA ACCCCTGTCTTTGAGACCCGAAGCTTGTCGACCAAGTCCCGGT GGCGCGGTTTACACGCGGGGGCTCGATTTTCGCTGTGATATA TATATATGGGCCCCATTGGCTGGAACATGCGGAGTATTGCTTC TGAGCCTGGTGATTACCTCTACTGTAAGAGAGGGCCGGAAGA AACTTCTTTATATATTCAAGCAGCCCTTTATGCGACCCGTTCA GACTACCCAAGAGGAAGATGGATGCAGTTGCCGCTTTCCAGA AGAGGAGGAGGGCGGGTGCGAACTGAGAGTGAAGTTCAGCA GATCCGCGGATGCTCCCGCTATAAGCAGGGGCCAAAACCAGC TGTACAACGAGCTGAACCTGGGGAGAAGAGAAGAGTACGAC GTGCTGGACAAGCGGAGAGGCAGAGATCCTGAAATGGGCGG CAAGCCCAGACGGAAGAATCCTCAAGAGGGCCTGTATAATGA GCTGCAGAAAGACAAGATGGCCGAGGCCTACAGCGAGATCG GAATGAAGGGCGAGCGCAGAAGAGGCAAGGGACACGATGGA CTGTACCAGGGACTGAGCACCGCCACCAAGGATACCTATGAC GCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa
315	CircM971 - CD22	ATGCTGCTGCTGGTACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCAGGTTCAACTCCAGCAGTC TGGTCCCGGCCTCGTTAAACCAAGCCAGACTTTTGTCTCTTACC TGTGCTATCAGTGGCGATAGCGTGTCTAGTAATTCAGCCGCAT GGAAGTGGATCCGACAATCACCGAGTAGGGGACTTGAATGGC TGGGTAGAACCTATTACCGGTCCAAATGGTACAATGACTATG CAGTGTCTGTAAAAGCAGGATCAGATCAACCCTGATACGT CTAAAAACCAAGTTTTCTCTGCAACTTAATAGTGTGACCCCTGA AGACACCGCTGTGTATTACTGTGCACGGGAGGTTACCGGTGA TCTTGAAGATGCTTTTGATATATGGGGCCAAGGTACGATGGT CACGGTGTCTAGTgggggaggcggcagcGACATACAGATGACGCAG AGCCCATCCAGTCTCTCCGCGTCTGTTGGTGACAGAGTGACTA TTACATGTAGGGCGTCTCAGACCAATTTGGTCTTACCTCAATTG GTATCAACAGCGACCAAGGCAAAGCACCGAACTTGCTCATTTA

		<p>CGCTGCCAGCTCACTCCAAAGTGGTGTGCCGTCCAGATTTAGT GGTAGGGGCAGTGGCACTGATTTCACTCTGACTATTTCAAGTC TTCAAGCTGAGGATTTTGGCACATACTACTGCCAGCAAAGTT ACTCAATACCTCAGACTTTTGGACAGGGGACAAAATTGGAGA TTAAAtccggaACCACAACGCCCCGCCCCGCGACCGCTACTCCC GCTCCCACAATTGCATCACAACCCCTGTCTTTGAGACCCGAA GCTTGTGCGACCAGCTGCCGGTGGCGCGGTTACACGCGGGGG CTCGATTTTCGCTGTGATATATATATATATGGGCCCCATTGGCTG GAACATGCGGAGTATTGCTTCTGAGCCTGGTGATTACCCCTCTA CTGTAAGAGAGGGCCGGAAGAACTTCTTTATATATTCAAGCA GCCCTTTATGCGACCCGTTTCACTACCCAAAGAGGAAGATGG ATGCAGTTGCCGCTTTCCAGAAGAGGAGGAGGGCGGGTGCGA ACTGAGAGTGAAGTTTCACTCAGCAGATCCGCGCGATGCTCCCGCTA TAAGCAGGGGCCAAAACCAGCTGTACAACGAGCTGAACCTGG GGAGAAGAGAAGAGTACGACGTGCTGGACAAGCGGAGAGGC AGAGATCCTGAAATGGGCGGCAAGCCCAGACGGAAGAATCC TCAAGAGGGCCTGTATAATGAGCTGCAGAAAGACAAGATGG CCGAGGCCCTACAGCGAGATCGGAATGAAGGGCGAGCGCAGA AGAGGCAAGGGACACGATGGACTGTACCAGGGACTGAGCAC CGCCACCAAGGATACCTATGACGCCCTGCACATGCAGGCCCT GCCTCCAAGAtaa</p>
316	CircCD19_22 Bispecific 29	<p>ATGCTGCTGCTGGTACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCGACATCCAGATGACCCAGAC CACAAGCAGCCTGTCTGCCAGCCTGGGCGATAGAGTGACCAT CAGCTGTAGAGCCAGCCAGGACATCAGCAAGTACCTGAACCTG GTATCAGCAAAAAGCCCGACGGCACCGTGAAGCTGTGATCTA CCACACCAGCAGACTGCACAGCGGCGTGCCAAGCAGATTTTC TGGCAGCGGCTCTGGCACCGACTACAGCCTGACAATCAGCAA CCTGGAACAAGAGGATATCGCTACCTACTTCTGCCAGCAAGG CAACACCCCTGCCTTACACCTTTGGCGGAGGCACCAAGCTGGA AATCACCGggcgaggagatccCAGGTTCAACTCCAGCAGTCTGGTC CCGGCCTCGTTAAACCAAGCCAGACTTTGTCTCTTACCTGTGC TATCAGTGGCGATAGCGTGTCTAGTAATTCAGCCGCATGGAA CTGGATCCGACAATCACCGAGTAGGGGACTTGAATGGCTGGG TAGAACCTATFACCGGTCCAAATGGTACAATGACTATGCAGT GTCTGTAAAAAGCAGGATCACGATCAACCCTGATACGTCTAA AAACCAGTTTTCTCTGCAACTTAATAGTGTGACCCCTGAAGAC ACCGCTGTGTATTACTGTGCACGGGAGGTTACCGGTGATCTTG AAGATGCTTTTGATATATGGGGCCAAGGTACGATGGTCACGG TGTCTAGTGGCTCTACAAGCGGCAGCGGCAAACCTGGATCTG GCGAGGGATCTACCAAGGGCGACATACAGATGACGCGAGAGC CCATCCAGTCTCTCCGCGTCTGTTGGTGACAGAGTGACTATTA CATGTAGGGCGTCTCAGACCATTTGGTCTTACCTCAATTGGTA TCAACAGCGACCAGGCAAAGCACCGAACTTGCTCATTTACGC TGCCAGCTCACTCCAAAGTGGTGTGCCGTCCAGATTTAGTGGT AGGGGCAGTGGCACTGATTTCACTCTGACTATTTCAAGTCTTC AAGCTGAGGATTTTGGCACATACTACTGCCAGCAAAGTTACT CAATACCTCAGACTTTTGGACAGGGGACAAAATTGGAGATTA AAgggggagggcggcagcGAAGTGAACTGCAAGAGTCTGGCCCTGG ACTGGTGGCCCCATCTCAGTCTCTGAGCGTGACCTGTACAGTC AGCGGAGTGTCCCTGCCTGATTACGGCGTGTCTGGATCAGA CAGCCTCCTCGGAAAGGCCTGGAATGGCTGGGAGTGATCTGG GGCAGCGAGACAACCTACTACAACAGCGCCCTGAAGTCCCGG CTGACCATCATCAAGGACAACCTCAAGAGCCAGGTGTTCTCTG AAGATGAACAGCCTGCAGACCGACGACACCGCCATCTACTAT</p>

		<p> TCGGCCAAGCACTACTACTACGGCGGCAGCTACGCCATGGAT TATTGGGGCCAGGGCACCAGCGTGACCGTTTCTTCTTccggaACC ACAACGCCCCGCCCCGCGACCGCCTACTCCCGCTCCCACAATT GCATCACAACCCCTGTCTTTGAGACCCGAAGCTTGTGACCA GCTGCCGGTGGCGCGGTTACACGCGGGGGCTCGATTTCGCC TGTGATATATATATATATGGGCCCCATTGGCTGGAACATGCGGA GTATTGCTTCTGAGCCTGGTGATTACCCTCTACTGTAAGAGAG GCCGGAAGAACTTCTTTATATATTCAAGCAGCCCTTTATGCG ACCCGTTTCAGACTACCCAAGAGGAAGATGGATGCAGTTGCCG CTTTCCAGAAGAGGAGGAGGGCGGGTGCGAACTGAGAGTGA AGTTCAGCAGATCCGCGGATGCTCCCGCCTATAAGCAGGGCC AAAACCAGCTGTACAACGAGCTGAACCTGGGGAGAAGAGAA GAGTACGACGTGTGACAAAGCGGAGAGGCAGATCCCTGA AATGGGCGGCAAGCCCAGACGGAAGAATCCTCAAGAGGGCC TGTATAATGAGCTGCAGAAAGACAAGATGGCCGAGGCCTACA GCGAGATCGGAATGAAGGGCGAGCGCAGAAGAGGCAAGGGA CACGATGGACTGTACCAGGGACTGAGCACCGCCACCAAGGAT ACCTATGACGCCCTGCACATGCAGGCCCTGCCTCCAAGAtaa </p>
317	CircCD19_22 Bispecific 30	<p> ATGCTGCTGCTGGTCACATCTCTGCTGCTGTGCGAGCTGCCCC ATCCTGCCTTTCTGCTGATCCCCAGGTTCAACTCCAGCAGTC TGGTCCCGGCTCGTTAAACCAAGCCAGACTTTGTCTCTTACC TGTGCTATCAGTGGCGATAGCGTGTCTAGTAATTCAGCCGCAT GGAACCTGGATCCGACAATCACCGAGTAGGGGACTTGAATGGC TGGGTAGAACCTATTACCGGTCCAAATGGTACAATGACTATG CAGTGTCTGTAAAAAGCAGGATCACGATCAACCCTGATACGT CTAAAAACCAAGTTTTCTCTGCAACTTAATAGTGTGACCCCTGA AGACACCGCTGTGTATTACTGTGCACGGGAGGTTACCGGTGA TCTTGAAGATGCTTTTGATATATGGGGCCAAGGTACGATGGT CACGGTGTCTAGTgggggagggcgagcGACATACAGATGACGCAG AGCCCATCCAGTCTCTCCGCGTCTGTTGGTGACAGAGTGACTA TTACATGTAGGGCGTCTCAGACCATTTGGTCTTACCTCAATTG GTATCAACAGCGACCAAGGCAAAGCACCGAACTTGCTCATTTA CGCTGCCAGCTCACTCAAAGTGGTGTGCCGTCCAGATTTAGT GGTAGGGGCAGTGGCACTGATTTCACTCTGACTATTTCAAGTC TTCAAGCTGAGGATTTTGCCACATACTACTGCCAGCAAAGTT ACTCAATACCTCAGACTTTTGGACAGGGGACAAAATTGGAGA TTAAAGGGGGAGGCGGATCCGGCGGTGGTGGCTCCGGCGGTG GTGGTTCTGGAGGCGGCGGAAGCGGTGGGGGTGGTAGCGAC ATCCAGATGACCCAGACCACAAGCAGCCTGTCTGCCAGCCTG GGCGATAGAGTGACCATCAGCTGTAGAGCCAGCCAGGACATC AGCAAGTACCTGAACTGGTATCAGCAAAAGCCCCGACGGCACC GTGAAGCTGCTGATCTACCACACCAGCAGACTGCACAGCGGC GTGCCAAGCAGATTTTCTGGCAGCGGCTCTGGCACCGACTAC AGCCTGACAATCAGCAACCTGGAACAAGAGGATATCGCTACC TACTTCTGCCAGCAAGGCAACACCCTGCCTTACACCTTTGGCG GAGGCACCAAGCTGGAAATCACCGGCTCTACAAGCGGCAGC GGCAAACTGGATCTGGCGAGGGATCTACCAAGGGCGAAGT GAAACTGCAAGAGTCTGGCCCTGGACTGGTGGCCCCATCTCA GTCTCTGAGCGTGACCTGTACAGTCAGCGGAGTGTCCCTGCCT GATTACGGCGTGTCTGGATCAGACAGCCTCCTCGAAAGGC CTGGAATGGCTGGGAGTGATCTGGGGCAGCGAGACAACCTAC TACAACAGCGCCCTGAAGTCCCGGCTGACCATCATCAAGGAC AACTCCAAGAGCCAGGTGTTCTGAAGATGAACAGCCTGCAG ACCGACGACACCGCCATCTACTATTGCGCCAAGCACTACTAC TACGGCGGCAGCTACGCCATGGATTATTGGGGCCAGGGCACC </p>

		AGCGTGACCGTTTCTTC ^{Tccgga} ACCACAACGCCCGCCCCGCGA CCGCCTACTCCCGCTCCCACAATTGCATCACAACCCCTGTCTT TGAGACCCGAAGCTTGTGCGACCAGCTGCCGGTGGCGCGGTTT ACACGCGGGGGCTCGATTTGCGCTGTGATATATATATATGGG CCCCATTGGCTGGAACATGCGGAGTATTGCTTCTGAGCCTGGT GATTACCTCTACTGTAAGAGAGGGCCGGAAGAACTTCTTTA TATATTCAAGCAGCCCTTTATGCGACCCGTTTCCAGACTACCCAA GAGGAAGATGGATGCAGTTGCCGCTTTCCAGAAGAGGAGGA GGGCGGGTGCGAAGTGAAGTTCAGCAGATCCGCCG ATGCTCCCGCTATAAGCAGGGCCAAAACCAGCTGTACAACG AGCTGAACCTGGGGAGAAGAGAAGAGTACGACGTGCTGGAC AAGCGGAGAGGCAGAGATCCTGAAATGGGCGGCAAGCCCAG ACGGAAGAATCCTCAAGAGGGCCTGTATAATGAGCTGCAGAA AGACAAGATGGCCGAGGCCTACAGCGAGATCGGAATGAAGG GCGAGCGCAGAAGAGGCAAGGGACACGATGGACTGTACCAG GGACTGAGCACCGCCACCAAGGATACCTATGACGCCCTGCAC ATGCAGGCCCTGCCTCCAAGAtaa
318	FMC63 CD28 anti-CD19 human CAR	MLLLVTSLLLCELPHPAFLIPDIQMTQTTSSLSASLGDRVTISCR ASQDISKYLWYQQKPDGTVKLLIYHTSRLHSGVPSRFSGSGSG TDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGTKLEITGSGSG KPGSGEGSTKGEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYG VSWIRQPPRKGLEWLGVIWGSETTYNSALKSRLTIHKDNSKSQV FLKMNSLQTDDBTAIYYCAKHYYYGGSYAMDYWGQGTSTVTVSS AAAIEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPFPGPSKPFVW LVVVGGLACYSLLVTVAFIHFWVRSKRSRLHSDYMNMTPRRP GPTRKHYQPYAPPRDFAAYRSRVKFSRSADAPAYQQGQNQLYN ELNLGRREEYDVLDRRGRDPEMGGKPRRKNPQEGLYNELQKD KMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQ ALPPR
319	FMC63 4- 1BB anti- CD19 human CAR	MALPVTALLPLALLLHAARPDQMTQTTSSLSASLGDRVTISCR ASQDISKYLWYQQKPDGTVKLLIYHTSRLHSGVPSRFSGSGSG TDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGTKLEITGGGSG GGGSGGGGSEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVS WIRQPPRKGLEWLGVIWGSETTYNSALKSRLTIHKDNSKSQVFL KMNSLQTDDBTAIYYCAKHYYYGGSYAMDYWGQGTSTVTVSSTT TPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACDIYI WAPLAGTCGVLLLSLVITLYCKRGRKKLLYIFKQPFMRPVQTTQ EEDGCSCRFPEEEEGGCELRVKFSRSADAPAYKQGQNQLYNELN LGRREEYDVLDRRGRDPEMGGKPRRKNPQEGLYNELQKDKM AEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQAL PPR
320	Anti-BCMA human CAR	MALPVTALLPLALLLHAARPDIVLTQSPASLAVSLGERATINCR ASESVSVIGAHLIHWYQQKPGQPPKLLIYLASNLETGVPARFSGS GSGTDFTLTISSLQAEDAIIYYCQSRIFPRTFGGQTKLEIKGSTS GSGKPGSGEGSTKGQVQLVQSGSELKKPGASVKVSCKASGYTFT DYSINWVRQAPGQGLEWMGWINTETREPAYAYDFRGRFVFSLD TSVSTAYLQISSLKAEDTAVYYCARDYSYAMDYWGQGTSLVTVS SAAATTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLD FACDIYIWAPLAGTCGVLLLSLVITLYCKRGRKKLLYIFKQPFMR PVQTTQEEDGCSCRFPEEEEGGCELRVKFSRSADAPAYQQGQNQ LYNELNLGRREEYDVLDRRGRDPEMGGKPRRKNPQEGLYNEL QDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDAL HMQALPPR

321	1D3ζ mouse CD19 CAR; Caprine Kobuvirus IRES	MGVPTQLLGLLLLWITDAICDIQMTQSPASLSTSLGETVTIQCQA SEDIYSGLAWYQQKPGKSPQLLIYGASDLQDGVPSRFSGSGSGT QYSLKITSMTQTEDEGVYFCQQGLTYPRTFGGGKLELKGSGSG GGGSGGGGSEVQLQQSGAELVRPGTSVKLSCKVSGDTITFYIM HFVKQRPQGQLEWIGRIDPEDESTKYSEKFNKATLTADTSSNT AYLKLSSLTSEDATYFCIYGGYYFDYWGQGVMTVSSIEFMYP PPYLDNERSNGTIIHIKEKHLCHTQSSPKLFWALVVVAGVLFYCYG LLVTVALCVIWTNSRRNRGGQSDYMNMTPRRPGLTRKPYQPYA PARDFAAAYRPRAKFSRSAETAANLQDPNQLFNELNLGRREEFDV LEKKRARDPEMGGKQORRRNPQEGVYNALQKDKMAEAYSEIG TKGERRRGKGHDGLFQGLSTATKDTFDALHMQTLAPR
322	WTζ (wild type zeta) mouse CAR	MGVPTQLLGLLLLWITDAICDIQMTQSPASLSTSLGETVTIQCQA SEDIYSGLAWYQQKPGKSPQLLIYGASDLQDGVPSRFSGSGSGT QYSLKITSMTQTEDEGVYFCQQGLTYPRTFGGGKLELKGSGSG GGGSGGGGSEVQLQQSGAELVRPGTSVKLSCKVSGDTITFYIM HFVKQRPQGQLEWIGRIDPEDESTKYSEKFNKATLTADTSSNT AYLKLSSLTSEDATYFCIYGGYYFDYWGQGVMTVSSIEFMYP PPYLDNERSNGTIIHIKEKHLCHTQSSPKLFWALVVVAGVLFYCYG LLVTVALCVIWTNSRRNRGGQSDYMNMTPRRPGLTRKPYQPYA PARDFAAAYRPRAKFSRSAETAANLQDPNQLYNELNLGRREEYD VLEKKRARDPEMGGKQORRRNPQEGVYNALQKDKMAEAYSEI GTKGERRRGKGHDGLYQGLSTATKDTYDALHMQTLAPR

[0687] In some embodiments, a CAR is encoded by a nucleotide sequence as listed in Table 24.

Table 25. CAR domain sequences.

SEQ ID NO	Protein	Sequence
318	4-1BB	KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEEGGCEL
319	CD3ζ intracellular domain	RVKFSRSADAPAYKQGNQLYNELNLGRREEYDVLDRRRGRDPE MGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGHD GLYQGLSTATKDTYDALHMQALPPR
320	CD28 intracellular signaling domain	QVQLVQSGAEVEKPGASVKVSKASGYTFTDYYMHWVRQAPGQ GLEWMGWINPNSGGTNYAQKFQGRVTMTRDTSISTAYMELSRLR SDDTAVYYCASGWDFDYWGQGLTVTVSSGGGSGGGGSGGGGS GGGSDIVMTQSPSSLSASVGDRTITCRASQSIRYYLSWYQQKP GKAPKLLIYTASILQNGVPSRFSGSGSGTDFTLTISLQPEDFATYY CLQYTTTPDFGPGTKVEIK
321	FMC63 VH	EVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQPPRKGLE WLGVIWGSETTYNSALKSRLTIHKDNSKSQVFLKMNSLQTDGTA IYYCAKHYYYGGSYAMDYWGQGSVTVSS
322	FMC63 VL	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPDGTVK LLIYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDIATYFCQQGN TLPYTFGGGKLEIT

[0688] In some embodiments, a CAR domain encoded by an inventive polynucleotide has a sequence as listed in Table 25.

Table 26. PD-1 or PD-L1 sequences.

SEQ ID NO	Description	Sequence
323	Pembrolizumab heavy chain	QVQLVQSGVEVKKPGASVKVSCASGYTFTNYYMYWV RQAPGQGLEWMGGINPSNGGTNFKNEKFKNRVTLTDSST TTAYMELKSLQFDDTAVYYCARRDYRFDMGFDYWGQG TTVTVSSASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFP EPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPS SLGKTKYTCNVDPKPSNTKVDKRVESKYGPPCPPCPAPE FLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDSQEDPEV QFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQ DWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPREPQVY LPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE NYKTTTPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSV MHEALHNHYTQKSLSLGLK
324	Pembrolizumab light chain	EIVLTQSPATLSLSPGERATLSCRASKGVSTSGYSYLHWY QKPGQAPRLLIYLAAYLESQVPAFSGSGSGTDFTLTIS LEPEDFAVYYCQHSRDLPLTFGGGKVEIKRTVAAPSVFI FPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQS GNSQESVTEQDSKSTYLSSTLTLSKADYEKHKVYACE VTHQGLSSPVTKSFNRGEC
325	Nivolumab heavy chain	QVQLVESGGGVVQPGRSLRLDCKASGITFSNSGMHWVR QAPGKGLEWVAWIYDGSKRYADSVKGRFTISRDNK NTLFLQMNSLRAEDTAVYYCATNDDYWGQGLTVTVSSA STKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVT NSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGKTKY TCNVDPKPSNTKVDKRVESKYGPPCPPCPAPEFLGGPSV LFPKPKDTLMISRTPEVTCVVDVDSQEDPEVQFNWYVD GVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKE YKCKVSNKGLPSSIEKTISKAKGQPREPQVYTLPPSQEEM TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV LDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNH YTQKSLSLGLK
326	Nivolumab light chain	EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQKP GQAPRLLIYDASNRTGIPARFSGSGSGTDFTLTISSELE PEFAVYYCQSSNWPRTFGQGTVEIKRTVAAPSVFIFPPS DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNS QESVTEQDSKSTYLSSTLTLSKADYEKHKVYACEVTH QGLSSPVTKSFNRGEC
327	Atezolizumab heavy chain	EVQLVESGGGLVQPGGSLRLSCAASGFTFSDSWIHVVRQ APGKGLEWVAWISPYGGSTYYADSVKGRFTISADTSKNT AYLQMNSLRAEDTAVYYCARRHWPGGFDYWGQGLTVT VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEP VTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPS SLGKTKYTCNVDPKPSNTKVDKKVEPKSCDKTHTCPPCPAPE LLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDSHEDPEV KFNWYVDGVEVHNAKTKPREEQYASTYRVVSVLTVLHQ DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVY LPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE NYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFC SVSMHEALHNHYTQKSLSLSPGK
328	Atezolizumab light chain	DIQMTQSPSSLSASVGDRVTITCRASQDVSTAVAWYQK PGKAPKLLIYSASFLYSGVPSRFSGSGSGTDFTLTIS SLQPEDFATYYCQYLYHPATFGQGTVEIKRTVAAPSVFI FPPS

		DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNS QESVTEQDSKDYSLSSSTLTLSKADYEKHKVYACEVTH QGLSSPVTKSFNRGEC
329	Avelumab heavy chain	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYIMMWVRQ APGKGLEWVSSIYPSGGITFYADTVKGRFTISRDNKNTL YLQMNSLRAEDTAVYYCARIKLGTVTTVDYWGQGLVT VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPV TVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLG TQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE LLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEV KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQ DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYT LPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN YKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFCFSVMH EALHNHYTQKSLSLSPGK
330	Avelumab light chain	QSALTQPASVSGSPGQSITISCTGTSSDVGGYNYVSWYQQ HPGKAPKLMYDVSNRPSGVSNRFSKSGNTASLTISGL QAEDEADYYCSSYTSSTRVFGTGTQVTVLGQPKANPTV TLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADGSPV KAGVETTKPSKQSNKYAASSYLSLTPEQWKSHRSYSCQ VTHEGSTVEKTVAPTECS
331	Durvalumab heavy chain	EVQLVESGGGLVQPGGSLRLSCAASGFTFSRYWMSWVR QAPGKGLEWVANIKQDGSEKYYVDSVKGRFTISRDNK NSLYLQMNSLRAEDTAVYYCAREGGWFGELAFDYWGQ GTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDY FPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPS SSSLGTQTYICNVNHKPSNTKVDKRVKPKSCDKTHTCPPC PAPEFEGGSPVFLFPPKPKDTLMISRTPEVTCVVDVSHED DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLT VLHQDWLNGKEYKCKVSNKALPASIEKTISKAKGQPREP QVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNG QPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFS CSVMHEALHNHYTQKSLSLSPGK
332	Durvalumab light chain	EIVLTQSPGTLSPGERATLSCRASQRVSSSYLAWYQQK PGQAPRLLIYDASSRATGIPDRFSGSGSGTDFTLTISRLEPE DFAVYYCQYGSFPWTFGQGTQVEIKRTVAAPSVFIFPPS DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNS QESVTEQDSKDYSLSSSTLTLSKADYEKHKVYACEVTH QGLSSPVTKSFNRGEC

[0689] In some embodiments, a cleavage site separating expression sequences encoded by an inventive polynucleotide has a sequence listed in Table 26.

Table 27. Cytokine sequences.

SEQ ID NO	Cytokine	Sequence
333	IL-2	APTSSSTKKTQLQLEHLLLDLQMLNGINNYKNPKLTRMLTFKFY MPKKATELKHLCLEELKPLEEVLNLAQSKNFHLRPRDLISNIN IVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIHSLT
334	IL-12A	RNLPVATPDGFMFCLHHSQNLRAVSNMLQKARQTLFYPCTSE EIDHEDITKDKTSTVEACLPLELTKNESCLNSRETSFITNGSCLASR KTSFMMALCLSSIYEDLKMYQVEFKTMNAKLLMDPKRQIFLDQN

		MLAVIDELMQALNFNSETVPQKSSLEEPDFYKTKIKLCILLHAFRIR AVTIDRVMSYLNAS
335	IL-12B	IWELKKDVYVVELDWYPDAPGEMVVLTCDTPEEDGITWTLQSS EVLGSGKTLTIQVKEFGDAGQYTCHKGGEVLSSHLLLLHKKEDGI WSTDILKDQKEPKNKTFLRCEAKNYSGRFTCWWTITISTDLTFSV KSSRGSSDPQGVTCGAATLSAERVGRDNKEYEYSVEQCEDSACPA AEESLPIEVMVDAVHKLKYENYTSSFFIRDIIKPDPPKNLQLKPLKN SRQVEVSWEYPDTWSTPHSYFSLTFCVQVQGKSKREKKDRVFTD KTSATVICRKNASISVRAQDRYSSSWSEWASVPCS
336	IL-7	DCDIEGKDGKQYESVLMVSIQLLDSMKEIGSNCLNNEFNFFKRHI CDANKEGMFLFRAARKLRQFLKMNSTGDFDLHLLKVSEGTTILLN CTGQVKGRKPAALGEAQPTKSLEENKSLKEQKKLNDLCFLKRL QEIKTCWNKILMGTEH
337	IL-10	SPGQGTQSENSCTHFPGNLPNMLRDLRDAFSRVKTFQMKDQLD NLLLKESLLEDFKGYLGCCQALSEMIQFYLEEVMQAEQDPDIKA HVNSLGENLKTLLRLRRLRRCHRFLPCENKSKAVEQVKNAFNKLQE KGIYKAMSEFDIFINYIEAYMTMKIRN
338	IL-15	NWVNVISDLKKIEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLL ELQVISLES GDASHDTVENLIILANNSLSSNGNVTESGCKECEELE EKNIKEFLQSFVHIVQMFINTS
339	IL-18	YFGKLESKLSVIRNLNDQVLFIDQGNRPLFEDMTSDSCRDNAPRTI FHSMYKDSQPRGMAVTISVKCEKISTLSCENKIISFKEMNPPDNIKD TKSDIIFQRSVPGHDNKMQFESSSYEGYFLACEKERDLFKLILKK EDELGDRSIMFTVQNE
340	IL-27beta	RKGPPAALTLPVQCRASRYPIAVDCSWTLPPAPNSTSPVSFIATY RLGMAARGHSWPCLQQTPTSTCTITDVQLFSMAPYVLNVTAHV PWGSSSFVPPFITEHIIKPDPEGVRLSPLAERQLQVQWEPGSPWF PEIFSLKYWIRYKRQGAARFHRVGPIEATSFILRAVRPRARYVQV AAQDLTDYGELSDWSLPATATMSLGK
341	IFNgamma	QDPYVKEAENLKKYFNAGHSDVADNGTLFLGILKNWKEESDRKI MQSQIVSFYFKLFKNFKDDQSIQKSVETIKEDMNVKFFNSNKKKR DDFEKLTNYSVTDLNVQRKAIHELIVMAELSPA AKTGKRKRSQ MLFRG
342	TGFbeta1	ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFC LGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIV YYVGRKPKVEQLSNMIVRSCKCS

[0690] In some embodiments, a cytokine encoded by an inventive polynucleotide has a sequence as listed in Table 27.

Table 28. Transcription factor sequences.

SEQ ID NO	Transcription factor	Sequence
343	FOXP3	MPNPRPGKPSAPSLALGPSGASPSWRAAPKASDLLGARGPGGT FQGRDLRGGAHASSSSLNPMPPSQLQLPTLPLVMVAPSGARLGP LPHLQALLQDRPHFMHQLSTVDAHARTPVLVHPLESPAMISLT PPTTATGVFSLKARPLPPGINVASLEWVSREPALLCTFPNPSAPR KDSTLSAVPQSSYPLLANGVCKWPGCEKVFEEDFLKHCQADH LLDEKGRAQCLLQREMVQSLEQQLVLEKEKLSAMQAHLAGKM ALTKASSVASSDKGCCIVAAGSQGPVVPWWSGPREAPDSLFAV RRHLWGSHGNSTFPEFLHNMDYFKFHNMRPPFTYATLIRWAILE

		APEKQRTLNEIYHWFTRMF AFFRNHPATWKN AIRHNLSLHKCFV RVESEKGAVWTVDELEFRKKRSQRPSRCSNPTPGP
344	FOXP3	MPNPRPGKPSAPSLALGPSGASPSWRAAPKASDLLGARGPGGT FQGRDLRGGAHASSSSSLNPMPPSQLQLPTLPLVMVAPSGARLGP LPHLQALLQDRPHFMHQLSTVDAHARTPVLQVHPLESPAMISLT PPTTATGVFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPR KDSTLSAVPQSSYPLLANGVCKWPGCEKVFEEDFLKHCQADH LLDEKGRAQCLLQREMOVQSLEQQLVLEKEKLSAMQAHLAGKM ALTKASSVASSDKGSCCIVAAGSQGPVVPWWSGPREAPDSLFAV RRHLWGSHGNSTFPEFLHNMDYFKFHNMRPPFTYATLIRWAILE APEKQRTLNEIYHWFTRMF AFFRNHPATWKN AIRHNLSLHKCFV RVESEKGAVWTVDELEFRKKR
345	FOXP3	GGAHASSSSSLNPMPPSQLQLPTLPLVMVAPSGARLGPLPHLQAL LQDRPHFMHQLSTVDAHARTPVLQVHPLESPAMISLTPPTTATG VFSLKARPGLPPGINVASLEWVSREPALLCTFPNPSAPRKDSTLS AVPQSSYPLLANGVCKWPGCEKVFEEDFLKHCQADHLLDEK GRAQCLLQREMOVQSLEQQLVLEKEKLSAMQAHLAGKMAITKA SSVASSDKGSCCIVAAGSQGPVVPWWSGPREAPDSLFAVRRHLW GSHGNSTFPEFLHNMDYFKFHNMRPPFTYATLIRWAILEAPEKQ RTLNEIYHWFTRMF AFFRNHPATWKN AIRHNLSLHKCFVRVESE KGAVWTVDELEFRKKR
346	STAT5B	MAVWIAQQQLQGEALHQMQUALYGOHFPIEVRHYLSQWIESQA WDSVDLDNPQENIKATQLLEGLVQELQKKAHQVGEDGFLKI KLGHYATQLQNTYDRCPMELVRCIRHILYNEQRLVREANNSSP AGSLADAMSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFII QYQESLRIQAQFGPLAQLSPQERLSRETALQKQVSLEAWLQRE AQTLLQYRVELAEKHQKTLQLLRKQQTILDDDELQWKRRQQLA GNGGPEGSLLDVLQSWCEKLAHWNQNRQQIRRAEHLCCQLPIPG PVEEMLAEVNATITDIISALVTSTFHEKQPPQVLKQTQKFAATVR LLVGGKLNVMHNPPQVKATIISEQQAQSLLKNENTRNDYSGEIL NNCCVMEYHQATGTL SAHFRNMSLKRIKRSRRGAESVTEEFK TILFESQFSVGGNELVFQVKTLSPVVVIVHGSQDNNATATVLW DNAFAEPGRVPFAVPDKVLWPQLCEALNMKFAEVQSNRGLTK ENLVFLAQKLFNNSSSHLEDYSGLSVWSQFNRENLPGRNYTFW QWFDGVMELVKHLKPHWNDGAILGFVNKQQAHDLLINKPDG TFLLRFSDEIGGITIAWKFDSDQERMFWNLMPTTTRDFSIRSLADR LGDNLNLIYVFPDRPKDEVYSKYYPVPCESATAKAVDGYVKPQ IKQVVPEFVNASADAGGGSATYMDQAPSPAVCPQAHYNMYPQ NPDSVLDTDGDFLEDTMDVARRVEELLGRPMDSQWIPHAQS
347	HELIOS	METEADGYITCDNELSPEREHSNMAIDLTSSTPNGQHASPSHMT STNSVKLEMQSDEECDRKPLSREDEIRGHDEGSSLEEPLIESSEVA DNRKVQELQGEQGIRLPNGKLLKCDVCGMVCIGPNVLMVHKRSH TGERPFHCNQCASFTQKGNLLRHIKLSGEEKPFKCPFCSYACRR RDALTGHLRTHSVGKPHKCNCGRSYKQRSSLEEHEKCHNYL QNVSMEEAGQVMSSHVPPMEDCKEQEPIMDNNISLVPFERPAVI EKL TGNMGKRKSSTPQKFEVGEKLMRFSYPDIHFDMLNLTYEKEA ELMQSHMMDQAINNAITYLGAEALHPLMQHPPSTIAEVAPVISS AYSQVYHPNRIERPISRETADSHENNMDGPISLIRPKSRPQEREAS PSNSCLDSTDSESHDDHQSYQGH PALNPKRKQSPAYMKEDVK ALDTTKAPKGSCLKDIYKVFNGEGEQIRAFKCEHCRVFLDHVVMY TIHMGCHGYRDPLECNICGYRSQDRYEFSSHIVRGEHTFH

[0691] In some embodiments, a transcription factor encoded by an inventive polynucleotide has a sequence as listed in Table 28.

Table 29. Additional Accessory Sequences

SEQ ID NO	IRES	Sequence
390	CK 3' UTR Scr	ccctgcagccgtcacgtaagtttgaagtiaccgcatacagcctctgcttccagcgctccaatt cctgttcttattgttccctccaggcggttacgcgtgacgaagaactgtgtcgcagctaccacattatt ccggagccttcattctcgggctctgacgt
391	CK 3' UTR S2M	ggagaccgcgccacgcccagtaggacgagggtagagtctcc
392	CK 3' UTR	gacaccaggatcactcttctctgacccgcctgtgtagaataagactcatgcttccctaagacctgg atttcttccaggcactttcaccgctgcccctgctcctcagtggaactgcacccaggaggcggtc ctgactgtcctttactttctattctggattgc
393	CK 5' UTR 1	AAACCCCCCTAAGCCGCGCCGCCGCCACC
394	CK 5' UTR 2	CCCCCCCCAACCCGTCACG
395	CK 5' UTR 3	GTCACG
396	SZ1 3' UTR Scr	tctgcgcactcgtaatcagtactaaccccccttctgaggacactatgcgataatcgatccgcttttc accgccttcggaattttattacctaactgatcctggagctctcttctggtttcacggaggcctccgcc ca
397	SZ1 S2M	ggagaccgcgccacgcccagtaggacgagggtagagtctcc
398	SZ1 3' UTR	ccccctgaacccccgccccaggttcagttctcttccctctgtctctcatgttgatacaaaagac cctttgtggaccctaagccatgtagtgtctgctcctccttcagttgtgaatattggtttctgttaatca ca
399	SZ1 5' UTR 1	AAACCCCCCTAAGCCGCGCCGCCGCCACC
400	SZ1 5' UTR 2	CCCCCCCCAACCCGTCACG
401	SZ1 5' UTR 3	GTCACG
402	UTR1	gTcacG
403	UTR2	AATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGC CACC
404	UTR3	cgaactagtagttcttctgtgtcccccacagactcagagagaacccgccacc
405	UTR4	Agccacc
406	STOP1	tgatAGctAaCtaG
407	STOP2	tagtAGctAaCtaG
408	STOP3	tGatGActGaGtGA
409	STOP4	tagtagctagGtag
410	STOP5	taa
411	STOP6	taatagCtaaCtag
412	STOP7	taaCtagCtaaCtag
718	CCL1	MQIITTAALVCLLLAGMWPEDVDSKSMQVPFSRCCFSFAEQEIP LRAILCYRNTSSICSNEGLIFKLKRGKEACALDITVGWVQRHRK MLRHCPSKRK
719	CCL2	MKVSAALLCLLLIAATFIPQGLAQPDAINAPVTCCYNFTNRKIS VQLAS
720	CCL3	MQVSTAALAVLLCTMALCNQFSASLAADTPTACCFSYTSRQI PQNFIADYFETSSQCSKPGVIFLTRSRQVCADPSEEWVQKYV SDLELSA
721	CCL4	MKLCVTVLSLLMLVAAFCSPALAPMGSDPPTACCFSTARK LPRNFVVDYYETSSLCSPAVVFQTKRSKQVCADPSESWWQE YVYDLELN
722	CCL5	MKVSAAALAVILIATALCAPASAPYSSDTPCCFAYIARPLPR AHIKEYFYTSKCSNPVVFVTRKNRQVCANPEKKWVREYIN SLEMS

723	CCL6	MRNSKTAISFFILVAVLGSQAAGLIQEMEKEDEDRRYNPPIIHQGFQ DTSSDCCFSYATQIPCKRFIYFPTSGGCIKPGHIFISRRGTQVCA DPSDRRVQRCLSTLKQGPGRSGNKVIA
724	CCL7	MKASAALLCLLLTAAAFSPQGLAQPVGINTSTTCCYRFINKKI PKQRLESYRRITSSHCPCREAVIFKTKLDKEICADPTQKWVQDF MKHLDKKTQTPKL
725	CCL8	MKVSAALLCLLLMAATFSPQGLAQPDVSIPITCCFNVINRKIP IQRLESYTRITNIQCPKEAVIFKTKRGKEVCADPKERWVRDSM KHLDQIFQNLKP
726	CCL9/CCL10	MKPFHTALSFLILTALGIWAQITHATETKEVQSSLKAQQGLEI EMFHMFGQDSSDCCLSYNSRIQCSRFIGYFPTSGGCTRPQIIFIS KRGFQVCANPSDRRVQRCIERLEQNSQPRTYKQ
727	CCL11	MKVSAALLWLLIAAAFSPQGLAGPASVPTTCCFNLANRKIPL QRLESYRRITSGKCPQKAVIFKTKLAKDICADPKKKWVQDSM KYLDQKSPTPKP
728	CCL12	MKISTLLCLLLIATTISPQVLAGPDAVSTPVTCCYNVVKQKIHV RKLKSYRRITSSQCPCREAVIFRTILDKEICADPKEKWVKNINH LDKTSQTFILEPSCLG
729	CCL13	MKVSALLCLLLMTAAFNPOGLAQPDALNPSTCCFTFSSKKI SLQRLKSYVITTSRCPQKAVIFRTKLKKEICADPKEKWVQNY MKHLGRKAHTLKT
730	CCL14	MKISVAAIPFLLITIALGKTTESSRGPYHPSECCFTYTTYKIPR QRIMDYETNSQCSKPGIVFITKRGHVCTNPSPDKWVQDYIKD MKEN
731	CCL15	MKVSVAALSCLMLVAVLGSQAQFINDAETELMMSKLPLENP VVLNSFHFAADCCTSYISQIPCSLMKSYFETSSECSKPGVIFLT KKGRQVCAKPSGPGVQDCMKKLPYSI
732	CCL16	MKVSEAALLVLLIITSASRSQPKVPEWVNTPSTCCLKYEYK VLPRLVVGYSKALNCHLPAIIFVTKRNREVCTNPNDWVQE YIKDPNLPLPTRNLSTVKIITAKNGQPQLLSQ
733	CCL17	MAPLKMLALVTLGLGASLQHIHAARGTNVGRECCLEYFKGAI PLRKLKTWYQTSSEDCSRDAIVFVTVQGRAICSDPNKRKVKNA VKYLQSLERS
734	CCL18	MKGLAAALLVVLCTMALCSCAQVGTNKLCLVYTSWQIPQ KFIVDYSETSPQCPKPGVILLTKRGRQICADPNKKWVQKYISD LKLNA
735	CCL19	MALLLALSLLVLWTSPAPTLSGTNDACCLSVTQKPIPGYIV RNFHYLLIKDGRVPAVVFTTLRGRQLCAPPDQPWVERIIRL QRTSAKMKRRSS
736	CCL20	MCCTKSLLLAALMSVLLHLGGESEAAASNFDCCCLGYTDRILH PKFIVGFTRQLANEGCDINAIIFHTKKKLSVCANPKQTWVKYI VRLLSKKVKNM
737	CCL21	MAQSLALSLLVLVAFGIPRTQGSDDGAQDCCLKYSQKIPAK VRSYRKQEPSLGCIPAILFLPRKRSQAEICADPKELWVQQL MQHLDKTPSPQKPAQGCRCRKGASKTGKKGKSGSKGCKRTER SQTPKGP
738	CCL22	MDRLQTALLVVLVLLAVALQATEAGPYGANMEDSVCCRDIY VRYRLPLRVVKHFYWTSDSCPRPGVVLLTFRDKEICADPRVP WVKMILNKLSQ
739	CCL23	MKVSVAALSCLMLVTALGSQARVTKDAETEFMMSKLPLENP VLLDRFHATSADCCISYTPRSIPCSLLESYFETNSECSPKGVIFL TKKGRRFCANPSDKQVQVCVRMLKLDTRIKTRKN

740	CCL24	MAGLMTIVTSLFLGVCACHHIPTGSVIPSPCCMFFVSKRIPE NRVVSQYLSSRSTCLKAGVIFTTKKGQFCGDPKQEWVQRY MKNLDAKQKKASPRARAVAVKGPVQRYPGNQTTT
741	CCL25	MNLWLLACLAVAGFLGAWAPAVHTQGVFEDCCLAYHYPIGW AVLRRAWTYRIQEVSGSCNLPAAIFYLPKRHRKVCGNPKSRE VQRAMKLLDARNKVFALHHTNTQTFQAGPHAVKKLSSGNSK LSSSKFSNPISSSKRNVSLLISANSGL
742	CCL26	MMGLSLASAVLLASLLSLHLGTATRGSDISKTCFQYSHKPLP WTWVRSYEFTSNSCSQRAVIFTTKRGKKVCTHPRKKWVQKYI SLLKTPKQL
743	CCL27	MKGPPFTFCSLLLLSLLSPDPTAAFLPPSTACCTQLYRKPLSD KLLRKVIQVELQEADGDCHLQAFVLHLAQRSICHPQNPSLSQ WFEHQERKLHGTLPKLNFGLRKMKG
744	CCL28	MQQRGLAIVALAVCAALHASEAILPIASSCCTEVSHHISRRLL RVNMCRIQRADGDCDLAAVILHVKRRRICVSPHNHTVKQWM KVQAACKNGKGNVCHRRKHHGKRNSNRAHQGKHETYGHK TPY
745	CXCL1	MARAALSAAPSNPRLLRVALLLLLLVAAGRRRAAGASVATELR CQCLQTLQGIHPKNIQSVNVKSPGPHCAQTEVIATLKNRKCAC LNPASPIVKKIJEKMLNSDKSN
746	CXCL2	MARATLSAAPSNPRLLRVALLLLLLVAASRRRAAGAPLATELR CQCLQTLQGIHLKNIQSVKVKSPGPHCAQTEVIATLKNQKA CLNPASPMVKKIJEKMLKNGKSN
747	CXCL3	MAHATLSAAPSNPRLLRVALLLLLLVAASRRRAAGASVATELR CQCLQTLQGIHLKNIQSVNVRSPGPHCAQTEVIATLKNKKAC LNPASPMVQKIIKILNKGSTN
748	CXCL4	MAYYEEIIFDFTDNSSEESGDFDVFQGPCERAPNYDFHKIFLP TVFGIIFILGIFGNGLVVVMGFQNKCKTSMTDKYRLHLSVAD LMFVLTLPFWAVDAASSWYFGGFLCKAVHMIYTVNLYSSVLI LAFISLDRLAVVRATATSSQATRKLLAGKVIYVCVWLPAAIL TVPDLVFASSFEVELDMGGSRMICQRIYPLESNHIWVAEFRFQ HILVGFVLPGLVILTCYCHISKLSQGSKGQALKRKALKTTVILI LCFFCCWLPYCAGIFVDTLNLEVIPHSCALEQGVMMKWISITE VLAYFHCCLNPILYAFGLGVKFKKSARNALTFSSRSSHKMLTKK RGAISSASTESESSSVLYS
749	CXCL5	MSLLSSRAARVPGPSSSLCALLVLLLLLTQPGPIASAGPAAVL RELRCVCLQTTQGVHPKMISNLQVFAIGPQCSKVEVVASLKN GKEICLDPEAPFLKKVIQKILDGGNKEN
750	CXCL6	MSLPSSRAARVPGPSGLCALLALLLLLLTPPGPLASAGPVS LTELCTCLRVTLRVNPKTIGKLQVFPAGPQCSKVEVVASLKN GKQVCLDPEAPFLKKVIQKILDGSGNKK
751	CXCL7	MSRLDTPPCNSARPLHALQVLLLLSLLLTALASSTKGQTKR NLAGKKEESLSDLYAELRCMCIKTTSGIHPKNIQSLEVIGKGT HCNQVEVIATLKDGRKICLDPDAPRIKKIVQKKLAGDESAD
752	CXCL8	MTSKLAVALLAFLISAALCEGAVLPRSAKELRCQCIKTYSKP FHPKFIKELRVIESGPHCANTEHVKLSDGRELCLDPKENWVQR VVEKFLKRAENS
753	CXCL9	MKKSGVLFLLGILLVLIGVQGTVPVVRKGRCSCISTNQGTIHLQ SLKDLKQFAPSPSCEKIEIATLKNQVQTCNPDSDADVKELIK WEKQVSQKKKQKNGKKHKKKVLKVRKSQRSRQKKT
754	CXCL10	MNQTAILICCLIFLTLSGIQGVPLSRTVRCTCISISNQPVNPSLE KLEIIPASQFCPRVEIATMKKKGEKRCLNPESKAIKNLLKAVS KERSKRSP

755	CXCL11	MSVKGMAIALAVILCATVVQGFPFMRGRCLCIGPGVKAVK VADIEKASIMYPSNNCDKIEVIITLKENKGQRCLNPKSKQARLI IKKVERKNF
756	CXCL12	MNAKVVVVLVLTALCLSDGKPVSLSYRCPFRFFESHVARA NVKHLKILNTPNCALQIVARLKNNNRQVCIDPKLKWIEYLE KALNKRFKM
757	CXCL13	MKFISTSLLLMLLVSSLSPVQGVLEVYYTSLRCRCVQESSVFIP RRFIDRIQILPRGNGCPRKEIIVWKKNKSIVCVDPQAEWIQRM MEVLRKRSSSTLPVPVFKRKIP
758	CXCL14	MSLLPRRAPVSMRLAAALLLLLLALYTARVDGSKCKCSRK GPKIRYSDVKKLEMKPKYPHCEEKMMVIITTSVSRYRGQEHCL HPKLQSTKRFIKWYNWNEKRRVYEE
759	CXCL15	MAAQGWSMLLLAVLNLGIFVRPCDTQELRCLCIQEHSEFIPLK LIKINIMVIFETIYCNRKEVIAVPKNGSMICLDPDAPWVKATVG PITNRFLEPDLKQKEFPFAMKLLYSVEHEKPLYLSFGREPENKRI FPFPIRETSRHFAADLAHNSDRNFLRDSSEVSLTGSDA
760	CXCL16	MGRDLRPGSRVLLLLLLLLLVYLTQPGNGNEGSVTGSCYCGK RISSDSPPSVQFMNRLRKHLRAYHRCLYYTRFQLLSWSVCGG NKDPWVQELMSCLDLKECGHAYSGIVAHQKHLPTSPISQA SEGASSDIHTPAQMLLSTLQSTQRPTLPVGLSSDKELTRPNET TIHTAGHSLAAGPEAGENQKQPEKNAGPTARTSATVPVLCCL AIIIFILTAALSYVLCKRRRRGQSPQSSPDLPVHYIPVAP
761	CXCL17	MKVLISLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRW LQEGGQECECKDWFLRAPRRKFMTVSGLPKKQPCDHFKN VKTRHQRHHRKPNKHSRACQQLKQCOLRSFALPL
762	XCL1	MRLILALLGICSLTAYIVEGVGSEVSDKRTCVSLLTQRLPVS IKTYTITEGSLRAVIFITKRGLKVCADPQATWVRDVRSMDRK SNTRNNMIQTKTGTQQTNTAVTLTG
763	XCL2	MRLILALLGICSLTAYIVEGVGSEVSHRRTCVSLLTQRLPVS IKTYTITEGSLRAVIFITKRGLKVCADPQATWVRDVRSMDRK SNTRNNMIQTKTGTQQTNTAVTLTG
764	CX3CL1	MAPISLSWLLRLATFCHLTVLLAGQHGHVTKCNITCSKMTSKI PVALLIHYQQNQASCGKRAIILETRQHRLFCADPKEQWVKDA MQHLDRQAAALTRNGGTFEKQIGEVKPRTPPAAGGMDES VLEPEATGESSSLEPTPSSQEAQALGTSPELPTGVTGSSGTRLPP TPKAQDGGPVGTLEFRVPPVSTAATWQSSAPHQPGPSLWAEA KTSEAPSTQDPSTQASTASSPAPEENAPSEGQRVWGGQSPRP ENSLEREEMGPVPAHTDAFQDWGPGSMAHVSVPVSSEGTPS REPVASGSWTPKAEPIHATMDPQRLGVLITPVPDAQAATR QAVGLLAFLGLLFCLGVAMFTYQSLQGCPKRMAGEMAEGLR YIPRSCGSNSYVLPV
765	Tbet	HLLWSKFNQHQTEMIITKQGRMFPLSFTVAGLEPTSHYRMF VDVVLVDQHHWRYQSGKWVQCGKAEGSMGPNRLYVHPDSP NTGAHWMRQEVSFGLKLTNNKGASNNVTQMIVLQSLHKY QPRLHIVEVNEGEPETVCNASNTHIFTQETQFIQAVTAYQNAEI TQLKIDNNPFAKGFRENFESMYASVDTSPSPPGPNCQLLGGD PYSPLLSNQYPVP
766	GATA3	MEVTADQPRWVSHHHPAVLNGQHPDTHHPGLSHSYMDAAQ YPLPEEVDVLFNIDGQGNHVPYGYNSVRATVQRYPPTHHGS QVCRPPLLHGSLPWLDGGKALGSHHTASPWNLSPFKSTSIHH GSPGPLSVYPPASSSSLSGGHASPFLFTFPPTPKDVSPDPSLST PGSAGSARQDEKECLKYQVPLPDSMKLESSHSRGSMTALGGA SSSTHHPITTYPPYVPEYSSGLFPSSLLGGSPTGFGCKSRPKAR SSTGRECVNCGATSTPLWRRDGTGHYLCNACGLYHKMNGQN

		RPLIKPKRRLSAARRAGTSCANCQTTTTTLWRRNANGDPVCN ACGLYYKLHNINRPLTMKKEGIQTRNRKMSSKSKCKKVHD SLEDFPKNSSFNPAALSRHMSSLSHISPFSSSHMLTTPTPMHP PSSLSFGPHHPSSMVTAMG
767	RORgt	MDRAPQRQHRASRELLAAKKTHTSQIEVIPCKICGDKSSGIHY GVITCEGCKGFFRRSQRCNAAYSCTRQQNCPIDRTSRNRCQHC RLQKCLALGMSRDAVKFGRMSKKQRDSLHAEVQKQLQQRQ QQQQEPVVKTPPAGAQQGADTLTYTLGLPDGQLPLGSSPDLPE ASACPPGLLKASGSGPSYSNNLAKAGLNGASCHLEYS PERGK AEGRESFYSTGSQLTPDRCGLRFEHRHPGLGELGQGPD SYGS PSFRSTPEAPYASLTEIEHLVQSVCKSYRETCQLRLEDLLQRS NIFSREEVTGYQRKSMWEMWERCAHHLTEAIQYVVEFAKRL SGFMELCQNDQIVLLKAGAMEVVLVRMCRAYNADNRTVFE GKYGGMELFRAIGCSELISSIFDFSHLSALHFSEDEIALYTAL VLINAHRLPGLQEKRKVEQLQYNLELAFHHHLCKTHRQSILAK LPPKGKLRSLCSQHVERLQIFQHLHPVVQAAPPLYKELFSTE TESPVGLSK
768	Cd25	MEPCLLMWGILTFITVSGYTTDLCDDDPPNLKHATFKALTYK TGTVLNCDCEGFRRISSYMHCTGNSSHASWENKCQCKSISPE NRKGKVTTKPEEQKGENPTMQSQTPPMDEVLDLVGHCREPP WEHENSRIYHFVVGQTLHYQCMQGFALHRGPAKSICKTIF GKTRWTQPLKCISESQFPDDEELQASTDAPAGRDTSSPFITTS TPDFHKHTEVATTMESFIFTTEYQIAVASCVLLISIVLVSGLT WQRRRRKSRTI
769	Skpl	MPSIKLQSSDGEIFEVDVEIAKQSVTIKTMLEDLGMDDEGDDD PVPLPNVNAAILKKVIQWCTHHKDDPPPPEDDENKEKRTDDIP VWDQEFLLKVDQGTLFELILAANYLDIKGLLDVTECKTVANMIK GKTPEEIRKTFNIKNDFTEEEEAQVRKENQWCEEK
770	Spy	MRKLTALFVASTLALGAANLAHAADTTTAAAPADAKPMMHH KGKFGPHQDMMFKDLNLTDQKQQIREIMKGQRDQMKRPPL EERRAMHDIIASDTFDKAKAEQIAKMEEQRKANMLAHMET QNKIYNILTPEQKKQFNANFEKRLTERPAAKGKMPATAE
771	Fkpa	MKSLFKVTLLATTMAVALHAPITFAAEAAKPATAADSKAAFK NDDQKSAYALGASLGRYMENSLKEQEKLGKLDKDQLIAGV QDAFADKSKLSDQEIEQTLQAFEARVKSSAQAKMEKDAADN EAKGKEYREKFAKEKGKVTSSSTGLVYQVVEAGKGEAPKDS TVVVNYKGTLDGKEFDNSYTRGEPLSFRLDGVIPGWTEGLK NIKKGGKIKLVIPPELAYGKAGVPGIPPNSTLVFDVELLDVKPA PKADAKPEADAKAADSACK
772	Sura	MKNWKTLLLGIAMIANSTFAAPQVVDKVAADVNNGVVLESD VDGLMQSVKLNAAQARQQLPDDATLRHQIMERLIMDQIILQM GQKMGVKISDEQLDQAIANIAKQNNMTLDQMRSLAYDGLN YNTYRNQIRKEMIISEVRNNEVRRRITILPQEVESLAQQVGNQ NDASTELNLSHILPLPENPTSDQVNEAESQARAIVDQARNGA DFGKLAIAHSADQQALNGGQMGWGRIQELPGIFAQALSTAKK GDIVGPIRSGVGFHILKVNDLRGESKNISVTEVHARHILLKPSPI MTDEQARVKLEQIAADIKSGKTTFAAAAKEFSQDPGSANQGG DLGWATPDIFDPAFRDALTRLNKGQMSAPVHSSFGWHLELL DTRNVDKTDAAQKDRAYRMLMNRKFSEEAASWMQEQRASA YVKILSN
773	Hsp60	MLRVNSKSSIKTFVRHLHSHKELKFGVEGRAALLKGVNTLADA VSVTLGPKGRNVLIEQQFGAPKITKDGVTVAKAITLEDKFEDL GAKLLQEVASKTNESAGDGTTSATVLGRSIFTESVKNVAAGC NPMDLRRGSQA AVEAVIEFLQKNKKEITTSEEIAQVATISANG

		DKHIGDLLANAMEKVGKEGVITVKEGKTLEDELEVTEGMKF DRGFISPYFITNTKTGKVEFENPLILLSEKKISSIQDILPSLELSN QTRRPLLIIAEDVDGEALAAACILNKLRGQVQVCAVKAPGFGD NRKNTLGDIAILSGGTVFTEELDIKPENATIEQLGSAGAVTITK EDTVLLNGEGSKDNLEARCEQIRSVIADVHTTEYEKEKLQERL AKLSGGVAVIKVGGASEVEVGEKKDRYEDALNATRAAVEEG ILPGGGTALIKATKILDEVKEKAVNFDQKLGVDTIRAAITKPA KRIHENAGEEGAVIVGKIYDEPEFNKGYSQKGEFTDMIAAGII DPFKVVKNGLVDSAGVASLLATTECAIVDAPQPKGSPAAPPA PGMGGMPGMF
774	Hsp70	MAPAVGIDLGTTYSCVGIFRDDRRIEIIHANDQGNRTTPSFVAFTD TERLIGDAAKNQVAMNPANTVFDKRLIGRKVFADPEVQADM KHFPFKITDKGGKPVQVEFKGETKEFTPEEISSMVLTKMRETA EAYLGTVNNNAVVTVPAYFNDSQRQATKDAGLIAGLNVLRRI NEPTAAAIAYGLDKKADGERNVLIFDLGGGTDFVSLLTIEEGIF EVKSTAGDTHLGGEDFDNRLVNHVSEFKRKFKKISPAERAR ALRRSPTACERAKRTLSSAAQTSIEIDSLYEGIDFYTSITRARFE ELCQDLFRSTMEPVERVLRDAKIDKSSVHEIVLVGGSTRIPRIQ KLVSDFENGKEPNKSINPDEAVAYGAAVQAAILSGDTSSKSTN EILLLDVAPLSLGIETAGGVMTPLIKRNTTIPTKKSETFSTFSDN QPGVLIQVFEGERARTKDNNLLGKFELTGIPRARGVPQIEVTF DVDANGIMNVSALEKGTTRKTNKIVITNDKGRLSKEEIERMLA EAEKYKAEDEAEASRIRPKNGLESYAYSRLNSLRHSHKVDEKL EAGDKEKLKSEIDKTVQWLDENQTATKEEYESQKKELEAVA NPIMMKFYAGGEGAPGGFPGAGGPGGPGGPGAGHASGGGD DGPTVEEVDLKFPMPLPWQLSVRKMRPFLLFLFLIFLIFLIL FLFYFFLPVRFNESCFS
775	GroEL	MAAKDVKFGNDARVKMLRGVNVLADAVKVTGLGPKGRNVV LDKSFGAPTTITKDGVSVAEREIELEDKFENMGAQMVKEVASKA NDAAGDGTITATVLAQAITEGLKAVAAGMNPMDLKRIGDK AVTAAVEELKALSVPSCSDSKAIAQVGTISANSDETGVKLIAEA MDKVGKEGVITVEDGTGLQDELDDVEGMQFDRGYLSPYFIN KPETGAVELESPFILLADKKISNIREMLPVLEAVAKAGKPLLIIA EDVEGEALATLVVNTMRGIVKVAAVKAPGFGDRRKAMLQDI ATLTGGTVISEEIGMELEKATLEDLGQAKRVVINKDTTTIIDGV GEEAAIQGRVAQIRQQIEEATSDYDREKLQERVAKLAGGVAV IKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVA LIRVASKLADLRGQNEQNVGIKVALRAMEAPLRQIVLNCGE EPSVVANTVKGGDGNYGYNAAATEEYGNMIDMGILDPTKVTR SALQYAASVAGLMITTECMVTDLPKNDAAADLGAAGGMGGM GGMGGMM
776	GroES	MNIRPLHDRVIVKRKEVETKSAGGIVLTGSAAAKSTRGEVLA VGNGRILENGEVKPLDVKVGDIVIFNDGYGVKSEKIDNEEVLI MSSEDILAIVEA
777	Hsp90-alpha	MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIF LRELISNSSDALDKIRYESLTDPSKLDSGKELHINLIPNKQDRTL TIVDTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIG QFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRT DTGPEPMGRGTKVILHLKEDQTEYLEE RRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEKEDKEEE KEKEEKESEDKPEIEDVGSDEEEEEKDGDKKKKKKIKKEYIDQ EELNKTPIWTRNPDDITNEEYGEFYKSLTNDWEDHLAVKHF SVEGQLEFRALLFVPRRAPFDLFENRKKKNNIKLYVRRVFIMD NCEELIPEYLNfirGVVDSEDLPLNISR

		<p>EMLQQSKILKVIRKNLVKKCLELFTELAEDKENYKFFYEQFSK NIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSCLKDYCTRMK ENQKHIIYYITGETKDQVANSASFVERLRKHGLEVIYMIPIDEY CVQQLKEFEGKTLVSVTKEGLELPEDEEEKKKQEEKKTKFEN LCKIMKDILEKKVEKVVSNNRLVTSPCCIV TSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHS IETLRQKAEADKNDKSVKDLVILLYETALLSSGFSLEDPPQTHA NRIYRMKLGLEDDEDDPTADDTSAAVTEEMPPLEGDDDDTSR MEEVD</p>
778	HtpG	<p>MKGQETRFGFQSEVKQLLHLMHSLYSNKEIFLRELISNASDAA DKLRFRLSNPDLYEGDGELRVVVSFDKDKRTLITSDNGVGM TRDEVIDHLGTIAKSGTKSFLESLGSDQAKDSQLIGQFGVGFYS AFIVADKVTVRTRAAGEKPEGVFWESAGEGEYTVADITKED RGTEITLHLREGEDEFLLDDWRVRSIISKY SDHIALPVEIEKREEKDGETVISWEKINKAQALWTRNKSEITDE EYKEFYKHIAHDFNDPLTWSHNRVEGKQEYTSLLYIPSQAPW DMWNRDHHKGLKLYVQRFVIMDDAEQFMPNYLRFVRGLIDS SDLPLNVSREILODSTVTRNLRN ALTKRVLQMLEKLAKDDAEKYQTFWQQFGLVLKEGPAEDFA NQEAIAKLLRFASHTDSSAQTVSLEDYVSRMKEGQEKIYYIT ADSYAAAKSSPHLELLRKKGIEVLLLSDRIDEWMMNYLTFED GKPFQSVSKVDESLEKLADEVDESACEAEKALTPFIDRVKALL GERVKDVRRLTHRLTDTPAIVSTDADEMSTQM AKLFAAAGQKVPEVKYIFELNPDHVLVKRAADTEDEAKFSE WVELLLDQALLAERGTLDEPNLFIRRMNQLLVS</p>
779	Hsp40	<p>MGKDYQTLGLARGASDEEIKRAYRRQALRYHPDKNKEPGA EEKFKEIAEAYDVLSDPRKREIFDRYGEEGLKSGSPSGSGGG ANGTSFSYTFHGDPHAMFAEFFGGRNPFDTFFGQRNGEEMD IDDPFSGFPMGMGGFTNVNFGRSRQAEPARKKQDPPVTHDL RVSLEEIYSGCTKKMKISHKRLNPDGKSIRNEDKILIEVKKG WKEGTKITFPKEGDQTSNNIPADIVFVLKDKPHNIFKRDGSDVI YPARISLREALCGCTVNVPTLDGRTIPVVFVDVIRPGMRKVP GEGPLPKTPEKRGDLIEFEVIFPERIPQTSRTVLEQVLP</p>
780	ClpP	<p>MWPGILVGGARVASCYPALGPRLAAHFPAQRPPQRTLQNGL ALQRCLHATATRALPLIPVVEQTGRGERAYDIYSRLLRERIVC VMGPIDDSVASLVIAQLLFLQSESNNKPIHMYINSPGGVVTAG LAIYDTMQYILNPICTWCVGQAASMGSLLLAAGTPGMRHSLP NSRIMIHQPSGGARGQATDIAIQAEIIMKLKKQLYNIYAKHTK QSLQVIESAMERDRYMSPMEAQEFGLDKVLVHPPQDGEDEP TLVQKEPVEAAPAAEPVPAST</p>
781	ClpX	<p>MPSCGACTCGAAVRLITSSLASAQRGISGGRIHMSVLGRLGT FETQILQRAPLRSFTETPAYFASKDGIKDGSGDGNKKSASEGS SKKSGSGNSGKGG NQLRCPKCGDLCTHVETFVSSTRFVKCEKCHHFFVVLSEADS KKSIIKEPESAAEAVKLAFQKPPPPPKKIYNYLDKYVVGQSF AKKVLSVAVYNHYKRIYNNIPANLRQQAEEVEKQTSITPRELEI RRREDEYRFTKLLQIAGISPHG NALGASMQQQVNQQIPQEKRGGEVLDSSHDDIKLEKSNILL GPTGSGKTLLAQTLAKCLDVPFAICDCTTLTQAGYVGEDIESV IAKLLQDANYNVEKAQQGIVFLDEVVKIGSVPGIHQLRDVGG EGVQQGLLKLLEGITVNVPEKNSRKLKRGTVQVDTTNILFVAS GAFNGLDRIISRRKNEKYLGFGTSPNLGKG RRAAAAADLANRSGESNTHQDIEEKDRLLRHVEARDLIEFGM IPEFVGRLPVVPLHSLDEKTLVQILTEPRNAVIPQYQALFSMD</p>

		KCELVNTEADAKAIARLALERKTGARGLRISIMEKLLLEPMFEV PNSDIVCVEVDKEVVEGKKEPGYIRAPTKESSEEEYDSGVEEE GWPRQADAANS
782	CDKN1C	MSDASLRSTSTMERLVARGTFPVLVRTSACRSLFGPVDHEELS RELQARLAELNAEDQNRWDYDFQQDMPLRGPGRLQWTEVD SDSVPAFYRETVQVGRCLLLAPRPVAVAVAVSPPLEPAAESL DGLLEAPEQLPSVPVPAPASTPPPVPVLAPAPAPAPAPVAA AAPVAVAVLAPAPAPAPAPAPAPVAAAPAPAPAPAPAPAP PAPAPDAAPQESAEQGANQGGQRGQEPLADQLHSGISGRPAAG TAAASANGAAIKKLSGPLISDFFAKRKRSAPKSSGDVPAPCPS PSAAPGVGSVEQTTPRKRLR
783	BAX inhibitor	MNIFDRKINFDAALLKFSHTPSTQQHLKKVYASFALCMFVAAA GAYVHMTVTHFIQAGLLSALGSLILMIWLMATPHSHETEQRKL GLLAGFAFLTGVGLGPALEFCIAVNPSILPTAFMGTA MIFTCFT LSALYARRRSYLFLLGGILMSALSLLLLSSLGNVFFGSIWLFQAN LYVGLVVMCGFVLFDLQLIIEKAEHGD
784	codA	MSNNALQTIINARLPGEGLWQIHLQDGKISAIDAQSGVMPIT ENSLDAEQGLVPPFVEPHIHLDTTQTAGQPNWNQSGTLFEGIE RWAERKALLTHDDVKQRAWQTLKWQIANGIQHVTRTHVDVS DATLTALKAMLEVKEVAPWIDLQIVAFPQEGIQDYIWHCIDL FLDFITVFRKLMMILAMNEKDKKKEKKLSYPNGEALLEEALR LGADVVGAIHPHEFTREYGVESLHKTFALAQKYDRLIDVHCD EIDDEQSRFVETVAALAHHEGMGARVTASHTTAMHSYNGAY TSRLFRLLKMSGINFVANPLVNIHLQGRFDTPKRRGITRVKE MLESINVCFGHDDVDFDPWYPLGTANMLQVLHMLHVCQL MGYGGQINDGLNLITHHSARTLNLQDYGIAAGNSANLILPAEN GFDALRRQVPVRYSVRGKVIASQTQPAQTTVYLEQPEAIDYK R

[0692] In some embodiments, a circular RNA or a precursor RNA (*e.g.*, linear precursor RNA) disclosed herein comprises a sequence as listed in Table 29.

[0693] In some embodiments, a polynucleotide or a protein encoded by a polynucleotide contains a sequence with at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% similarity to one or more sequences disclosed herein. In some embodiments, a polynucleotide or a protein encoded by a polynucleotide contains a sequence that is identical to one or more sequences disclosed herein.

[0694] Preferred embodiments are described herein. Variations of those preferred embodiments may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible

variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

EXAMPLES

[0695] Wesselhoeft *et al.*, (2019) RNA Circularization Diminishes Immunogenicity and Can Extend Translation Duration *In vivo*. Molecular Cell. 74(3), 508-520 and Wesselhoeft *et al.*, (2018) Engineering circular RNA for Potent and Stable Translation in Eukaryotic Cells. Nature Communications. 9, 2629 are incorporated by reference in their entirety.

[0696] The invention is further described in detail by reference to the following examples but are not intended to be limited to the following examples. These examples encompass any and all variations of the illustrations with the intention of providing those of ordinary skill in the art with complete disclosure and description of how to make and use the subject invention and are not intended to limit the scope of what is regarded as the invention.

EXAMPLE 1

Example 1A: External homology regions allow for circularization of long precursor RNA using the permuted intron exon (PIE) circularization strategy.

[0697] A 1,100nt sequence containing a full-length encephalomyocarditis virus (EMCV) IRES, a *Gaussia* luciferase (GLuc) expression sequence, and two short exon fragments of the permuted intron-exon (PIE) construct were inserted between the 3' and 5' introns of the permuted group I catalytic intron in the thymidylate synthase (Td) gene of the T4 phage. Precursor RNA was synthesized by run-off transcription. Circularization was attempted by heating the precursor RNA in the presence of magnesium ions and GTP, but splicing products were not obtained.

[0698] Perfectly complementary 9 nucleotide and 19 nucleotide long homology regions were designed and added at the 5' and 3' ends of the precursor RNA. Addition of these homology arms increased splicing efficiency from 0 to 16% for 9 nucleotide homology regions and to 48% for 19 nucleotide homology regions as assessed by disappearance of the precursor RNA band.

[0699] The splicing product was treated with RNase R. Sequencing across the putative splice junction of RNase R-treated splicing reactions revealed ligated exons, and digestion of the RNase R-treated splicing reaction with oligonucleotide-targeted RNase H produced a single band in contrast to two bands yielded by RNase H-digested linear precursor. This

shows that circular RNA is a major product of the splicing reactions of precursor RNA containing the 9 or 19 nucleotide long external homology regions.

Example 1B: Spacers that conserve secondary structures of IRES and PIE splice sites increase circularization efficiency.

[0700] A series of spacers was designed and inserted between the 3' PIE splice site and the IRES. These spacers were designed to either conserve or disrupt secondary structures within intron sequences in the IRES, 3' PIE splice site, and/or 5' splice site. The addition of spacer sequences designed to conserve secondary structures resulted in 87% splicing efficiency, while the addition of a disruptive spacer sequences resulted in no detectable splicing.

EXAMPLE 2

Example 2A: Internal homology regions in addition to external homology regions creates a splicing bubble and allows for translation of several expression sequences.

[0701] Spacers were designed to be unstructured, non-homologous to the intron and IRES sequences, and to contain spacer-spacer homology regions. These were inserted between the 5' exon and IRES and between the 3' exon and expression sequence in constructs containing external homology regions, EMCV IRES, and expression sequences for *Gaussia* luciferase (total length: 1289nt), Firefly luciferase (2384nt), eGFP (1451nt), human erythropoietin (1313nt), and Cas9 endonuclease (4934nt). Circularization of all 5 constructs was achieved. Circularization of constructs utilizing T4 phage and Anabaena introns were roughly equal. Circularization efficiency was higher for shorter sequences. To measure translation, each construct was transfected into HEK293 cells. *Gaussia* and Firefly luciferase transfected cells produced a robust response as measured by luminescence, human erythropoietin was detectable in the media of cells transfected with erythropoietin circRNA, and EGFP fluorescence was observed from cells transfected with EGFP circRNA. Co-transfection of Cas9 circRNA with sgRNA directed against GFP into cells constitutively expressing GFP resulted in ablated fluorescence in up to 97% of cells in comparison to an sgRNA-only control.

Example 2B: Use of CVB3 IRES increases protein production.

[0702] Constructs with internal and external homology regions and differing IRES

containing either *Gaussia* luciferase or Firefly luciferase expression sequences were made. Protein production was measured by luminescence in the supernatant of HEK293 cells 24 hours after transfection. The Coxsackievirus B3 (CVB3) IRES construct produced the most protein in both cases.

Example 2C: Use of polyA or polyAC spacers increases protein production.

[0703] Thirty nucleotide long polyA or polyAC spacers were added between the IRES and splice junction in a construct with each IRES that produced protein in example 2B. *Gaussia* luciferase activity was measured by luminescence in the supernatant of HEK293 cells 24 hours after transfection. Both spacers improved expression in every construct over control constructs without spacers.

EXAMPLE 3

HEK293 or HeLa cells transfected with circular RNA produce more protein than those transfected with comparable unmodified or modified linear RNA.

[0704] HPLC-purified *Gaussia* luciferase-coding circRNA (CVB3-GLuc-pAC) was compared with a canonical unmodified 5' methylguanosine-capped and 3' polyA-tailed linear GLuc mRNA, and a commercially available nucleoside-modified (pseudouridine, 5-methylcytosine) linear GLuc mRNA (from Trilink). Luminescence was measured 24 h post-transfection, revealing that circRNA produced 811.2% more protein than the unmodified linear mRNA in HEK293 cells and 54.5% more protein than the modified mRNA. Similar results were obtained in HeLa cells and a comparison of optimized circRNA coding for human erythropoietin with linear mRNA modified with 5-methoxyuridine.

[0705] Luminescence data was collected over 6 days. In HEK293 cells, circRNA transfection resulted in a protein production half-life of 80 hours, in comparison with the 43 hours of unmodified linear mRNA and 45 hours of modified linear mRNA. In HeLa cells, circRNA transfection resulted in a protein production half-life of 116 hours, in comparison with the 44 hours of unmodified linear mRNA and 49 hours of modified linear mRNA. CircRNA produced substantially more protein than both the unmodified and modified linear mRNAs over its lifetime in both cell types.

EXAMPLE 4

Example 4A: Purification of circRNA by RNase digestion, HPLC purification, and phosphatase treatment decreases immunogenicity. Completely purified circular RNA is

significantly less immunogenic than unpurified or partially purified circular RNA. Protein expression stability and cell viability are dependent on cell type and circular RNA purity.

[0706] Human embryonic kidney 293 (HEK293) and human lung carcinoma A549 cells were transfected with:

- products of an unpurified GLuc circular RNA splicing reaction,
- products of RNase R digestion of the splicing reaction,
- products of RNase R digestion and HPLC purification of the splicing reaction, or
- products of RNase digestion, HPLC purification, and phosphatase treatment of the splicing reaction.

[0707] RNase R digestion of splicing reactions was insufficient to prevent cytokine release in A549 cells in comparison to untransfected controls.

[0708] The addition of HPLC purification was also insufficient to prevent cytokine release, although there was a significant reduction in interleukin-6 (IL-6) and a significant increase in interferon- α 1 (IFN- α 1) compared to the unpurified splicing reaction.

[0709] The addition of a phosphatase treatment after HPLC purification and before RNase R digestion dramatically reduced the expression of all upregulated cytokines assessed in A549 cells. Secreted monocyte chemoattractant protein 1 (MCP1), IL-6, IFN- α 1, tumor necrosis factor α (TNF α), and IFN γ inducible protein-10 (IP-10) fell to undetectable or untransfected baseline levels.

[0710] There was no substantial cytokine release in HEK293 cells. A549 cells had increased GLuc expression stability and cell viability when transfected with higher purity circular RNA. Completely purified circular RNA had a stability phenotype similar to that of transfected 293 cells.

Example 4B: Circular RNA does not cause significant immunogenicity and is not a RIG-I ligand.

[0711] A549 cells were transfected with the products of a splicing reaction:

[0712] A549 cells were transfected with:

- unpurified circular RNA,
- high molecular weight (linear and circular concatenations) RNA,
- circular (nicked) RNA,
- an early fraction of purified circular RNA (more overlap with nicked RNA peak),
- a late fraction of purified circular RNA (less overlap with nicked RNA peak),

- introns excised during circularization, or
- vehicle (*i.e.* untransfected control).

[0713] Precursor RNA was separately synthesized and purified in the form of the splice site deletion mutant (DS) due to difficulties in obtaining suitably pure linear precursor RNA from the splicing reaction. Cytokine release and cell viability was measured in each case.

[0714] Robust IL-6, RANTES, and IP-10 release was observed in response to most of the species present within the splicing reaction, as well as precursor RNA. Early circRNA fractions elicited cytokine responses comparable to other non-circRNA fractions, indicating that even relatively small quantities of linear RNA contaminants are able to induce a substantial cellular immune response in A549 cells. Late circRNA fractions elicited no cytokine response in excess of that from untransfected controls. A549 cell viability 36 hours post-transfection was significantly greater for late circRNA fractions compared with all of the other fractions.

[0715] RIG-I and IFN- β 1 transcript induction upon transfection of A549 cells with late circRNA HPLC fractions, precursor RNA or unpurified splicing reactions were analyzed. Induction of both RIG-I and IFN- β 1 transcripts were weaker for late circRNA fractions than precursor RNA and unpurified splicing reactions. RNase R treatment of splicing reactions alone was not sufficient to ablate this effect. Addition of very small quantities of the RIG-I ligand 3p-hpRNA to circular RNA induced substantial RIG-I transcription. In HeLa cells, transfection of RNase R-digested splicing reactions induced RIG-I and IFN- β 1, but purified circRNA did not. Overall, HeLa cells were less sensitive to contaminating RNA species than A549 cells.

[0716] A time course experiment monitoring RIG-I, IFN- β 1, IL-6, and RANTES transcript induction within the first 8 hours after transfection of A549 cells with splicing reactions or fully purified circRNA did not reveal a transient response to circRNA. Purified circRNA similarly failed to induce pro-inflammatory transcripts in RAW264.7 murine macrophages.

[0717] A549 cells were transfected with purified circRNA containing an EMCV IRES and EGFP expression sequence. This failed to produce substantial induction of pro-inflammatory transcripts. These data demonstrate that non-circular components of the splicing reaction are responsible for the immunogenicity observed in previous studies and that circRNA is not a natural ligand for RIG-I.

EXAMPLE 5

Circular RNA avoids detection by TLRs.

[0718] TLR 3, 7, and 8 reporter cell lines were transfected with multiple linear or circular RNA constructs and secreted embryonic alkaline phosphatase (SEAP) was measured.

[0719] Linearized RNA was constructed by deleting the intron and homology arm sequences. The linear RNA constructs were then treated with phosphatase (in the case of capped RNAs, after capping) and purified by HPLC.

[0720] None of the attempted transfections produced a response in TLR7 reporter cells. TLR3 and TLR8 reporter cells were activated by capped linearized RNA, polyadenylated linearized RNA, the nicked circRNA HPLC fraction, and the early circRNA fraction. The late circRNA fraction and m1ψ-mRNA did not provoke TLR-mediated response in any cell line.

[0721] In a second experiment, circRNA was linearized using two methods: treatment of circRNA with heat in the presence of magnesium ions and DNA oligonucleotide-guided RNase H digestion. Both methods yielded a majority of full-length linear RNA with small amounts of intact circRNA. TLR3, 7, and 8 reporter cells were transfected with circular RNA, circular RNA degraded by heat, or circular RNA degraded by RNase H, and SEAP secretion was measured 36 hours after transfection. TLR8 reporter cells secreted SEAP in response to both forms of degraded circular RNA, but did not produce a greater response to circular RNA transfection than mock transfection. No activation was observed in TLR3 and TLR7 reporter cells for degraded or intact conditions, despite the activation of TLR3 by *in vitro* transcribed linearized RNA.

EXAMPLE 6

Unmodified circular RNA produces increased sustained in vivo protein expression than linear RNA.

[0722] Mice were injected and HEK293 cells were transfected with unmodified and m1ψ-modified human erythropoietin (hEpo) linear mRNAs and circRNAs. Equimolar transfection of m1ψ-mRNA and unmodified circRNA resulted in robust protein expression in HEK293 cells. hEpo linear mRNA and circRNA displayed similar relative protein expression patterns and cell viabilities in comparison to GLuc linear mRNA and circRNA upon equal weight transfection of HEK293 and A549 cells.

[0723] In mice, hEpo was detected in serum after the injection of hEpo circRNA or linear mRNA into visceral adipose. hEpo detected after the injection of unmodified circRNA

decayed more slowly than that from unmodified or m1ψ-mRNA and was still present 42 hours post-injection. Serum hEpo rapidly declined upon the injection of unpurified circRNA splicing reactions or unmodified linear mRNA. Injection of unpurified splicing reactions produced a cytokine response detectable in serum that was not observed for the other RNAs, including purified circRNA.

EXAMPLE 7

Circular RNA can be effectively delivered in vivo or in vitro via lipid nanoparticles.

[0724] Purified circular RNA was formulated into lipid nanoparticles (LNPs) with the ionizable lipidoid cKK-E12 (Dong *et al.*, 2014; Kauffman *et al.*, 2015). The particles formed uniform multilamellar structures with an average size, polydispersity index, and encapsulation efficiency similar to that of particles containing commercially available control linear mRNA modified with 5moU.

[0725] Purified hEpo circRNA displayed greater expression than 5moU-mRNA when encapsulated in LNPs and added to HEK293 cells. Expression stability from LNP-RNA in HEK293 cells was similar to that of RNA delivered by transfection reagent, with the exception of a slight delay in decay for both 5moU-mRNA and circRNA. Both unmodified circRNA and 5moU-mRNA failed to activate RIG-I/IFN-β1 *in vitro*.

[0726] In mice, LNP-RNA was delivered by local injection into visceral adipose tissue or intravenous delivery to the liver. Serum hEpo expression from circRNA was lower but comparable with that from 5moU-mRNA 6 hours after delivery in both cases. Serum hEpo detected after adipose injection of unmodified LNP-circRNA decayed more slowly than that from LNP-5moU-mRNA, with a delay in expression decay present in serum that was similar to that noted *in vitro*, but serum hEpo after intravenous injection of LNP-circRNA or LNP-5moU-mRNA decayed at approximately the same rate. There was no increase in serum cytokines or local RIG-I, TNFα, or IL-6 transcript induction in any of these cases.

EXAMPLE 8

Example 8A: Expression and functional stability by IRES in HEK293, HepG2, and 1C1C7 cells.

[0727] Constructs including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and varying IRES were circularized. 100 ng of each circularization reaction was separately transfected into 20,000 HEK293 cells, HepG2 cells, and 1C1C7 cells using Lipofectamine MessengerMax. Luminescence in each supernatant was assessed after

24 hours as a measure of protein expression. In HEK293 cells, constructs including Crohivirus B, Salivirus FHB, Aichi Virus, Salivirus HG-J1, and Enterovirus J IRES produced the most luminescence at 24 hours (**FIG. 1A**). In HepG2 cells, constructs including Aichi Virus, Salivirus FHB, EMCV-Cf, and CVA3 IRES produced high luminescence at 24 hours (**FIG. 1B**). In 1C1C7 cells, constructs including Salivirus FHB, Aichi Virus, Salivirus NG-J1, and Salivirus A SZ-1 IRES produced high luminescence at 24 hours (**FIG. 1C**).

[0728] A trend of larger IRES producing greater luminescence at 24 hours was observed. Shorter total sequence length tends to increase circularization efficiency, so selecting a high expression and relatively short IRES may result in an improved construct. In HEK293 cells, a construct using the Crohivirus B IRES produced the highest luminescence, especially in comparison to other IRES of similar length (**FIG. 2A**). Expression from IRES constructs in HepG2 and 1C1C7 cells plotted against IRES size are in **FIGs. 2B and 2C**.

[0729] Functional stability of select IRES constructs in HepG2 and 1C1C7 cells were measured over 3 days. Luminescence from secreted *Gaussia* luciferase in supernatant was measured every 24 hours after transfection of 20,000 cells with 100 ng of each circularization reaction, followed by complete media replacement. Salivirus A GUT and Salivirus FHB exhibited the highest functional stability in HepG2 cells, and Salivirus N-J1 and Salivirus FHB produced the most stable expression in 1C1C7 cells (**FIGs. 3A and 3B**).

Example 8B: Screening of additional IRES

[0730] Functional stability of additional IRES constructs in HEK293 cells were measured. Briefly, 5' untranslated regions (UTRs) of interest were identified from GenBank. Selected UTRs were truncated to 675nt from the 5' end and inserted into a circular RNA backbone construct encoding *Gaussia* Luciferase (Gluc) and in front of the Gluc coding region. The circular RNAs were transfected into HEK293 cells. After 24 hours, the supernatants were collected and the luminescence from secreted Gluc protein was measured using commercially available reagents. The results are depicted in **FIGs. 1D and 1E** and Table 30, suggesting that many natural IRES sequences enhance the protein expression in a circular RNA context.

Table 30

SEQ ID NO	IRES	Expression	SEQ ID NO	IRES	Expression
413	RhPV	1.10E+05	553	Pronghorn antelope pestivirus	1.35E+06
414	Halastavi arva (1x mut)	9.46E+04	554	Porcine pestivirus isolate Bungowannah	1.10E+07

SEQ ID NO	IRES	Expression	SEQ ID NO	IRES	Expression
415	Oscivirus	4.55E+07	555	Porcine pestivirus 1	9.46E+04
416	Cadicivirus B	2.10E+05	556	Pestivirus giraffe-1	4.72E+05
417	PSIV (2x mut for XbaI)	9.70E+04	557	Classical swine fever virus	3.16E+05
418	PSIV IGR	1.01E+05	558	Human pegivirus isolate JD2B11	6.85E+05
419	PV Mahoney	1.09E+05	559	Human pegivirus isolate GBV-C-ZJ	N/A
420	REV A	9.44E+04	560	Human pegivirus isolate JD2B8C	5.36E+05
421	Tropivirus A	9.52E+04	561	Hepatitis GB virus A	N/A
422	Symapivirus A	1.27E+05	562	Simian pegivirus	8.56E+04
423	Sakobuvirus A FFUP1 (1x mut)	8.82E+06	563	Pegivirus I	8.02E+04
424	Rosavirus C NFSM6F	6.84E+05	564	Pegivirus K	8.07E+04
425	Rosavirus 2 GA7403	5.05E+06	565	Theiler's disease-associated virus	7.84E+04
426	Rhinovirus A	8.42E+05	566	Rodent pegivirus	1.79E+05
427	Rafivirus LPXYC222841	2.22E+05	567	Human pegivirus 2	3.14E+05
428	Rafivirus WHWGGF74766	4.53E+06	568	GB virus C/Hepatitis G virus	1.36E+05
429	Poecivirus BCCH-449	3.43E+05	569	Equine Pegivirus 1	8.80E+04
430	Megirivirus A LY	1.80E+06	570	Culex theileri flavivirus	8.52E+04
431	Megirivirus E	1.10E+07	571	Bussuquara virus	8.20E+04
432	Megirivirus C	1.24E+05	572	Zika Virus	8.61E+04
433	Ludopivirus	1.05E+05	573	Yokose virus	8.55E+04
434	Livupivirus	2.10E+05	574	Wesselsbron virus	N/A
435	Aichivirus A FSS693	6.25E+07	575	Equine hepacivirus	8.40E+04
436	Aichivirus KVGH	1.72E+07	576	Hepacivirus B	8.84E+04
437	Aichivirus DV	7.79E+07	577	Hepacivirus I	7.50E+04
438	Murine Kobuvirus 1	1.60E+07	578	Hepacivirus J	7.65E+04
439	Porcine Kobuvirus K-30	N/A	579	Hepacivirus K	8.91E+04
440	Porcine Kobuvirus XX	1.32E+07	580	Icavirus	4.41E+06
441	Caprine Kobuvirus 12Q108	2.87E+08	581	Antarctic penguin virus A	8.42E+04
442	Rabbit Kobuvirus	3.73E+07	582	Forest pouched giant rat arterivirus	N/A
443	Aalivirus	2.65E+05	583	Avisivirus Pf-CHK1	1.19E+05
444	Grusopivirus A	1.09E+05	584	Avian paramyxovirus penguin	9.91E+04
445	Grusopivirus B	2.12E+05	585	Newcastle disease virus	8.86E+04
446	Yancheng osbecks grenadier anchovy picornavirus	1.57E+06	586	Bat Hp-betacoronavirus	8.47E+04

SEQ ID NO	IRES	Expression	SEQ ID NO	IRES	Expression
447	Turkey Gallivirus M176	4.37E+05	587	Basella alba endornavirus	7.65E+04
448	Falcovirus A1	1.48E+05	588	Ball python nidovirus	8.25E+04
449	Tremovirus B	1.31E+05	589	Bat sapelovirus	8.05E+04
450	Didelphis aurita HAV	1.38E+05	590	Bat Picornavirus 3	N/A
451	Hepatovirus G1	1.41E+05	591	Bat Picornavirus 2	7.99E+07
452	Hepatovirus D	1.47E+06	592	Bat Picornavirus 1	1.85E+07
453	Hepatovirus H2	1.08E+05	593	Bat Iflavivirus	9.76E+04
454	Hepatovirus I	8.79E+05	594	Bat dicibavirus	7.43E+04
455	Hepatovirus C	5.08E+05	595	Betacoronavirus HKU24	8.96E+04
456	Fipivirus A	2.69E+05	596	Betacoronavirus England 1	8.74E+04
457	Fipivirus C	1.09E+05	597	Boone cardiovirus 1	2.62E+06
458	Fipivirus E	1.10E+05	598	Breda virus	1.16E+05
459	Aquamavirus	4.51E+06	599	Bovine viral diarrhea virus 3	2.70E+06
460	Avisivirus A	1.91E+05	600	Bovine rhinitis A virus	3.62E+06
461	Avisivirus B	8.68E+04	601	Bovine picornavirus isolate TCH6	1.21E+05
462	Crohivirus A	9.96E+04	602	Bovine nidovirus TCH5	1.17E+05
463	Kunsagivirus B	8.01E+04	603	Bovine hepacivirus	1.89E+05
464	Limnipivirus A	8.30E+04	604	Botrytis cinerea mitovirus 4 RdRp	9.68E+04
465	Limnipivirus C	1.35E+05	605	Botrytis cinerea mitovirus 2 RdRp	8.73E+04
466	Orivirus	6.09E+05	606	Canine picodistovirus strain 209	2.79E+06
467	HAV FHI	1.24E+05	607	Canine distemper virus	3.02E+05
468	HAV HMI75	4.96E+05	608	Canine kobuvirus	1.48E+08
469	Parechovirus F	6.56E+05	609	Camel alphacoronavirus	2.48E+05
470	Parechovirus D	3.10E+05	610	Cripavirus	1.95E+05
471	Parechovirus C	1.24E+06	611	Human coxsackievirus A2	7.75E+07
472	Ljungan Virus 87-012	2.00E+06	612	Coronavirus AcCoV-JC34	1.82E+05
473	Parechovirus A2	1.80E+07	613	Chicken picornavirus 3	9.13E+04
474	Parechovirus A3	3.58E+06	614	Chicken picornavirus 1	1.21E+05
475	Parechovirus A8	1.61E+07	615	Chicken orivirus 1	3.16E+05
476	Parechovirus A17	1.20E+06	616	Chicken gallivirus 1	1.51E+07
477	Potamipivirus A	8.43E+05	617	Chicken calicivirus	1.28E+05
478	Potamipivirus B	7.20E+05	618	Carp picornavirus 1	1.13E+05
479	Beihai Conger Picornavirus	1.15E+06	619	Falcon picornavirus	3.08E+06

SEQ ID NO	IRES	Expression	SEQ ID NO	IRES	Expression
480	Porcine Sapelovirus JD2011	N/A	620	Equine rhinitis B virus 1	1.01E+05
481	Porcine Sapelovirus A2	4.34E+06	621	Equine rhinitis A virus	3.73E+05
482	Simian Sapelovirus 1	6.55E+07	622	Equine arteritis virus	1.89E+05
483	Simian Sapelovirus 2	4.24E+07	623	Enterovirus sp. isolate CPML	6.83E+07
484	Rabovirus C	2.49E+06	624	Enterovirus AN12	3.87E+06
485	Rabovirus A NYC-B10	1.24E+06	625	Dolphin morbillivirus	1.22E+05
486	Parabovirus C	1.83E+07	626	Dianke virus	1.35E+05
487	Parabovirus B	7.85E+06	627	Guereza hepacivirus	1.38E+05
488	Parabovirus A3	2.44E+08	628	Grapevine associated narnavirus-1	1.30E+05
489	Felipivirus 127F	8.92E+06	629	Goat torovirus	1.19E+05
490	Boosepivirus A	7.07E+07	630	Foot-and-mouth disease virus O isolate	1.12E+05
491	Boosepivirus B	1.17E+08	631	Feline infectious peritonitis virus	1.35E+05
492	Phacovirus PF-CHK1	5.87E+06	632	Farmington virus	1.22E+05
493	HRVC3 QPM	1.64E+07	633	Avian infectious bronchitis virus	2.84E+05
494	HRVB27	2.04E+08	634	Human rhinovirus 1	7.40E+07
495	HRVA73	1.08E+08	635	EV22	1.95E+07
496	EV L	6.49E+07	636	Human TMEV-like cardiovirus	4.48E+07
497	EV K	7.52E+07	637	Human coronavirus 229E	N/A
498	EV J 1631	9.88E+07	638	Hubei zhaovirus-like virus 1	1.03E+05
499	EV J N125	2.90E+07	639	Hubei tombus-like virus 9	9.28E+04
500	EV I	1.31E+08	640	Hubei tombus-like virus 32	9.23E+04
501	EV F1 BEV 261	1.12E+07	641	Hubei sobemo-like virus 3	1.17E+05
502	EV D94	9.25E+07	642	Hubei picorna-like virus 2	1.95E+05
503	PV3	1.25E+08	643	Hepacivirus P	6.04E+05
504	EV C102	8.85E+07	644	Harrier picornavirus 1	1.47E+05
505	EV 30	5.48E+06	645	Kunsagivirus 1	4.15E+05
506	SA5	1.61E+08	646	Kagoshima-2-24-KoV	9.30E+07
507	EV A114	1.50E+08	647	Kashmir bee virus	1.65E+05
508	Mobovirus A	3.44E+06	648	Jingmen picorna-like virus	9.32E+04
509	Burpengary Virus	1.09E+07	649	Mumps virus	1.47E+05
510	Hunnivirus A1	1.61E+06	650	Mouse Mosavirus	9.00E+04

SEQ ID NO	IRES	Expression	SEQ ID NO	IRES	Expression
511	Hunnivirus A2	6.38E+06	651	Miniopterus schreibersii picornavirus 1	6.05E+06
512	Ia Io	1.35E+06	652	Linda virus	7.37E+05
513	Taura Syndrome Virus	8.30E+05	653	Lesavirus 2	3.67E+07
514	ABPV	6.48E+05	654	Lesavirus 1	6.37E+06
515	BRAV-2	3.98E+06	655	Phopivirus strain NewEngland	1.06E+05
516	BRBV-1	3.34E+06	656	Pestivirus strain Aydin	3.11E+06
517	ERAV-1 U188	N/A	657	Quail picornavirus QPV1	6.55E+07
518	GFTV	1.23E+06	658	Porcine sapelovirus 1	N/A
519	SAFV V13C	9.32E+07	659	Porcine reproductive and respiratory syndrome virus 2	1.29E+05
520	SAV P-113	4.37E+07	660	Porcine enterovirus 9	3.20E+07
521	VHEV	1.74E+08	661	Pigeon picornavirus B	1.24E+05
522	TRV NGS910	3.84E+07	662	Picornavirus HK21	4.09E+05
523	EMCV2 RD1338	1.97E+06	663	Picornavirales Tottori-HG1	9.54E+04
524	EMCV1 JZ1203	N/A	664	Rodent hepatovirus	1.39E+05
525	EMCV1 AnrB-3741	2.55E+06	665	Rinderpest virus	4.26E+05
526	Cosavirus D1	2.11E+06	666	Rabovirus A	2.88E+06
527	Cosavirus B1	1.91E+06	667	Shingleback nidovirus 1	2.62E+05
528	Cosavirus A SH1	2.16E+06	668	Seneca valley virus	1.46E+07
529	Malagasivirus B	5.05E+06	669	Sclerotinia sclerotiorum dsRNA mycovirus-L	1.69E+05
530	Mosavirus A2 SZAL6	8.27E+06	670	Yak enterovirus	6.19E+06
531	SVV	1.06E+06	671	Wobbly possum disease virus	2.60E+05
532	PTV A	7.29E+05	672	Avian orthoreovirus segment S1	4.37E+05
533	PTV B	6.02E+06	673	Caprine Kobuvirus d10	2.20E+08
534	Tottorivirus	2.76E+07	674	Caprine Kobuvirus d20	2.00E+08
535	Posavirus 1	1.55E+06	675	Caprine Kobuvirus d30	1.87E+08
536	A105-675	2.18E+07	676	Caprine Kobuvirus d40	2.15E+08
537	A110-675	1.24E+08	677	Caprine Kobuvirus d50	9.65E+07
538	18-675	6.04E+07	678	Picornavirales sp. isolate RtMruf-PicoV	2.26E+08
539	A115-675	5.93E+07	679	Apodemus agrarius picornavirus strain Longquan-Aa118	1.90E+08
540	A73-675	1.30E+08	680	Niviventer confucianus picornavirus	6.10E+07
541	Kobuvirus 16317	2.03E+07	681	Bat picornavirus isolate BtRs-PicoV	1.13E+06

SEQ ID NO	IRES	Expression	SEQ ID NO	IRES	Expression
542	Aichivirus Chshc7	1.87E+07	682	Rhinolophus picornavirus strain Guizhou-Rr100	N/A
543	Aichivirus Goiania	1.66E+07	683	Rhinolophus picornavirus strain Henan-Rf265	3.85E+05
544	Aichivirus EHP4	1.78E+07	684	Human enterovirus C105	5.49E+05
545	Aichivirus DVI2169	2.98E+06	685	Human poliovirus 1 strain NIE1116623	3.94E+05
546	Aichivirus DVI2321	6.63E+07	686	Human enterovirus 109	4.92E+05
547	Aichivirus rat08	3.51E+07	687	Human poliovirus 2 strain NIE0811460	2.59E+07
548	Aichivirus Rt386	5.71E+07	688	Bovine picornavirus	3.82E+06
549	Norway Rat Pestivirus	N/A	689	Human poliovirus 1 strain EQG1419328	2.44E+05
550	Porcine Kobuvirus GS2	44200000	690	Human poliovirus 2 isolate IS_061	5.84E+06
551	Kobuvirus SZAL6	98850000	691	Coxsackievirus B5	N/A
552	Kobuvirus sheep TB3	N/A	692	Coxsackievirus A10	N/A

EXAMPLE 9

Expression and functional stability by IRES in Jurkat cells.

[0731] 2 sets of constructs including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and a subset of previously tested IRES were circularized. 60,000 Jurkat cells were electroporated with 1 µg of each circularization reaction. Luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation. A CVB3 IRES construct was included in both sets for comparison between sets and to previously defined IRES efficacy. CVB1 and Salivirus A SZ1 IRES constructs produced the most expression at 24h. Data can be found in **FIGs. 4A and 4B**.

[0732] Functional stability of the IRES constructs in each round of electroporated Jurkat cells was measured over 3 days. Luminescence from secreted *Gaussia* luciferase in supernatant was measured every 24 hours after electroporation of 60,000 cells with 1 µg of each circularization reaction, followed by complete media replacement (**FIGs. 5A and 5B**).

[0733] Salivirus A SZ1 and Salivirus A BN2 IRES constructs had high functional stability compared to other constructs.

EXAMPLE 10

Expression, functional stability, and cytokine release of circular and linear RNA in Jurkat

cells.

[0734] A construct including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and a Salivirus FHB IRES was circularized. mRNA including a *Gaussia* luciferase expression sequence and a ~150nt polyA tail, and modified to replace 100% of uridine with 5-methoxy uridine (5moU) is commercially available and was purchased from Trilink. 5moU nucleotide modifications have been shown to improve mRNA stability and expression (Bioconjug Chem. 2016 Mar 16;27(3):849-53). Expression of modified mRNA, circularization reactions (unpure), and circRNA purified by size exclusion HPLC (pure) in Jurkat cells were measured and compared (**FIG. 6A**). Luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation of 60,000 cells with 1 µg of each RNA species.

[0735] Luminescence from secreted *Gaussia* luciferase in supernatant was measured every 24 hours after electroporation of 60,000 cells with 1 µg of each RNA species, followed by complete media replacement. A comparison of functional stability data of modified mRNA and circRNA in Jurkat cells over 3 days is in **FIG. 6B**.

[0736] IFN γ (**FIG. 7A**), IL-6 (**FIG. 7B**), IL-2 (**FIG. 7C**), RIG-I (**FIG. 7D**), IFN- β 1 (**FIG. 7E**), and TNF α (**FIG. 7F**) transcript induction was measured 18 hours after electroporation of 60,000 Jurkat cells with 1 µg of each RNA species described above and 3p-hpRNA (5' triphosphate hairpin RNA, which is a known RIG-I agonist).

EXAMPLE 11

Expression of circular and linear RNA in monocytes and macrophages.

[0737] A construct including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and a Salivirus FHB IRES was circularized. mRNA including a *Gaussia* luciferase expression sequence and a ~150nt polyA tail, and modified to replace 100% of uridine with 5-methoxy uridine (5moU) was purchased from Trilink. Expression of circular and modified mRNA was measured in human primary monocytes (**FIG. 8A**) and human primary macrophages (**FIG. 8B**). Luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation of 60,000 cells with 1 µg of each RNA species. Luminescence was also measured 4 days after electroporation of human primary macrophages with media changes every 24 hours (**FIG. 8C**). The results can be found in **FIG. 8**. The difference in luminescence was statistically significant in each case ($p < 0.05$).

EXAMPLE 12

Expression and functional stability by IRES in primary T cells.

[0738] Constructs including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and a subset of previously tested IRES were circularized and reaction products were purified by size exclusion HPLC. 150,000 primary human CD3+ T cells were electroporated with 1 µg of each circRNA. Luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation (**FIG. 9A**). Aichi Virus and CVB3 IRES constructs had the most expression at 24 hours.

[0739] Luminescence was also measured every 24 hours after electroporation for 3 days in order to compare functional stability of each construct (**FIG. 9B**). The construct with a Salivirus A SZ1 IRES was the most stable.

EXAMPLE 13

Expression and functional stability of circular and linear RNA in primary T cells and PBMCs.

[0740] Constructs including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and a Salivirus A SZ1 IRES or Salivirus FHB IRES were circularized. mRNA including a *Gaussia* luciferase expression sequence and a ~150nt polyA tail, and modified to replace 100% of uridine with 5-methoxy uridine (5moU) and was purchased from Trilink. Expression of Salivirus A SZ1 IRES HPLC purified circular and modified mRNA was measured in human primary CD3+ T cells. Expression of Salivirus FHB HPLC purified circular, unpurified circular and modified mRNA was measured in human PBMCs. Luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation of 150,000 cells with 1 µg of each RNA species. Data for primary human T cells is shown in FIGs. 10A and 10B, and data for PBMCs is shown in FIG. 10C. The difference in expression between the purified circular RNA and unpurified circular RNA or linear RNA was significant in each case ($p < 0.05$).

[0741] Luminescence from secreted *Gaussia* luciferase in primary T cell supernatant was measured every 24 hours after electroporation over 3 days in order to compare construct functional stability. Data is shown in FIG. 10B. The difference in relative luminescence from the day 1 measurement between purified circular RNA and linear RNA was significant at both day 2 and day 3 for primary T cells.

EXAMPLE 14

Circularization efficiency by permutation site in Anabaena intron.

[0742] RNA constructs including a CVB3 IRES, a *Gaussia* luciferase expression sequence, anabaena intron / exon regions, spacers, internal homology regions, and homology arms were produced. Circularization efficiency of constructs using the traditional anabaena intron permutation site and 5 consecutive permutations sites in P9 was measured by HPLC. HPLC chromatograms for the 5 consecutive permutation sites in P9 are shown in FIG. 11A.

[0743] Circularization efficiency was measured at a variety of permutation sites. Circularization efficiency is defined as the area under the HPLC chromatogram curve for each of: circRNA / (circRNA + precursor RNA). Ranked quantification of circularization efficiency at each permutation site is in FIG. 11B. 3 permutation sites (indicated in FIG. 11B) were selected for further investigation.

[0744] Circular RNA in this example was circularized by *in vitro* transcription (IVT) then purified via spin column. Circularization efficiency for all constructs would likely be higher if the additional step of incubation with Mg²⁺ and guanosine nucleotide were included; however, removing this step allowed for comparison between, and optimization of, circular RNA constructs. This level of optimization is especially useful for maintaining high circularization efficiency with large RNA constructs, such as those encoding chimeric antigen receptors.

EXAMPLE 15*Circularization efficiency of alternative introns.*

[0745] Precursor RNA containing a permuted group 1 intron of variable species origin or permutation site and several constant elements including: a CVB3 IRES, a *Gaussia* luciferase expression sequence, spacers, internal homology regions, and homology arms were created. Circularization data can be found in FIG. 12. FIG. 12A shows chromatograms resolving precursor, CircRNA and introns. Fig. 12B provides ranked quantification of circularization efficiency, based on the chromatograms shown in Fig. 12A, as a function of intron construct.

[0746] Circular RNA in this example was circularized by *in vitro* transcription (IVT) then spin column purification. Circularization efficiency for all constructs would likely be higher if the additional step of incubation with Mg²⁺ and guanosine nucleotide were included; however, removing this step allows for comparison between, and optimization of, circular RNA constructs. This level of optimization is especially useful for maintaining high circularization efficiency with large RNA constructs, such as those encoding chimeric antigen receptors.

EXAMPLE 16

Circularization efficiency by homology arm presence or length.

[0747] RNA constructs including a CVB3 IRES, a *Gaussia* luciferase expression sequence, anabaena intron / exon regions, spacers, and internal homology regions were produced. Constructs representing 3 anabaena intron permutation sites were tested with 30 nt, 25% GC homology arms or without homology arms ("NA"). These constructs were allowed to circularize without an Mg^{2+} incubation step. Circularization efficiency was measured and compared. Data can be found in FIGs. 13A and 13B. Circularization efficiency was higher for each construct lacking homology arms. FIG. 13A provides ranked quantification of circularization efficiency; FIG. 13B provides chromatograms resolving precursor, circRNA and introns.

[0748] For each of the 3 permutation sites, constructs were created with 10 nt, 20 nt, and 30 nt arm lengths and 25%, 50%, and 75% GC content. Splicing efficiency of these constructs was measured and compared to constructs without homology arms (FIG. 14). Splicing efficiency is defined as the proportion of free introns relative to the total RNA in the splicing reaction.

[0749] FIG. 15 A (left) shows HPLC chromatograms indicating the contribution of strong homology arms to improved splicing efficiency. Top left: 75% GC content, 10 nt homology arms. Center left: 75% GC content, 20 nt homology arms. Bottom left: 75% GC content, 30 nt homology arms.

[0750] FIG. 15 A (right) shows HPLC chromatograms showing increased splicing efficiency paired with increased nicking, appearing as a shoulder on the circRNA peak. Top right: 75% GC content, 10 nt homology arms. Center right: 75% GC content, 20 nt homology arms. Bottom right: 75% GC content, 30 nt homology arms.

[0751] FIG. 15 B (left) shows select combinations of permutation sites and homology arms hypothesized to demonstrate improved circularization efficiency.

[0752] FIG. 15 B (right) shows select combinations of permutation sites and homology arms hypothesized to demonstrate improved circularization efficiency, treated with *E. coli* polyA polymerase.

[0753] Circular RNA in this example was circularized by *in vitro* transcription (IVT) then spin-column purified. Circularization efficiency for all constructs would likely be higher if an additional Mg^{2+} incubation step with guanosine nucleotide were included; however, removing this step allowed for comparison between, and optimization of, circular RNA

constructs. This level of optimization is especially useful for maintaining high circularization efficiency with large RNA constructs, such as those encoding chimeric antigen receptors.

EXAMPLE 17

Circular RNA encoding chimeric antigen receptors

[0754] Constructs including anabaena intron / exon regions, a Kymriah chimeric antigen receptors (CAR) expression sequence, and a CVB3 IRES were circularized. 100,000 human primary CD3⁺ T cells were electroporated with 500ng of circRNA and co-cultured for 24 hours with Raji cells stably expressing GFP and firefly luciferase. Effector to target ratio (E:T ratio) 0.75:1. 100,000 human primary CD3⁺ T cells were mock electroporated and co-cultured as a control (FIG. 16).

[0755] Sets of 100,000 human primary CD3⁺ T cells were mock electroporated or electroporated with 1 µg of circRNA then co-cultured for 48 hours with Raji cells stably expressing GFP and firefly luciferase E:T ratio 10:1 (FIG. 17).

[0756] Quantification of specific lysis of Raji target cells was determined by detection of firefly luminescence (FIG. 18). 100,000 human primary CD3⁺ T cells either mock electroporated or electroporated with circRNA encoding different CAR sequences were co-cultured for 48 hours with Raji cells stably expressing GFP and firefly luciferase. % Specific lysis defined as $1 - [\text{CAR condition luminescence}] / [\text{mock condition luminescence}]$. E:T ratio 10:1.

EXAMPLE 18

Expression and functional stability of circular and linear RNA in Jurkat cells and resting human T cells.

[0757] Constructs including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and a subset of previously tested IRES were circularized and reaction products were purified by size exclusion HPLC. 150,000 Jurkat cells were electroporated with 1 µg of circular RNA or 5moU-mRNA. Luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation (FIG. 19A left). 150,000 resting primary human CD3⁺ T cells (10 days post-stimulation) were electroporated with 1 µg of circular RNA or 5moU-mRNA. Luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation (FIG. 19A right).

[0758] Luminescence from secreted *Gaussia* luciferase in supernatant was measured every 24 hours after electroporation, followed by complete media replacement. Functional

stability data shown in FIG. 19B. Circular RNA had more functional stability than linear RNA in each case, with a more pronounced difference in Jurkat cells.

EXAMPLE 19

IFN- β 1, RIG-I, IL-2, IL-6, IFN γ , and TNF α transcript induction of cells electroporated with linear RNA or varying circular RNA constructs.

[0759] Constructs including anabaena intron / exon regions, a *Gaussia* luciferase expression sequence, and a subset of previously tested IRES were circularized and reaction products were purified by size exclusion HPLC. 150,000 CD3+ human T cells were electroporated with 1 μ g of circular RNA, 5moU-mRNA, or immunostimulatory positive control poly inosine:cytosine. IFN- β 1 (FIG. 20A), RIG-I (FIG. 20B), IL-2 (FIG. 20C), IL-6 (FIG. 20D), IFN γ (FIG. 20E), and TNF α (FIG. 20F) transcript induction was measured 18 hours after electroporation.

EXAMPLE 20

Specific lysis of target cells and IFN γ transcript induction by CAR expressing cells electroporated with different amounts of circular or linear RNA; specific lysis of target and non-target cells by CAR expressing cells at different E:T ratios.

[0760] Constructs including anabaena intron / exon regions, an anti-CD19 CAR expression sequence, and a CVB3 IRES were circularized and reaction products were purified by size exclusion HPLC. 150,000 human primary CD3+ T cells either mock electroporated or electroporated with different quantities of circRNA encoding an anti-CD19 CAR sequence were co-cultured for 12 hours with Raji cells stably expressing GFP and firefly luciferase at an E:T ratio of 2:1. Specific lysis of Raji target cells was determined by detection of firefly luminescence (FIG. 21A). %Specific lysis was defined as 1-[CAR condition luminescence]/[mock condition luminescence]. IFN γ transcript induction was measured 24 hours after electroporation (FIG. 21B).

[0761] 150,000 human primary CD3+ T cells were either mock electroporated or electroporated with 500ng circRNA or m1 ψ -mRNA encoding an anti-CD19 CAR sequence, then co-cultured for 24 hours with Raji cells stably expressing firefly luciferase at different E:T ratios. % Specific lysis of Raji target cells was determined by detection of firefly luminescence (FIG. 22A). % Specific lysis was defined as 1-[CAR condition luminescence]/[mock condition luminescence].

[0762] CAR expressing T cells were also co-cultured for 24 hours with Raji or K562 cells

stably expressing firefly luciferase at different E:T ratios. Specific lysis of Raji target cells or K562 non-target cells was determined by detection of firefly luminescence (FIG. 22B). % Specific lysis is defined as $1 - [\text{CAR condition luminescence}] / [\text{mock condition luminescence}]$.

EXAMPLE 21

Specific lysis of target cells by T cells electroporated with circular RNA or linear RNA encoding a CAR.

[0763] Constructs including anabaena intron / exon regions, an anti-CD19 CAR expression sequence, and a CVB3 IRES were circularized and reaction products were purified by size exclusion HPLC. Human primary CD3⁺ T cells were electroporated with 500 ng of circular RNA or an equimolar quantity of m1ψ-mRNA, each encoding a CD19-targeted CAR. Raji cells were added to CAR-T cell cultures over 7 days at an E:T ratio of 10:1. % Specific lysis was measured for both constructs at 1, 3, 5, and 7 days (FIG. 23).

EXAMPLE 22

Specific lysis of Raji cells by T cells expressing an anti-CD19 CAR or an anti-BCMA CAR.

[0764] Constructs including anabaena intron / exon regions, anti-CD19 or anti-BCMA CAR expression sequence, and a CVB3 IRES were circularized and reaction products were purified by size exclusion HPLC. 150,000 primary human CD3⁺ T cells were electroporated with 500ng of circRNA, then were co-cultured with Raji cells at an E:T ratio of 2:1. % Specific lysis was measured 12 hours after electroporation (FIG. 24).

EXAMPLE 23

Example 23A: Synthesis of compounds

[0765] Synthesis of representative ionizable lipids of the invention are described in PCT applications PCT/US2016/052352, PCT/US2016/068300, PCT/US2010/061058, PCT/US2018/058555, PCT/US2018/053569, PCT/US2017/028981, PCT/US2019/025246, PCT/US2018/035419, PCT/US2019/015913, and US applications with publication numbers 20190314524, 20190321489, and 20190314284, the contents of each of which are incorporated herein by reference in their entireties.

Example 23B: Synthesis of compounds

[0766] Synthesis of representative ionizable lipids of the invention are described in US patent publication number US20170210697A1, the contents of of which is incorporated herein by reference in its entirety.

EXAMPLE 24*Protein expression by organ*

[0767] Circular or linear RNA encoding FLuc was generated and loaded into transfer vehicles with the following formulation: 50% ionizable lipid 15 in Table 10b, 10% DSPC, 1.5% PEG-DMG, 38.5% cholesterol. CD-1 mice were dosed at 0.2 mg/kg and luminescence was measured at 6 hours (live IVIS) and 24 hours (live IVIS and *ex vivo* IVIS). Total Flux (photons/second over a region of interest) of the liver, spleen, kidney, lung, and heart was measured (FIGs. 25 and 26).

EXAMPLE 25*Distribution of expression in the spleen*

[0768] Circular or linear RNA encoding GFP is generated and loaded into transfer vehicles with the following formulation: 50% ionizable lipid 15 in Table 10b, 10% DSPC, 1.5% PEG-DMG, 38.5% cholesterol. The formulation is administered to CD-1 mice. Flow cytometry is run on spleen cells to determine the distribution of expression across cell types.

EXAMPLE 26*Production of nanoparticle compositions*

[0769] In order to investigate safe and efficacious nanoparticle compositions for use in the delivery of circular RNA to cells, a range of formulations are prepared and tested. Specifically, the particular elements and ratios thereof in the lipid component of nanoparticle compositions are optimized.

[0770] Nanoparticles can be made in a 1 fluid stream or with mixing processes such as microfluidics and T-junction mixing of two fluid streams, one of which contains the circular RNA and the other has the lipid components.

[0771] Lipid compositions are prepared by combining an ionizable lipid, optionally a helper lipid (such as DOPE, DSPC, or oleic acid obtainable from Avanti Polar Lipids, Alabaster, AL), a PEG lipid (such as 1,2-dimyristoyl-sn-glycerol methoxypolyethylene glycol, also known as PEG-DMG, obtainable from Avanti Polar Lipids, Alabaster, AL), and a structural lipid such as cholesterol at concentrations of about, *e.g.*, 40 or 50 mM in a solvent, *e.g.*, ethanol. Solutions should be refrigerated for storage at, for example, -20 °C. Lipids are combined to yield desired molar ratios (see, for example, Tables 31a and 31b below) and diluted with water and ethanol to a final lipid concentration of *e.g.*, between about 5.5 mM and about 25 mM.

Table 31a

Formulation number	Description
1	Aliquots of 50 mg/mL ethanolic solutions of C12-200, DOPE, Chol and DMG-PEG2K (40:30:25:5) are mixed and diluted with ethanol to 3 mL final volume. Separately, an aqueous buffered solution (10 mM citrate/150 mM NaCl, pH 4.5) of circRNA is prepared from a 1 mg/mL stock. The lipid solution is injected rapidly into the aqueous circRNA solution and shaken to yield a final suspension in 20% ethanol. The resulting nanoparticle suspension is filtered, diafiltrated with 1×PBS (pH 7.4), concentrated and stored at 2-8° C.
2	Aliquots of 50 mg/mL ethanolic solutions of DODAP, DOPE, cholesterol and DMG-PEG2K (18:56:20:6) are mixed and diluted with ethanol to 3 mL final volume. Separately, an aqueous buffered solution (10 mM citrate/150 mM NaCl, pH 4.5) of EPO circRNA is prepared from a 1 mg/mL stock. The lipid solution is injected rapidly into the aqueous circRNA solution and shaken to yield a final suspension in 20% ethanol. The resulting nanoparticle suspension is filtered, diafiltrated with 1×PBS (pH 7.4), concentrated and stored at 2-8° C. Final concentration=1.35 mg/mL EPO circRNA (encapsulated). Zave=75.9 nm (Dv(50)=57.3 nm; Dv(90)=92.1 nm).
3	Aliquots of 50 mg/mL ethanolic solutions of HGT4003, DOPE, cholesterol and DMG-PEG2K (50:25:20:5) are mixed and diluted with ethanol to 3 mL final volume. Separately, an aqueous buffered solution (10 mM citrate/150 mM NaCl, pH 4.5) of circRNA is prepared from a 1 mg/mL stock. The lipid solution is injected rapidly into the aqueous circRNA solution and shaken to yield a final suspension in 20% ethanol. The resulting nanoparticle suspension is filtered, diafiltrated with 1×PBS (pH 7.4), concentrated and stored at 2-8° C.
4	Aliquots of 50 mg/mL ethanolic solutions of ICE, DOPE and DMG-PEG2K (70:25:5) are mixed and diluted with ethanol to 3 mL final volume. Separately, an aqueous buffered solution (10 mM citrate/150 mM NaCl, pH 4.5) of circRNA is prepared from a 1 mg/mL stock. The lipid solution is injected rapidly into the aqueous circRNA solution and shaken to yield a final suspension in 20% ethanol. The resulting nanoparticle suspension is filtered, diafiltrated with 1×PBS (pH 7.4), concentrated and stored at 2-8° C.
5	Aliquots of 50 mg/mL ethanolic solutions of HGT5000, DOPE, cholesterol and DMG-PEG2K (40:20:35:5) are mixed and diluted with ethanol to 3 mL final volume. Separately, an aqueous buffered solution (10 mM citrate/150 mM NaCl, pH 4.5) of EPO circRNA is prepared from a 1 mg/mL stock. The lipid solution is injected rapidly into the aqueous circRNA solution and shaken to yield a final suspension in 20% ethanol. The resulting nanoparticle suspension is filtered, diafiltrated with 1×PBS (pH 7.4), concentrated and stored at 2-8° C. Final concentration=1.82 mg/mL EPO mRNA (encapsulated). Zave=105.6 nm (Dv(50)=53.7 nm; Dv(90)=157 nm).
6	Aliquots of 50 mg/mL ethanolic solutions of HGT5001, DOPE, cholesterol and DMG-PEG2K (40:20:35:5) are mixed and diluted with ethanol to 3 mL final volume. Separately, an aqueous buffered solution (10 mM citrate/150 mM NaCl, pH 4.5) of EPO circRNA is prepared from a 1 mg/mL stock. The lipid solution is injected rapidly into the aqueous circRNA solution and shaken to yield a final suspension in 20% ethanol. The resulting nanoparticle suspension is filtered, diafiltrated with 1×PBS (pH 7.4), concentrated and stored at 2-8° C.

[0772] In some embodiments, transfer vehicle has a formulation as described in Table

31a.

Table 31b

Composition (mol %)	Components
40:20:38.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
45:15:38.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
50:10:38.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
55:5:38.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG

Composition (mol %)	Components
60:5:33.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
45:20:33.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
50:20:28.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
55:20:23.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
60:20:18.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
40:15:43.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
50:15:33.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
55:15:28.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
60:15:23.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
40:10:48.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
45:10:43.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
55:10:33.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
60:10:28.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
40:5:53.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
45:5:48.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
50:5:43.5:1.5	Compound:Phospholipid:Phytosterol*:PEG-DMG
40:20:40:0	Compound:Phospholipid:Phytosterol*:PEG-DMG
45:20:35:0	Compound:Phospholipid:Phytosterol*:PEG-DMG
50:20:30:0	Compound:Phospholipid:Phytosterol*:PEG-DMG
55:20:25:0	Compound:Phospholipid:Phytosterol*:PEG-DMG
60:20:20:0	Compound:Phospholipid:Phytosterol*:PEG-DMG
40:15:45:0	Compound:Phospholipid:Phytosterol*:PEG-DMG

[0773] In some embodiments, transfer vehicle has a formulation as described in Table 31b.

[0774] For nanoparticle compositions including circRNA, solutions of the circRNA at concentrations of 0.1 mg/ml in deionized water are diluted in a buffer, *e.g.*, 50 mM sodium citrate buffer at a pH between 3 and 4 to form a stock solution. Alternatively, solutions of the circRNA at concentrations of 0.15 mg/ml in deionized water are diluted in a buffer, *e.g.*, 6.25 mM sodium acetate buffer at a pH between 3 and 4.5 to form a stock solution.

[0775] Nanoparticle compositions including a circular RNA and a lipid component are prepared by combining the lipid solution with a solution including the circular RNA at lipid component to circRNA wt:wt ratios between about 5:1 and about 50:1. The lipid solution is rapidly injected using, *e.g.*, a NanoAssembler microfluidic based system at flow rates between about 10 ml/min and about 18 ml/min or between about 5 ml/min and about 18 ml/min into the circRNA solution, to produce a suspension with a water to ethanol ratio between about 1:1 and about 4:1.

[0776] Nanoparticle compositions can be processed by dialysis to remove ethanol and achieve buffer exchange. Formulations are dialyzed twice against phosphate buffered saline (PBS), pH 7.4, at volumes 200 times that of the primary product using Slide-A-Lyzer cassettes (Thermo Fisher Scientific Inc., Rockford, IL) with a molecular weight cutoff of 10 kDa or 20 kDa. The formulations are then dialyzed overnight at 4° C. The resulting nanoparticle suspension is filtered through 0.2 µm sterile filters (Sarstedt, Nümbrecht, Germany) into glass vials and sealed with crimp closures. Nanoparticle composition solutions of 0.01 mg/ml to 0.15 mg/ml are generally obtained.

[0777] The method described above induces nano-precipitation and particle formation.

[0778] Alternative processes including, but not limited to, T-junction and direct injection, may be used to achieve the same nano-precipitation. B. Characterization of nanoparticle compositions

[0779] A Zetasizer Nano ZS (Malvern Instruments Ltd, Malvern, Worcestershire, UK) can be used to determine the particle size, the polydispersity index (PDI) and the zeta potential of the nanoparticle compositions in 1×PBS in determining particle size and 15 mM PBS in determining zeta potential.

[0780] Ultraviolet-visible spectroscopy can be used to determine the concentration of circRNA in nanoparticle compositions. 100 µL of the diluted formulation in 1×PBS is added to 900 µL of a 4:1 (v/v) mixture of methanol and chloroform. After mixing, the absorbance spectrum of the solution is recorded, for example, between 230 nm and 330 nm on a DU 800

spectrophotometer (Beckman Coulter, Beckman Coulter, Inc., Brea, CA). The concentration of circRNA in the nanoparticle composition can be calculated based on the extinction coefficient of the circRNA used in the composition and on the difference between the absorbance at a wavelength of, for example, 260 nm and the baseline value at a wavelength of, for example, 330 nm.

[0781] A QUANT-IT™ RIBOGREEN® RNA assay (Invitrogen Corporation Carlsbad, CA) can be used to evaluate the encapsulation of circRNA by the nanoparticle composition. The samples are diluted to a concentration of approximately 5 µg/mL or 1 µg/mL in a TE buffer solution (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). 50 µL of the diluted samples are transferred to a polystyrene 96 well plate and either 50 µL of TE buffer or 50 µL of a 2-4% Triton X-100 solution is added to the wells. The plate is incubated at a temperature of 37° C for 15 minutes. The RIBOGREEN® reagent is diluted 1:100 or 1:200 in TE buffer, and 100 µL of this solution is added to each well. The fluorescence intensity can be measured using a fluorescence plate reader (Wallac Victor 1420 Multilabel Counter; Perkin Elmer, Waltham, MA) at an excitation wavelength of, for example, about 480 nm and an emission wavelength of, for example, about 520 nm. The fluorescence values of the reagent blank are subtracted from that of each of the samples and the percentage of free circRNA is determined by dividing the fluorescence intensity of the intact sample (without addition of Triton X-100) by the fluorescence value of the disrupted sample (caused by the addition of Triton X-100). C.

In vivo formulation studies:

[0782] In order to monitor how effectively various nanoparticle compositions deliver circRNA to targeted cells, different nanoparticle compositions including circRNA are prepared and administered to rodent populations. Mice are intravenously, intramuscularly, intraarterially, or intratumorally administered a single dose including a nanoparticle composition with a lipid nanoparticle formulation. In some instances, mice may be made to inhale doses. Dose sizes may range from 0.001 mg/kg to 10 mg/kg, where 10 mg/kg describes a dose including 10 mg of a circRNA in a nanoparticle composition for each 1 kg of body mass of the mouse. A control composition including PBS may also be employed.

[0783] Upon administration of nanoparticle compositions to mice, dose delivery profiles, dose responses, and toxicity of particular formulations and doses thereof can be measured by enzyme-linked immunosorbent assays (ELISA), bioluminescent imaging, or other methods. Time courses of protein expression can also be evaluated. Samples collected from the rodents for evaluation may include blood and tissue (for example, muscle tissue from the site of an

intramuscular injection and internal tissue); sample collection may involve sacrifice of the animals.

Higher levels of protein expression induced by administration of a composition including a circRNA will be indicative of higher circRNA translation and/or nanoparticle composition circRNA delivery efficiencies. As the non-RNA components are not thought to affect translational machineries themselves, a higher level of protein expression is likely indicative of a higher efficiency of delivery of the circRNA by a given nanoparticle composition relative to other nanoparticle compositions or the absence thereof.

EXAMPLE 27

Characterization of nanoparticle compositions

[0784] A Zetasizer Nano ZS (Malvern Instruments Ltd, Malvern, Worcestershire, UK) can be used to determine the particle size, the polydispersity index (PDI) and the zeta potential of the transfer vehicle compositions in 1×PBS in determining particle size and 15 mM PBS in determining zeta potential.

[0785] Ultraviolet-visible spectroscopy can be used to determine the concentration of a therapeutic and/or prophylactic (*e.g.*, RNA) in transfer vehicle compositions. 100 µL of the diluted formulation in 1×PBS is added to 900 µL of a 4:1 (v/v) mixture of methanol and chloroform. After mixing, the absorbance spectrum of the solution is recorded, for example, between 230 nm and 330 nm on a DU 800 spectrophotometer (Beckman Coulter, Beckman Coulter, Inc., Brea, CA). The concentration of therapeutic and/or prophylactic in the transfer vehicle composition can be calculated based on the extinction coefficient of the therapeutic and/or prophylactic used in the composition and on the difference between the absorbance at a wavelength of, for example, 260 nm and the baseline value at a wavelength of, for example, 330 nm.

[0786] For transfer vehicle compositions including RNA, a QUANT-IT™ RIBOGREEN® RNA assay (Invitrogen Corporation Carlsbad, CA) can be used to evaluate the encapsulation of RNA by the transfer vehicle composition. The samples are diluted to a concentration of approximately 5 µg/mL or 1 µg/mL in a TE buffer solution (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). 50 µL of the diluted samples are transferred to a polystyrene 96 well plate and either 50 µL of TE buffer or 50 µL of a 2-4% Triton X-100 solution is added to the wells. The plate is incubated at a temperature of 37° C for 15 minutes. The RIBOGREEN® reagent is diluted 1:100 or 1:200 in TE buffer, and 100 µL of this solution is added to each well. The fluorescence intensity can be measured using a fluorescence plate

reader (Wallac Victor 1420 Multilabel Counter; Perkin Elmer, Waltham, MA) at an excitation wavelength of, for example, about 480 nm and an emission wavelength of, for example, about 520 nm. The fluorescence values of the reagent blank are [0787] subtracted from that of each of the samples and the percentage of free RNA is determined by dividing the fluorescence intensity of the intact sample (without addition of Triton X-100) by the fluorescence value of the disrupted sample (caused by the addition of Triton X-100).

EXAMPLE 28

T cell targeting

[0788] To target transfer vehicles to T-cells, T cell antigen binders, *e.g.*, anti-CD8 antibodies, are coupled to the surface of the transfer vehicle. Anti-T cell antigen antibodies are mildly reduced with an excess of DTT in the presence of EDTA in PBS to expose free hinge region thiols. To remove DTT, antibodies are passed through a desalting column. The heterobifunctional cross-linker SM(PEG)24 is used to anchor antibodies to the surface of circRNA-loaded transfer vehicles (Amine groups are present in the head groups of PEG lipids, free thiol groups on antibodies were created by DTT, SM(PEG)24 cross-links between amines and thiol groups). Transfer vehicles are first incubated with an excess of SM(PEG)24 and centrifuged to remove unreacted cross-linker. Activated transfer vehicles are then incubated with an excess of reduced anti-T cell antigen antibody. Unbound antibody is removed using a centrifugal filtration device.

EXAMPLE 29

RNA containing transfer vehicle using RV88.

[0789] In this example RNA containing transfer vehicles are synthesized using the 2-D vortex microfluidic chip with the cationic lipid RV88 for delivery of circRNA.

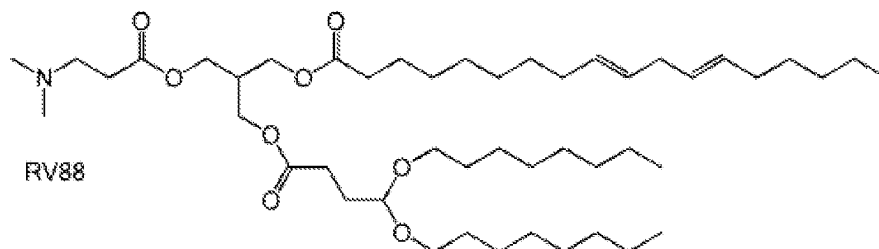


Table 32a

Materials and Instrument	Vendor	Cat #
1M Tris-HCl, pH 8.0, Sterile	Teknova	T1080
5M Sodium Chloride solution	Teknova	S0250
QB Citrate buffer, pH 6.0 (100 mM)	Teknova	Q2446
Nuclease-free water	Ambion	AM9937
Triton X-100	Sigma-Aldrich	T8787-100ML
RV88	GVK bio	
DSPC	Lipoid	556500
Cholesterol	Sigma	C3045-5G
PEG2K	Avanti Polar Lipids	880150
Ethanol	Acros Organic	615090010
5 mL Borosilicate glass vials	Thermo Scientific	ST5-20
PD MiniTrap G-25 Desalting Columns	GE Healthcare	VWR Cat. #95055-984
Quant-iT RiboGreen RNA Assay kit	Molecular Probes/ Life Technologies	R11490
Black 96-well microplates	Greiner	655900

[0790] RV88, DSPC, and cholesterol all being prepared in ethanol at a concentration of 10 mg/ml in borosilica vials. The lipid 14:0-PEG2K PE is prepared at a concentration of 4 mg/ml also in a borosilica glass vial. Dissolution of lipids at stock concentrations is attained by sonication of the lipids in ethanol for 2 min. The solutions are then heated on an orbital tilting shaker set at 170 rpm at 37 °C for 10 min. Vials are then equilibrated at 26 °C for a minimum of 45 min. The lipids are then mixed by adding volumes of stock lipid as shown in Table 32b. The solution is then adjusted with ethanol such that the final lipid concentration was 7.92 mg/ml.

Table 32b

Composition	MW	%	nmoles	mg	Stock (mg/ml)	ul	Ethanol (ul)
RV88	794.2	40%	7200	5.72	10	571.8	155.3
DSPC	790.15	10%	1800	1.42	10	142.2	
Cholesterol	386.67	48%	8640	3.34	10	334.1	
PEG2K	2693.3	2%	360	0.97	4	242.4	

[0791] RNA is prepared as a stock solution with 75 mM Citrate buffer at pH 6.0 and a concentration of RNA at 1.250 mg/ml. The concentration of the RNA is then adjusted to 0.1037 mg/ml with 75 mM citrate buffer at pH 6.0, equilibrated to 26 °C. The solution is then incubated at 26 °C for a minimum of 25 min.

[0792] The microfluidic chamber is cleaned with ethanol and neMYSIS syringe pumps are prepared by loading a syringe with the RNA solution and another syringe with the ethanolic lipid. Both syringes are loaded and under the control of neMESYS software. The solutions are then applied to the mixing chip at an aqueous to organic phase ratio of 2 and a total flow rate of 22 ml/min (14.67 ml/min for RNA and 7.33 ml/min for the lipid solution). Both pumps are started synchronously. The mixer solution that flowed from the microfluidic chip is collected in 4x1 ml fractions with the first fraction being discarded as waste. The remaining solution containing the RNA-liposomes is exchanged by using G-25 mini desalting columns to 10 mM Tris-HCl, 1 mM EDTA, at pH 7.5. Following buffer exchange, the materials are characterized for size, and RNA entrapment through DLS analysis and Ribogreen assays, respectively.

EXAMPLE 30

RNA containing transfer vehicle using RV94.

[0793] In this example, RNA containing liposome are synthesized using the 2-D vortex microfluidic chip with the cationic lipid RV94 for delivery of circRNA.

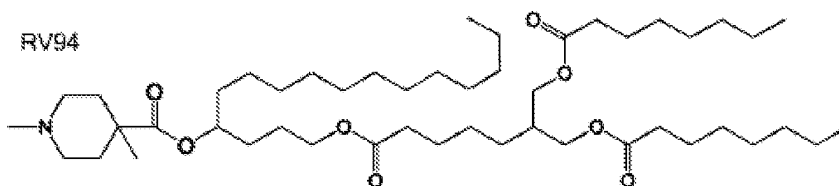


Table 33

<u>Materials and Instrument</u>	<u>Vendor</u>	<u>Cat #</u>
1M Tris-HCl, pH 8.0, Sterile	Teknova	T1080
5M Sodium Chloride solution	Teknova	S0250
QB Citrate buffer, pH 6.0 (100 mM)	Teknova	Q2446
Nuclease-free water	Ambion	AM9937
Triton X-100	Sigma-Aldrich	T8787-100ML
RV94	GVKbio	
DSPC	Lipoid	556500
Cholesterol	Sigma	C3045-5G
PEG2K	Avanti Polar Lipids	880150
Ethanol	Acros Organic	615090010
5 mL Borosilicate glass vials	Thermo Scientific	ST5-20
PD MiniTrap G-25 Desalting Columns	GE Healthcare	VWR Cat. #95055-984
Quant-IT RiboGreen RNA Assay kit	Molecular Probes/Life Technologies	R11490
Black 96-well microplates	Greiner	655900

[0794] The lipids were prepared as in Example 29 using the material amounts named in Table 34 to a final lipid concentration of 7.92 mg/ml.

Table 34

<u>Composition</u>	<u>MW</u>	<u>%</u>	<u>nmoles</u>	<u>mg</u>	<u>Stock (mg/ml)</u>	<u>ul</u>	<u>Ethanol (ul)</u>
RV94	808.22	40%	2880	2.33	10	232.8	155.3
DSPC	790.15	10%	720	0.57	10	56.9	
Cholesterol	386.67	48%	3456	1.34	10	133.6	
PEG2K	2693.3	2%	144	0.39	4	97.0	

[0795] The aqueous solution of circRNA is prepared as a stock solution with 75 mM Citrate buffer at pH 6.0 the circRNA at 1.250 mg/ml. The concentration of the RNA is then adjusted to 0.1037 mg/ml with 75 mM citrate buffer at pH 6.0, equilibrated to 26 °C. The solution is then incubated at 26 °C for a minimum of 25 min.

[0796] The microfluidic chamber is cleaned with ethanol and neMYSIS syringe pumps are prepared by loading a syringe with the RNA solution and another syringe with the ethanolic lipid. Both syringes are loaded and under the control of neMESYS software. The solutions are then applied to the mixing chip at an aqueous to organic phase ratio of 2 and a total flow rate of 22 ml/min (14.67 ml/min for RNA and 7.33 ml/min for the lipid solution). Both pumps are started synchronously. The mixer solution that flowed from the microfluidic chip is collected in 4x1 ml fractions with the first fraction being discarded as waste. The remaining solution containing the circRNA-transfer vehicles is exchanged by using G-25 mini desalting columns to 10 mM Tris-HCl, 1 mM EDTA, at pH 7.5, as described above. Following buffer exchange, the materials are characterized for size, and RNA entrapment through DLS analysis and Ribogreen assays, respectively. The biophysical analysis of the liposomes is shown in Table 35.

Table 35

<u>Sample Name</u>	<u>N:P Ratio</u>	<u>TFR</u>	<u>Ratio</u>	<u>RNA encapsulation amount</u>	<u>RNA encapsulation yield</u>	<u>size</u>	
		ml/min	(aqueous/ org phase)	(μ g/ml)	%	d.nm	PDI
SAM-RV94	8	22	2	31.46	86.9	113.1	0.12

EXAMPLE 31

General protocol for in line mixing.

[0797] Individual and separate stock solutions are prepared - one containing lipid and the other circRNA. Lipid stock containing a desired lipid or lipid mixture, DSPC, cholesterol and PEG lipid is prepared by solubilized in 90% ethanol. The remaining 10% is low pH citrate buffer. The concentration of the lipid stock is 4 mg/mL. The pH of this citrate buffer can range between pH 3 and pH 5, depending on the type of lipid employed. The circRNA is also solubilized in citrate buffer at a concentration of 4 mg/mL. 5 mL of each stock solution is prepared.

[0798] Stock solutions are completely clear and lipids are ensured to be completely solubilized before combining with circRNA. Stock solutions may be heated to completely

solubilize the lipids. The circRNAs used in the process may be unmodified or modified oligonucleotides and may be conjugated with lipophilic moieties such as cholesterol.

[0799] The individual stocks are combined by pumping each solution to a T-junction. A dual-head Watson-Marlow pump was used to simultaneously control the start and stop of the two streams. A 1.6mm polypropylene tubing is further downsized to 0.8mm tubing in order to increase the linear flow rate. The polypropylene line (ID = 0.8mm) are attached to either side of a T-junction. The polypropylene T has a linear edge of 1.6mm for a resultant volume of 4.1 mm³. Each of the large ends (1.6mm) of polypropylene line is placed into test tubes containing either solubilized lipid stock or solubilized circRNA. After the T-junction, a single tubing is placed where the combined stream exited. The tubing is then extended into a container with 2x volume of PBS, which is rapidly stirred. The flow rate for the pump is at a setting of 300 rpm or 110 mL/min. Ethanol is removed and exchanged for PBS by dialysis. The lipid formulations are then concentrated using centrifugation or diafiltration to an appropriate working concentration.

[0800] C57BL/6 mice (Charles River Labs, MA) receive either saline or formulated circRNA via tail vein injection. At various time points after administration, serum samples are collected by retroorbital bleed. Serum levels of Factor VII protein are determined in samples using a chromogenic assay (Biophen FVTI, Anicara Corporation, OH). To determine liver RNA levels of Factor VII, animals are sacrificed and livers are harvested and snap frozen in liquid nitrogen. Tissue lysates are prepared from the frozen tissues and liver RNA levels of Factor VII are quantified using a branched DNA assay (QuantiGene Assay, Panomics, CA).

[0801] FVII activity is evaluated in FVTI siRNA-treated animals at 48 hours after intravenous (bolus) injection in C57BL/6 mice. FVII is measured using a commercially available kit for determining protein levels in serum or tissue, following the manufacturer's instructions at a microplate scale. FVII reduction is determined against untreated control mice, and the results are expressed as % Residual FVII. Two dose levels (0.05 and 0.005 mg/kg FVII siRNA) are used in the screen of each novel liposome composition.

EXAMPLE 32

circRNA formulation using preformed vesicles.

[0802] Cationic lipid containing transfer vehicles are made using the preformed vesicle method. Cationic lipid, DSPC, cholesterol and PEG-lipid are solubilized in ethanol at a molar

ratio of 40/10/40/10, respectively. The lipid mixture is added to an aqueous buffer (50 mM citrate, pH 4) with mixing to a final ethanol and lipid concentration of 30% (vol/vol) and 6.1 mg/mL respectively and allowed to equilibrate at room temperature for 2 min before extrusion. The hydrated lipids are extruded through two stacked 80 nm pore-sized filters (Nuclepore) at 22°C using a Lipex Extruder (Northern Lipids, Vancouver, BC) until a vesicle diameter of 70-90 nm, as determined by Nicomp analysis, is obtained. For cationic lipid mixtures which do not form small vesicles, hydrating the lipid mixture with a lower pH buffer (50mM citrate, pH 3) to protonate the phosphate group on the DSPC headgroup helps form stable 70-90 nm vesicles.

[0803] The FVII circRNA (solubilised in a 50mM citrate, pH 4 aqueous solution containing 30% ethanol) is added to the vesicles, pre-equilibrated to 35°C, at a rate of ~5mL/min with mixing. After a final target circRNA/lipid ratio of 0.06 (wt wt) is achieved, the mixture is incubated for a further 30 min at 35°C to allow vesicle re-organization and encapsulation of the FVII RNA. The ethanol is then removed and the external buffer replaced with PBS (155mM NaCl, 3mM Na₂HPO₄, 1mM KH₂P0₄, pH 7.5) by either dialysis or tangential flow diafiltration. The final encapsulated circRNA-to-lipid ratio is determined after removal of unencapsulated RNA using size-exclusion spin columns or ion exchange spin columns.

EXAMPLE 33

Expression of trispecific antigen binding proteins from engineered circular RNA

[0804] Circular RNAs are designed to include: (1) a 3' post splicing group I intron fragment; (2) an Internal Ribosome Entry Site (IRES); (3) a trispecific antigen-binding protein coding region; and (4) a 3' homology region. The trispecific antigen-binding protein regions are constructed to produce an exemplary trispecific antigen-binding protein that will bind to a target antigen, *e.g.*, GPC3.

Generation of a scFv CD3 binding domain

[0805] The human CD3epsilon chain canonical sequence is Uniprot Accession No. P07766. The human CD3gamma chain canonical sequence is Uniprot Accession No. P09693. The human CD3delta chain canonical sequence is Uniprot Accession No. P043234. Antibodies against CD3epsilon, CD3gamma or CD3delta are generated via known technologies such as affinity maturation. Where murine anti-CD3 antibodies are used as a

starting material, humanization of murine anti-CD3 antibodies is desired for the clinical setting, where the mouse-specific residues may induce a human-anti-mouse antigen (HAMA) response in subjects who receive treatment of a trispecific antigen-binding protein described herein. Humanization is accomplished by grafting CDR regions from murine anti-CD3 antibody onto appropriate human germline acceptor frameworks, optionally including other modifications to CDR and/or framework regions.

[0806] Human or humanized anti-CD3 antibodies are therefore used to generate scFv sequences for CD3 binding domains of a trispecific antigen-binding protein. DNA sequences coding for human or humanized VL and VH domains are obtained, and the codons for the constructs are, optionally, optimized for expression in cells from *Homo sapiens*. The order in which the VL and VH domains appear in the scFv is varied (*i.e.* VL-VH, or VH-VL orientation), and three copies of the "G4S" or "G4S" subunit (G4S)₃ connect the variable domains to create the scFv domain. Anti-CD3 scFv plasmid constructs can have optional Flag, His or other affinity tags, and are electroporated into HEK293 or other suitable human or mammalian cell lines and purified. Validation assays include binding analysis by FACS, kinetic analysis using Proteon, and staining of CD3-expressing cells.

Generation of a scFv Glypican-3 (GPC3) binding domain

[0807] Glypican-3 (GPC3) is one of the cell surface proteins present on Hepatocellular Carcinoma but not on healthy normal liver tissue. It is frequently observed to be elevated in hepatocellular carcinoma and is associated with poor prognosis for HCC patients. It is known to activate Wnt signalling. GPC3 antibodies have been generated including MDX-1414, HN3, GC33, and YP7.

[0808] A scFv binding to GPC-3 or another target antigen is generated similarly to the above method for generation of a scFv binding domain to CD3.

Expression of trispecific antigen-binding proteins in vitro

[0809] A CHO cell expression system (Flp-In®, Life Technologies), a derivative of CHO-K1 Chinese Hamster ovary cells (ATCC, CCL-61) (Kao and Puck, Proc. Natl. Acad Sci USA 1968; 60(4):1275-81), is used. Adherent cells are subcultured according to standard cell culture protocols provided by Life Technologies.

[0810] For adaption to growth in suspension, cells are detached from tissue culture flasks and placed in serum-free medium. Suspension-adapted cells are cryopreserved in medium with 10% DMSO.

[0811] Recombinant CHO cell lines stably expressing secreted trispecific antigen-binding proteins are generated by transfection of suspension-adapted cells. During selection with the antibiotic Hygromycin B viable cell densities are measured twice a week, and cells are centrifuged and resuspended in fresh selection medium at a maximal density of 0.1×10^6 viable cells/mL. Cell pools stably expressing trispecific antigen-binding proteins are recovered after 2-3 weeks of selection at which point cells are transferred to standard culture medium in shake flasks. Expression of recombinant secreted proteins is confirmed by performing protein gel electrophoresis or flow cytometry. Stable cell pools are cryopreserved in DMSO containing medium.

[0812] Trispecific antigen-binding proteins are produced in 10-day fed-batch cultures of stably transfected CHO cell lines by secretion into the cell culture supernatant. Cell culture supernatants are harvested after 10 days at culture viabilities of typically >75%. Samples are collected from the production cultures every other day and cell density and viability are assessed. On day of harvest, cell culture supernatants are cleared by centrifugation and vacuum filtration before further use.

[0813] Protein expression titers and product integrity in cell culture supernatants are analyzed by SDS-PAGE.

Purification of trispecific antigen-binding proteins

[0814] Trispecific antigen-binding proteins are purified from CHO cell culture supernatants in a two-step procedure. The constructs are subjected to affinity chromatography in a first step followed by preparative size exclusion chromatography (SEC) on Superdex 200 in a second step. Samples are buffer-exchanged and concentrated by ultrafiltration to a typical concentration of >1 mg/mL. Purity and homogeneity (typically >90%) of final samples are assessed by SDS PAGE under reducing and non-reducing conditions, followed by immunoblotting using an anti-(half-life extension domain) or anti idiotype antibody as well as by analytical SEC, respectively. Purified proteins are stored at aliquots at -80 °C until use.

EXAMPLE 34

Expression of engineered circular RNA with a half-life extension domain has improved pharmacokinetic parameters than without a half-life extension domain

[0815] The trispecific antigen-binding protein encoded on a circRNA molecule of example 23 is administered to cynomolgus monkeys as a 0.5 mg/kg bolus injection

intramuscularly. Another cynomolgus monkey group receives a comparable protein encoded on a circRNA molecule in size with binding domains to CD3 and GPC-3, but lacking a half-life extension domain. A third and fourth group receive a protein encoded on a circRNA molecule with CD3 and half-life extension domain binding domains and a protein with GPC-3 and half-life extension domains, respectively. Both proteins encoded by circRNA are comparable in size to the trispecific antigen-binding protein. Each test group consists of 5 monkeys. Serum samples are taken at indicated time points, serially diluted, and the concentration of the proteins is determined using a binding ELISA to CD3 and/or GPC-3.

[0816] Pharmacokinetic analysis is performed using the test article plasma concentrations. Group mean plasma data for each test article conforms to a multi-exponential profile when plotted against the time post-dosing. The data are fit by a standard two-compartment model with bolus input and first-order rate constants for distribution and elimination phases. The general equation for the best fit of the data for i.v. administration is: $c(t) = Ae^{-at} + Be^{-pt}$, where $c(t)$ is the plasma concentration at time t , A and B are intercepts on the Y-axis, and a and β are the apparent first-order rate constants for the distribution and elimination phases, respectively. The a -phase is the initial phase of the clearance and reflects distribution of the protein into all extracellular fluid of the animal, whereas the second or β -phase portion of the decay curve represents true plasma clearance. Methods for fitting such equations are well known in the art. For example, $A = D/V(a - k_{21})/(a - p)$, $B = D/V(p - k_{21})/(a - p)$, and a and β (for $\alpha > \beta$) are roots of the quadratic equation: $r^2 + (k_{12} + k_{21} + k_{10})r + k_{21}k_{10} = 0$ using estimated parameters of V =volume of distribution, k_{10} =elimination rate, k_{12} =transfer rate from compartment 1 to compartment 2 and k_{21} =transfer rate from compartment 2 to compartment 1, and D =the administered dose.

[0817] Data analysis: Graphs of concentration versus time profiles are made using KaleidaGraph (KaleidaGraph™ V. 3.09 Copyright 1986-1997. Synergy Software. Reading, Pa.). Values reported as less than reportable (LTR) are not included in the PK analysis and are not represented graphically. Pharmacokinetic parameters are determined by compartmental analysis using WinNonlin software (WinNonlin® Professional V. 3.1 WinNonlin™ Copyright 1998-1999. Pharsight Corporation. Mountain View, Calif). Pharmacokinetic parameters are computed as described in Ritschel W A and Kearns G L, 1999, EST: Handbook Of Basic Pharmacokinetics Including Clinical Applications, 5th edition, American Pharmaceutical Assoc., Washington, D C.

[0818] It is expected that the trispecific antigen-binding protein encoded on a circRNA molecule of Example 23 has improved pharmacokinetic parameters such as an

increase in elimination half-time as compared to proteins lacking a half-life extension domain.

EXAMPLE 35

Cytotoxicity of the Trispecific Antigen-Binding Protein

[0819] The trispecific antigen-binding protein encoded on a circRNA molecule of Example 23 is evaluated in vitro on its mediation of T cell dependent cytotoxicity to GPC-3+ target cells.

[0820] Fluorescence labeled GPC3 target cells are incubated with isolated PBMC of random donors or T-cells as effector cells in the presence of the trispecific antigen-binding protein of Example 23. After incubation for 4 h at 37 °C. in a humidified incubator, the release of the fluorescent dye from the target cells into the supernatant is determined in a spectrofluorimeter. Target cells incubated without the trispecific antigen-binding protein of Example 23 and target cells totally lysed by the addition of saponin at the end of the incubation serve as negative and positive controls, respectively.

[0821] Based on the measured remaining living target cells, the percentage of specific cell lysis is calculated according to the following formula: $[1 - (\text{number of living targets}(\text{sample}) / \text{number of living targets}(\text{spontaneous}))] \times 100\%$. Sigmoidal dose response curves and EC50 values are calculated by non-linear regression/4-parameter logistic fit using the GraphPad Software. The lysis values obtained for a given antibody concentration are used to calculate sigmoidal dose-response curves by 4 parameter logistic fit analysis using the Prism software.

EXAMPLE 36

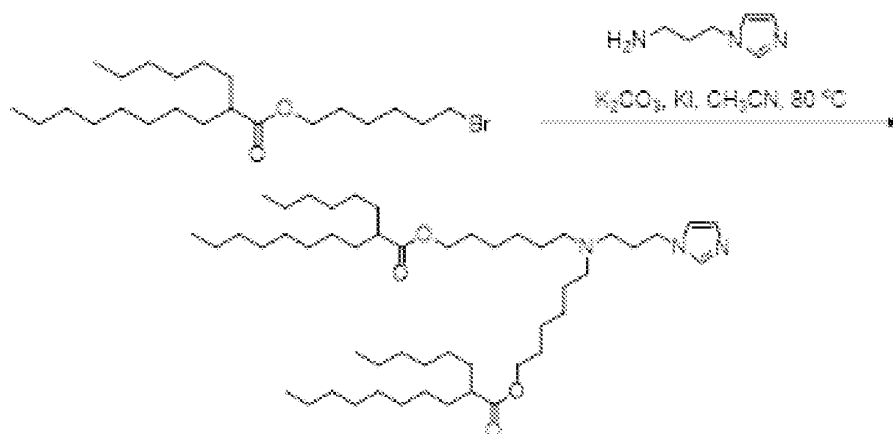
Synthesis of Ionizable Lipids

38.1 Synthesis of ((3-(2-methyl-1H-imidazol-1-yl)propyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate) (Lipid 27, Table 10a) and ((3-(1H-imidazol-1-yl)propyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate) (Lipid 26, Table 10a)

[0822] In a 100 mL round bottom flask connected with condenser, 3-(1H-imidazol-1-yl)propan-1-amine (100 mg, 0.799mmol) or 3-(2-methyl-1H-imidazol-1-yl)propan-1-amine (0.799mmol), 6-bromohexyl 2-hexyldecanoate (737.2 mg, 1.757 mmol), potassium carbonate (485 mg, 3.515 mmol) and potassium iodide (13 mg, 0.08 mmol) were mixed in acetonitrile (30 mL), and the reaction mixture was heated to 80 °C for 48 h. The mixture was

cooled to room temperature and was filtered through a pad of Celite. The filtrate was diluted with ethyl acetate. After washing with water, brine and dried over anhydrous sodium sulfate. The solvent was evaporated and the crude residue was purified by flash chromatography (SiO_2 : CH_2Cl_2 = 100% to 10% of methanol in CH_2Cl_2) and colorless oil product was obtained (92 mg, 15%). Molecular formula of ((3-(1*H*-imidazol-1-yl)propyl)azanediyl)bis(hexane-6,1-diyl) bis(2-hexyldecanoate)) is $\text{C}_{50}\text{H}_{95}\text{N}_3\text{O}_4$ and molecular weight (M_w) is 801.7.

[0823] Reaction scheme for synthesis of ((3-(1*H*-imidazol-1-yl)propyl)azanediyl)bis(hexane-6,1-diyl) bis(2-hexyldecanoate)) (Lipid 26, Table 10a).



[0824] Characterization of Lipid 26 was performed by LC-MS. FIG. 27A-C shows characterization of Lipid 26. FIG. 27A shows the proton NMR observed for Lipid 26. FIG. 27B is a representative LC/MS trace for Lipid 26 with total ion and UV chromatograms shown.

38.2 Synthesis of Lipid 22-S14

38.2.1 Synthesis of 2-(tetradecylthio)ethan-1-ol

[0825] To a mixture of 2-sulfanylethanol (5.40 g, 69.11 mmol, 4.82 mL, 0.871 *eq*) in acetonitrile (200 mL) was added 1-Bromotetradecane (22 g, 79.34 mmol, 23.66 mL, 1 *eq*) and potassium carbonate (17.55 g, 126.95 mmol, 1.6 *eq*) at 25°C. The reaction mixture was warmed to 40°C and stirred for 12 hr. TLC (ethyl acetate/petroleum ether = 25/1, R_f = 0.3, stained by I_2) showed the starting material was consumed completely and a new main spot was generated. The reaction mixture was filtered and the filter cake was washed with acetonitrile (50 mL) and then the filtrate was concentrated under vacuum to get a residue which was purified by column on silica gel (ethyl acetate/petroleum ether = 1/100 to 1/25) to

afford 2-(tetradecylthio)ethan-1-ol (14 g, yield 64.28%) as a white solid.

[0826] ¹H NMR (ET36387-45-P1A, 400 MHz, CHLOROFORM-*d*) δ 0.87 - 0.91 (m, 3 H) 1.27 (s, 20 H) 1.35 - 1.43 (m, 2 H) 1.53 - 1.64 (m, 2 H) 2.16 (br s, 1 H) 2.49 - 2.56 (m, 2 H) 2.74 (t, *J* = 5.93 Hz, 2 H) 3.72 (br d, *J* = 4.89 Hz, 2 H). FIG. 28 shows corresponding Nuclear Magnetic Resonance (NMR) spectrum.

38.2.2 Synthesis of 2-(tetradecylthio)ethyl acrylate

[0827] To a solution of 2-(tetradecylthio)ethan-1-ol (14 g, 51.00 mmol, 1 eq) in dichloromethane (240 mL) was added triethylamine (7.74 g, 76.50 mmol, 10.65 mL, 1.5 eq) and prop-2-enoyl chloride (5.54 g, 61.20 mmol, 4.99 mL, 1.2 eq) dropwise at 0°C under nitrogen. The reaction mixture was warmed to 25°C and stirred for 12 hr. TLC (ethyl acetate/petroleum ether = 25/1, *R_f* = 0.5, stained by I₂) showed the starting material was consumed completely and a new main spot was generated. The reaction solution was concentrated under vacuum to get crude which was purified by column on silica gel (ethyl acetate/petroleum ether = 1/100 to 1/25) to afford 2-(tetradecylthio)ethyl acrylate (12 g, yield 71.61%) as a colorless oil.

[0828] ¹H NMR (ET36387-49-P1A, 400 MHz, CHLOROFORM-*d*) δ 0.85 - 0.93 (m, 3 H) 1.26 (s, 19 H) 1.35 - 1.43 (m, 2 H) 1.53 - 1.65 (m, 2 H) 2.53 - 2.62 (m, 2 H) 2.79 (t, *J* = 7.03 Hz, 2 H) 4.32 (t, *J* = 7.03 Hz, 2 H) 5.86 (dd, *J* = 10.39, 1.47 Hz, 1 H) 6.09 - 6.19 (m, 1 H) 6.43 (dd, *J* = 17.30, 1.41 Hz, 1 H). FIG. 29 shows corresponding Nuclear Magnetic Resonance (NMR) spectrum.

38.2.3 Synthesis of bis(2-(tetradecylthio)ethyl) 3,3'-((3-(2-methyl-1H-imidazol-1-yl)propyl)azanediyl)dipropionate (Lipid 22-S14)

[0829] A flask was charged with 3-(2-methyl-1H-imidazol-1-yl)propan-1-amine (300 mg, 2.16 mmol) and 2-(tetradecylthio)ethyl acrylate (1.70 g, 5.17 mmol). The neat reaction mixture was heated to 80°C and stirred for 48 hr. TLC (ethyl acetate, *R_f* = 0.3, stained by I₂, one drop ammonium hydroxide added) showed the starting material was consumed completely and a new main spot was formed. The reaction mixture was diluted with dichloromethane (4 mL) and purified by column on silica gel (petroleum ether/ethyl acetate = 3/1 to 0/1, 0.1% ammonium hydroxide added) to get bis(2-(tetradecylthio)ethyl) 3,3'-((3-(2-methyl-1H-imidazol-1-yl)propyl)azanediyl)dipropionate (501 mg, yield 29.1%) as colorless oil.

[0830] ¹H NMR (ET36387-51-P1A, 400 MHz, CHLOROFORM-*d*) δ 0.87 (t, *J* = 6.73 Hz, 6 H) 1.25 (s, 40 H) 1.33 - 1.40 (m, 4 H) 1.52 - 1.61 (m, 4 H) 1.81 - 1.90 (m, 2 H) 2.36 (s, 3 H) 2.39 - 2.46 (m, 6 H) 2.53 (t, *J* = 7.39 Hz, 4 H) 2.70 - 2.78 (m, 8 H) 3.84 (t, *J* = 7.17 Hz,

2 H) 4.21 (t, $J = 6.95$ Hz, 4 H) 6.85 (s, 1 H) 6.89 (s, 1 H). FIG. 30 shows corresponding Nuclear Magnetic Resonance (NMR) spectrum.

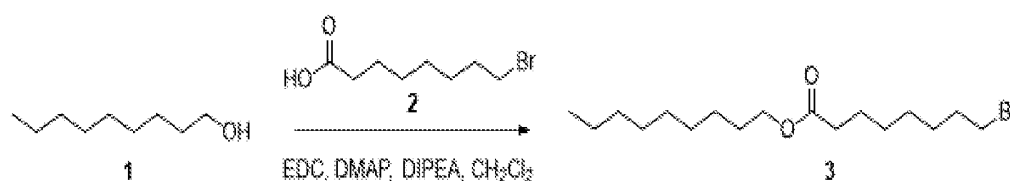
38.3 Synthesis of bis(2-(tetradecylthio)ethyl) 3,3'-((3-(1H-imidazol-1-yl)propyl)azanediyl)dipropionate (Lipid 93-S14)

[0831] A flask was charged with 3-(1H-imidazol-1-yl)propan-1-amine (300 mg, 2.40 mmol, 1 *eq*) and 2-(tetradecylthio)ethyl acrylate (1.89 g, 5.75 mmol, 2.4 *eq*). The neat reaction mixture was heated to 80°C and stirred for 48 hr. TLC (ethyl acetate, $R_f = 0.3$, stained by I_2 , one drop ammonium hydroxide added) showed the starting material was consumed completely and a new main spot was formed. The reaction mixture was diluted with dichloromethane (4 mL) and purified by column on silica gel (petroleum ether/ethyl acetate = 1/20 - 0/100, 0.1% ammonium hydroxide added) to get bis(2-(tetradecylthio)ethyl) 3,3'-((3-(1H-imidazol-1-yl)propyl)azanediyl)dipropionate (512 mg, yield 27.22%) as colorless oil.

[0832] ^1H NMR (ET36387-54-P1A, 400 MHz, CHLOROFORM-*d*) δ 0.89 (t, $J = 6.84$ Hz, 6 H) 1.26 (s, 40 H) 1.34 - 1.41 (m, 4 H) 1.58 (br t, $J = 7.50$ Hz, 4 H) 1.92 (t, $J = 6.62$ Hz, 2 H) 2.36 - 2.46 (m, 6 H) 2.55 (t, $J = 7.50$ Hz, 4 H) 2.75 (q, $J = 6.84$ Hz, 8 H) 3.97 (t, $J = 6.95$ Hz, 2 H) 4.23 (t, $J = 6.95$ Hz, 4 H) 6.95 (s, 1 H) 7.06 (s, 1 H) 7.51 (s, 1 H). FIG. 31 shows corresponding Nuclear Magnetic Resonance (NMR) spectrum.

38.4 Synthesis of heptadecan-9-yl 8-((3-(2-methyl-1H-imidazol-1-yl)propyl)(8-(nonyloxy)-8-oxooctyl)amino)octanoate (Lipid 54, Table 10a)

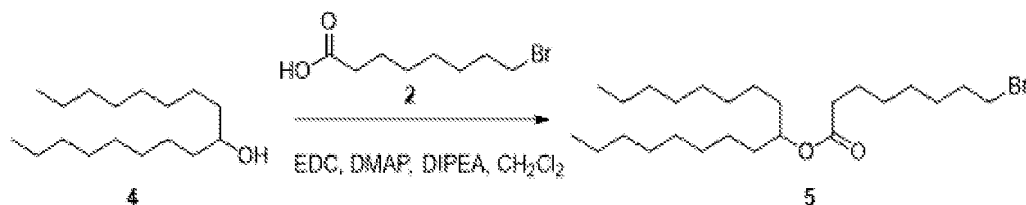
38.4.1 Synthesis of nonyl 8-bromooctanoate (3)



[0833] To a mixture of 8-bromooctanoic acid (2) (18.6 g, 83.18 mmol) and nonan-1-ol (1) (10 g, 69.32 mmol) in CH_2Cl_2 (500 mL) was added DMAP (1.7 g, 13.86 mmol), DIPEA (48 mL, 277.3 mmol) and EDC (16 g, 83.18 mmol). The reaction was stirred at room temperature overnight. After concentration of the reaction mixture, the crude residue was dissolved in ethyl acetate (500 mL), washed with 1N HCl, sat. NaHCO_3 , water and Brine.

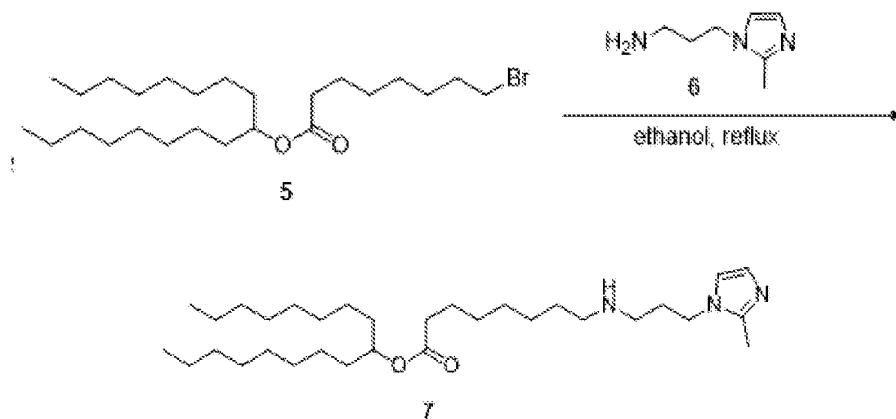
The organic layer was dried over anhydrous Na_2SO_4 . The solvent was evaporated and the crude residue was purified by flash chromatography (SiO_2 : Hexane = 100% to 30% of EtOAc in Hexane) and colorless oil product **3** was obtained (9 g, 37%).

38.4.2 Synthesis of heptadecan-9-yl 8-bromooctanoate (**5**)



[0834] To a mixture of 8-bromooctanoic acid (**2**) (10 g, 44.82 mmol) and heptadecan-9-ol (**4**) (9.6 g, 37.35 mmol) in CH_2Cl_2 (300 mL) was added DMAP (900 mg, 7.48 mmol), DIPEA (26 mL, 149.7 mmol) and EDC (10.7 g, 56.03 mmol). The reaction was stirred at room temperature overnight. After concentration of the reaction mixture, the crude residue was dissolved in ethyl acetate (300 mL), washed with 1N HCl, sat. NaHCO_3 , water and Brine. The organic layer was dried over anhydrous Na_2SO_4 . The solvent was evaporated and the crude residue was purified by flash chromatography (SiO_2 : Hexane = 100% to 30% of EtOAc in Hexane) and colorless oil product **5** was obtained (5 g, 29%).

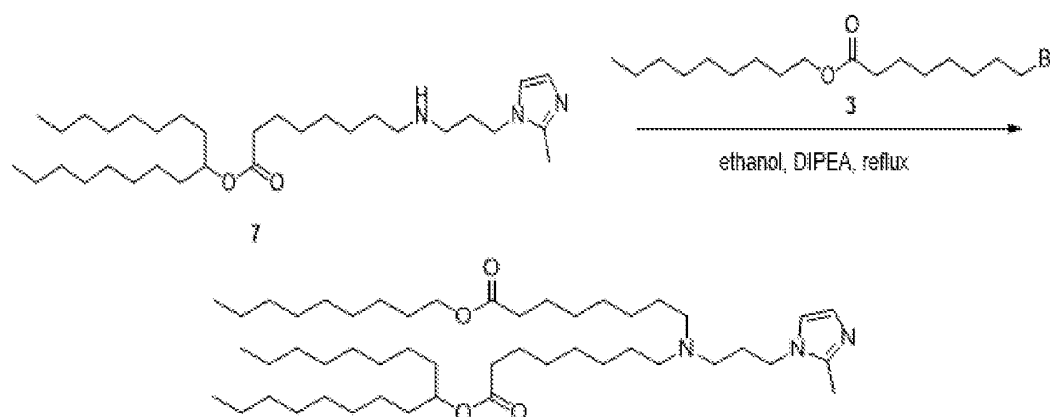
38.4.3 Synthesis of heptadecan-9-yl 8-((3-(2-methyl-1H-imidazol-1-yl)propyl)amino)octanoate (**7**)



[0835] In a 100 mL round bottom flask connected with condenser, heptadecan-9-yl 8-bromooctanoate (**5**) (860 mg, 1.868 mmol) and 3-(2-methyl-1H-imidazol-1-yl)propan-1-amine (**6**) (1.3 g, 9.339 mmol) were mixed in ethanol (10 mL). The reaction mixture was heated to reflux overnight. MS (APCI) showed the expected product. The mixture was cooled

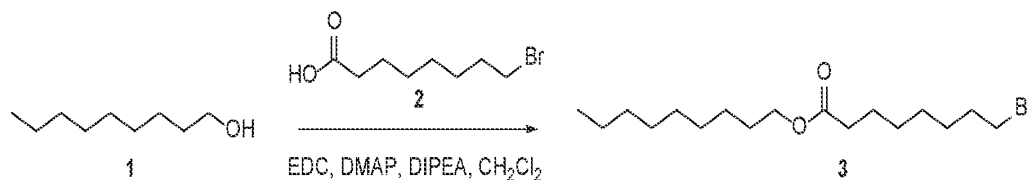
to room temperature and concentrated. The crude residue was purified by flash chromatography (SiO₂: CH₂Cl₂= 100% to 10% of methanol+1%NH₄OH in CH₂Cl₂) and colorless oil product **7** was obtained (665 mg, 69%).

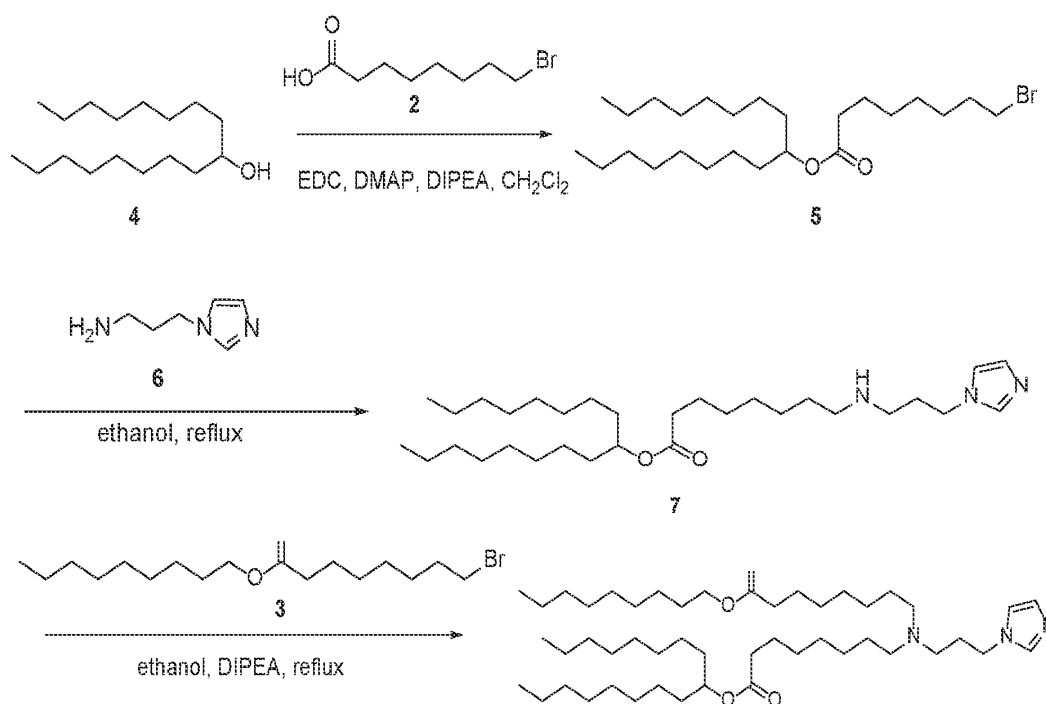
38.4.4 Synthesis of heptadecan-9-yl 8-((3-(2-methyl-1H-imidazol-1-yl)propyl)(8-(nonyloxy)-8-oxooctyl)amino)octanoate (Lipid 54, Table 10a)



[0836] In a 100 mL round bottom flask connected with condenser, heptadecan-9-yl 8-((3-(2-methyl-1H-imidazol-1-yl)propyl)amino)octanoate (**7**) (665 mg, 1.279 mmol) and nonyl 8-bromooctanoate (**3**) (536 mg, 1.535 mmol) were mixed in ethanol (10 mL), then DIPEA (0.55 mL, 3.198 mmol) was added. The reaction mixture was heated to reflux overnight. Both MS (APCI) and TLC (10%MeOH+1%NH₄OH in CH₂Cl₂) showed the product and some unreacted starting material. The mixture was cooled to room temperature and concentrated. The crude residue was purified by flash chromatography (SiO₂: CH₂Cl₂= 100% to 10% of methanol+1%NH₄OH in CH₂Cl₂) and colorless oil was obtained (170 mg, 17%).

38.5 Synthesis of heptadecan-9-yl 8-((3-(1H-imidazol-1-yl)propyl)(8-(nonyloxy)-8-oxooctyl)amino)octanoate (Lipid 53, Table 10a)





[0837] Lipid 53 from Table 10a is synthesized according to the scheme above. Reaction conditions are identical to Lipid 54 with the exception of 3-(1H-imidazol-1-yl)propan-1-amine as the imidazole amine.

EXAMPLE 37

Lipid nanoparticle formulation with circular RNA

[0838] Lipid Nanoparticles (LNPs) were formed using a Precision Nanosystems Ignite instrument with a 'NextGen' mixing chamber. Ethanol phase contained ionizable Lipid 26 from Table 10a, DSPC, Cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio was combined with an aqueous phase containing circular RNA and 25 mM sodium acetate buffer at pH 5.2. A 3:1 aqueous to ethanol mixing ratio was used. The formulated LNP then were dialyzed in 1L of water and exchanged 2 times over 18 hours. Dialyzed LNPs were filtered using 0.2 μm filter. Prior to *in vivo* dosing, LNPs were diluted in PBS. LNP sizes were determined by dynamic light scattering. A cuvette with 1 mL of 20 $\mu\text{g/mL}$ LNPs in PBS (pH 7.4) was measured for Z-average using the Malvern Panalytical Zetasizer Pro. The Z-average and polydispersity index

were recorded.

39.1 Formulation of Lipids 26 and 27 from Table 10a

[0839] Lipid Nanoparticles (LNPs) were formed using a Precision Nanosystems Ignite instrument with a 'NextGen' mixing chamber. Ethanol phase contained ionizable Lipid 26 or Lipid 27 from Table 10a, DOPE, Cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a weight ratio of 16:1:4:1 or 62:4:33:1 molar ratio was combined with an aqueous phase containing circular RNA and 25 mM sodium acetate buffer at pH 5.2. A 3:1 aqueous to ethanol mixing ratio was used. The formulated LNPs were then dialyzed in 1L of water and exchanged 2 times over 18 hours. Dialyzed LNPs were filtered using 0.2 µm filter. Prior to *in vivo* dosing, LNPs were diluted in PBS. LNP sizes were determined by dynamic light scattering. A cuvette with 1 mL of 20 µg/mL LNPs in PBS (pH 7.4) was measured for Z-average using the Malvern Panalytical Zetasizer Pro. The Z-average and polydispersity index were recorded.

39.2 Formulation of Lipids 53 and 54 from Table 10a

[0840] Lipid Nanoparticles (LNPs) were formed using a Precision Nanosystems Ignite instrument with a 'NextGen' mixing chamber. Ethanol phase contained ionizable Lipid 53 or 54 of Table 10a, DOPE, Cholesterol, and DSPE-PEG 2000 (Avanti Polar Lipids Inc.) at a molar ratio of 50:10:38.5:1.5 was combined with an aqueous phase containing circular RNA and 25 mM sodium acetate buffer at pH 5.2. A 3:1 aqueous to ethanol mixing ratio was used. The formulated LNPs were then dialyzed in 1L of 1x PBS and exchanged 2 times over 18 hours. Dialyzed LNPs were filtered using 0.2 µm filter. Prior to *in vivo* dosing, LNPs were diluted in PBS. LNP sizes were determined by dynamic light scattering. A cuvette with 1 mL of 20 µg/mL LNPs in PBS (pH 7.4) was measured for Z-average using the Malvern Panalytical Zetasizer Pro. The Z-average and polydispersity index were recorded.

[0841] LNP zeta potential was measured using the Malvern Panalytical Zetasizer Pro. A mixture containing 200 µL of the particle solution in water and 800 µL of distilled RNase-free water with a final particle concentration of 400 µg/mL was loaded into a zetasizer capillary cell for analysis.

[0842] RNA encapsulation was determined using a Ribogreen assay. Nanoparticle solutions were diluted in tris-ethylenediaminetetraacetic acid (TE) buffer at a theoretical oRNA concentration of 2 µg/mL. Standard oRNA solutions diluted in TE buffer were made

ranging from 2 µg/mL to 0.125 µg/mL. The particles and standards were added to all wells and a second incubation was performed (37 °C at 350 rpm for 3 minutes). Fluorescence was measured using a SPECTRAmax® GEMINI XS microplate spectrofluorometer. The concentration of circular RNA in each particle solution was calculated using the standard curve. The encapsulation efficiency was calculated from the ratio of oRNA detected between lysed and unlysed particles.

Table 36a. Characterization of LNPs

Ionizable Lipid	Size (nm)	PDI	Encapsulation Efficiency (%)	Zeta Data Potential (mV)
22-S14	88	0.09	96	3.968
93-S14	119	0.02	96	-6.071
Lipid 26, Table 10a	86	0.08	92	-15.24

Table 36b. Characterization of LNPs

Ionizable Lipid	Z-Average(nm)	PDI	RNA Entrapment(%)
22-S14	64	0.05	97
93-S14	74	0.04	95
Lipid 26, Table 10a	84	0.04	96

EXAMPLE 38

In Vivo Analysis

[0843] Female CD-1 or female c57BL/6J mice ranging from 22 – 25 g were dosed at 0.5 mg/kg RNA intravenously. Six hours after injection, mice were injected intraperitoneally with 200 µL of D-luciferin at 15 mg/mL concentration. 5 minutes after injection, mice were anesthetized using isoflurane, and placed inside the IVIS Spectrum In Vivo Imaging System (Perkin Elmer) with dorsal side up. Whole body total IVIS flux of Lipids 22-S14, 93-S14, Lipid 26 (Table 10a) is presented in FIG. 32A. Post 10 minutes injection, mice were scanned for luminescence. Mice were euthanized and organs were extracted within 25 minutes of luciferin injection to scan for luminescence in liver, spleen, kidneys, lungs, and heart. Images (FIGs. 33A-B, 34A-B, 35A-B) were analyzed using Living Images (Perkin Elmer) software.

Regions of interest were drawn to obtain flux and average radiance and analyzed for biodistribution of protein expression (FIG. 32A-B).

[0844] FIG. 32A illustrates the increased whole-body total flux observed from luciferase oRNA with Lipid 26 (Table 10a) LNPs compared to LNPs made with lipids 22-S14 and 93-S14. FIG. 32B shows the *ex vivo* IVIS analysis of tissues further highlighting the increased overall expression with Lipid 26 (Table 10a) while maintaining the desired spleen to liver ratios observed with lipids 22-S14 and 93-S14 despite the significant structural changes designed to improve expression. These data highlight the improvements afforded by Lipid 26 (Table 10a) compared to previously reported lipids.

[0845] Similar analysis as described above was also performed with oRNA encapsulated in LNPs formed with Lipid 15 from Table 10b or Lipid 53 or 54 from Table 10a. FIGs. 36A-C show the *ex vivo* IVIS analysis of tissues, respectively highlighting the overall expression with Lipid 15, 53, and 54 while maintaining the desired spleen to liver ratios despite the significant structural changes designed to improve expression. FIG. 36D shows the results for PBS control. These data demonstrates the improvements afforded by Lipids 15, 53, and 54 from Table 10a compared to previously reported lipids such as 93-S14 and 22-S14.

EXAMPLE 39

Delivery of Luciferase

[0846] Human peripheral blood mononuclear cells (PBMCs) (Stemcell Technologies) were transfected with lipid nanoparticles (LNP) encapsulating firefly luciferase (*fluc*) circular RNA and examined for luciferase expression. PBMCs from two different donors were incubated *in vitro* with five different LNP compositions, containing circular RNA encoding for firefly luciferase (200 ng), at 37°C in RPMI, 2% human serum, IL-2 (10 ng/mL), and 50 uM BME. PBMCs incubated without LNP were used as a negative control. After 24 hours, the cells were lysed and analyzed for firefly luciferase expression based on bioluminescence (Promega BrightGlo).

[0847] Representative data are presented in FIGs. 37A and 37B, showing that that the tested LNPs are capable of delivering circular RNA into primary human immune cells resulting in protein expression.

EXAMPLE 40

In Vitro Delivery of Green Fluorescent Protein (GFP) or Chimeric Antigen Receptor (CAR)

[0848] Human PBMCs (Stemcell Technologies) were transfected with LNP encapsulating GFP and examined by flow cytometry. PBMCs from five different donors (PBMC A-E) were incubated *in vitro* with one LNP composition, containing circular RNA encoding either GFP or CD19-CAR (200 ng), at 37°C in RPMI, 2% human serum, IL-2 (10 ng/mL), and 50 uM BME. PBMCs incubated without LNP were used as a negative control. After 24, 48, or 72 hours post-LNP incubation, cells were analyzed for CD3, CD19, CD56, CD14, CD11b, CD45, fixable live dead, and payload (GFP or CD19-CAR).

[0849] Representative data are presented in FIGs. 38A and 38B, showing that the tested LNP is capable of delivering circular RNA into primary human immune cells resulting in protein expression.

EXAMPLE 41*Multiple IRES variants can mediate expression of murine CD19 CAR in vitro*

[0850] Multiple circular RNA constructs, encoding anti-murine CD19 CAR, contains unique IRES sequences and were lipotransfected into 1C1C7 cell lines. Prior to lipotransfection, 1C1C7 cells are expanded for several days in complete RPMI. Once the cells expanded to appropriate numbers, 1C1C7 cells were lipotransfected (Invitrogen RNAiMAX) with four different circular RNA constructs. After 24 hours, 1C1C7 cells were incubated with His-tagged recombinant murine CD19 (Sino Biological) protein, then stained with a secondary anti-His antibody. Afterwards, the cells were analyzed via flow cytometry.

[0851] Representative data are presented in FIGs. 39, showing that IRES sourced from the indicated virus (apodemus agrarius picornavirus, caprine kobuvirus, parbovirus, and salivirus) are capable of driving expression of an anti-mouse CD19 CAR in murine T cells.

EXAMPLE 42*Murine CD19 CAR mediates tumor cell killing in vitro*

[0852] Circular RNA encoding anti-mouse CD19 CAR were electroporated into murine T cells to evaluate CAR-mediated cytotoxicity. For electroporation, T cells were electroporated with circular RNA encoding anti-mouse CD19 CAR using ThermoFisher's Neon Transfection System then rested overnight. For the cytotoxicity assay, electroporated T cells were co-cultured with Fluc⁺ target and non-target cells at 1:1 ratio in complete RPMI containing 10%

FBS, IL-2 (10 ng/mL), and 50 uM BME and incubated overnight at 37°C. Cytotoxicity was measured using a luciferase assay system 24 hours post-co-culture (Promega Brightglo Luciferase System) to detect lysis of Fluc+ target and non-target cells. Values shown are calculated relative to the untransfected mock signal.

[0853] Representative data are presented in FIG. 40, showing that an anti-mouse CD19 CAR expressed from circular RNA is functional in murine T cells *in vitro*.

EXAMPLE 43

Functional depletion of B cells with a lipid encapsulated circular RNA encoding murine CD19 CAR

[0854] C57BL/6J mice were injected with LNP formed with Lipid 15 in Table 10b, encapsulating circular RNA encoding anti-murine CD19 CAR. As a control, Lipid 15 in Table 10b encapsulating circular RNA encoding firefly luciferase (*fLuc*) were injected in different group of mice. Female C57BL/6J, ranging from 20-25 g, were injected intravenously with 5 doses of 0.5 mg/kg of LNP, every other day. Between injections, blood draws were analyzed via flow cytometry for fixable live/dead, CD45, TCRvb, B220, CD11b, and anti-murine CAR. Two days after the last injection, spleens were harvested and processed for flow cytometry analysis. Splenocytes were stained with fixable live/dead, CD45, TCRvb, B220, CD11b, NK1.1, F4/80, CD11c, and anti-murine CAR. Data from mice injected with anti-murine CD19 CAR LNP were normalized to mice that received *fLuc* LNP.

[0855] Representative data are presented in FIGs. 41A, 41B, and 41C, showing that an anti-mouse CD19 CAR expressed from circular oRNA delivered *in vivo* with LNPs is functional in murine T cells *in vivo*.

EXAMPLE 44

CD19 CAR expressed from circular RNA has higher yield and greater cytotoxic effect compared to that expressed from mRNA

[0856] Circular RNA encoding anti-CD19 chimeric antigen receptor, which includes, from N-terminus to C-terminus, a FMC63-derived scFv, a CD8 transmembrane domain, a 4-1BB costimulatory domain, and a CD3ζ intracellular domain, were electroporated into human peripheral T cells to evaluate surface expression and CAR-mediated cytotoxicity. For comparison, circular RNA-electroporated T cells were compared to mRNA-electroporated T cells in this experiment. For electroporation, CD3+ T cells were isolated from

human PBMCs using commercially available T cell isolation kits (Miltenyi Biotec) from donor human PBMCs. After isolation, T cells were stimulated with anti-CD3/anti-CD28 (Stemcell Technologies) and expanded over 5 days at 37°C in complete RPMI containing 10% FBS, IL-2 (10 ng/mL), and 50 uM BME. Five days post stimulation, T cells were electroporated with circular RNA encoding anti-human CD19 CAR using ThermoFisher's Neon Transfection System and then rested overnight. For the cytotoxicity assay, electroporated T cells were co-cultured with Fluc+ target and non-target cells at 1:1 ratio in complete RPMI containing 10% FBS, IL-2 (10 ng/mL), and 50 uM BME and incubated overnight at 37°C. Cytotoxicity was measured using a luciferase assay system 24 hours post-co-culture (Promega Brightglo Luciferase System) to detect lysis of Fluc+ target and non-target cells. Furthermore, an aliquot of electroporated T cells were taken and stained for live dead fixable staining, CD3, CD45, and chimeric antigen receptors (FMC63) at the day of analysis.

[0857] Representative data are presented in FIGs. 42 and 43. FIGs. 42A and 42B show that an anti-human CD19 CAR expressed from circular RNA is expressed at higher levels and longer than an anti-human CD19 CAR expressed from linear mRNA. FIGs. 43A and 43B show that an anti-human CD19 CAR expressed from circular RNA is exerts a greater cytotoxic effect relativea to anti-human CD19 CAR expressed from linear mRNA.

EXAMPLE 45

Functional Expression of Two CARs from a Single Circular RNA

[0858] Circular RNA encoding chimeric antigen receptors were electroporated into human peripheral T cells to evaluate surface expression and CAR-mediated cytotoxicity. The purpose of this study is to evaluate if circular RNA encoding for two CARs can be stochastically expressed with a 2A (P2A) or an IRES sequence. For electroporation, CD3+ T cells were commercially purchased (Cellero) and stimulated with anti-CD3/anti-CD28 (Stemcell Technologies) and expanded over 5 days at 37°C in complete RPMI containing 10% FBS, IL-2 (10 ng/mL), and 50 uM BME. Four days post stimulation, T cells were electroporated with circular RNA encoding anti-human CD19 CAR, anti-human CD19 CAR-2A-anti-human BCMA CAR, and anti-human CD19 CAR-IRES-anti-human BCMA CAR using ThermoFisher's Neon Transfection System then rested overnight. For the cytotoxicity assay, electroporated T cells were co-cultured with Fluc+ K562 cells expressing human CD19 or BCMA antigens at 1:1 ratio in complete RPMI containing 10% FBS, IL-2 (10 ng/mL), and 50 uM BME and incubated overnight at 37°C. Cytotoxicity was measured using a luciferase assay system 24 hours post-co-culture (Promega BrightGlo Luciferase System) to detect lysis of

Fluc⁺ target cells.

[0859] Representative data are presented in FIG. 44, showing that two CARs can be functionally expressed from the same circular RNA construct and exert cytotoxic effector function.

EXAMPLE 46

In vivo circular RNA transfection using Cre reporter mice

[0860] Circular RNAs encoding Cre recombinase (Cre) are encapsulated into lipid nanoparticles as previously described. Female, 6-8 week old B6.Cg-Gt(ROSA)26Sortm9(CAG-tdTomato)Hze/J (Ai9) mice were dosed with lipid nanoparticles at 0.5 mg/kg RNA intravenously. Fluorescent tdTomato protein was transcribed and translated in Ai9 mice upon Cre recombination, meaning circular RNAs have been delivered to and translated in tdTomato⁺ cells. After 48 hr, mice were euthanized and the spleens were harvested, processed into a single cell suspension, and stained with various fluorophore-conjugated antibodies for immunophenotyping via flow cytometry.

[0861] FIG. 45A shows representative FACS plots with frequencies of tdTomato expression in various spleen immune cell (CD45⁺, live) subsets, including total myeloid (CD11b⁺), B cells (CD19⁺), and T cells (TCR-B⁺) following treatment with LNPs formed with Lipid 27 or 26 from Table 10a or Lipid 15 from Table 10b. Ai9 mice injected with PBS represented background tdTomato fluorescence. FIG. 45B quantifies the proportion of myeloid cells, B cells, and T cells expressing tdTomato (mean + std. dev., n = 3), which is equivalent to the proportion of each cell population which has been successfully transfected with Cre circular RNA. LNPs made with Lipids 27 and 26 from Table 10a exhibit significantly higher myeloid and T cell transfection compared with Lipid 93-S14, highlighting the improvements conferred by lipid structural modifications.

[0862] FIG. 45C illustrates the proportion of additional splenic immune cell populations expressing tdTomato with Lipids 27 and 26 from Table 10a (mean + std. dev., n = 3), which also include NK cells (NKp46⁺, TCR-B⁻), classical monocytes (CD11b⁺, Ly-6G⁻, Ly-6C_{hi}), nonclassical monocytes (CD11b⁺, Ly-6G⁻, Ly-6C_{lo}), neutrophils (CD11b⁺, Ly-6G⁺), and dendritic cells (CD11c⁺, MHC-II⁺). These experiments demonstrate that LNPs made with Lipids 27 and 26 from Table 10a and Lipid 15 from Table 10b are effective at delivering circular RNAs to many splenic immune cell subsets in mice and lead to successful protein expression from the circular RNA in those cells.

EXAMPLE 47

Example 47A: *Built-in polyA sequences and affinity-purification to produce immune-silent circular RNA*

[0863] PolyA sequences (20-30nt) were inserted into the 5' and 3' ends of the RNA construct (precursor RNA with built-in polyA sequences in the introns). Precursor RNA and introns can alternatively be polyadenylated post-transcriptionally using, *e.g.*, *E. coli.* polyA polymerase or yeast polyA polymerase, which requires the use of an additional enzyme.

[0864] Circular RNA in this example was circularized by *in vitro* transcription (IVT) and affinity-purified by washing over a commercially available oligo-dT resin to selectively remove polyA-tagged sequences (including free introns and precursor RNA) from the splicing reaction. The IVT was performed with a commercial IVT kit (New England Biolabs) or a customized IVT mix (Orna Therapeutics), containing guanosine monophosphate (GMP) and guanosine triphosphate (GTP) at different ratios (GMP:GTP = 8, 12.5, or 13.75). In some embodiments, GMP at a high GMP:GTP ratio may be preferentially included as the first nucleotide, yielding a majority of monophosphate-capped precursor RNAs. As a comparison, the circular RNA product was alternatively purified by the treatment with Xrn1, Rnase R, and Dnase I (enzyme purification).

[0865] Immunogenicity of the circular RNAs prepared using the affinity purification or enzyme purification process were then assessed. Briefly, the prepared circular RNAs were transfected into A549 cells. After 24 hours, the cells were lysed and interferon beta-1 induction relative to mock samples was measured by qPCR. 3p-hpRNA, a triphosphorylated RNA, was used as a positive control.

[0866] FIGs. 46B and 46C show that the negative selection affinity purification removes non-circular products from splicing reactions when polyA sequences are included on elements that are removed during splicing and present in unspliced precursor molecules. FIG. 46D shows circular RNAs prepared with tested IVT conditions and purification methods are all immunoquiescent. These results suggest the negative selection affinity purification is equivalent or superior to enzyme purification for circular RNA purification and that customized circular RNA synthesis conditions (IVT conditions) may reduce the reliance on GMP excess to achieve maximal immunoquiescence.

Example 47B: *Dedicated binding site and affinity-purification for circular RNA production*

[0867] Instead of polyA tags, one can include specifically design sequences (DBS, dedicated binding site).

[0868] Instead of a polyA tag, a dedicated binding site (DBS), such as a specifically designed complementary oligonucleotide that can bind to a resin, may be used to selectively deplete precursor RNA and free introns. In this example, DBS sequences (30nt) were inserted into the 5' and 3' ends of the precursor RNA. RNA was transcribed and the transcribed product was washed over a custom complementary oligonucleotide linked to a resin.

[0869] FIGs. 47B and 47C demonstrates that including the designed DBS sequence in elements that are removed during splicing enables the removal of unspliced precursor RNA and free intron components in a splicing reaction, via negative affinity purification.

Example 47C: Production of a circular RNA encoding dystrophin

[0870] A 12kb12,000nt circular RNA encoding dystrophin was produced by *in vitro* transcription of RNA precursors followed by enzyme purification using a mixture of Xrn1, DNase 1, and RNase R to degrade remaining linear components. FIG. 48 shows that the circular RNA encoding dystrophin was successfully produced.

EXAMPLE 48

5' spacer between 3' intron fragment and the IRES improves circular RNA expression

[0871] Expression level of purified circRNAs with different 5' spacers between the 3' intron fragment and the IRES in Jurkat cells were compared. Briefly, luminescence from secreted *Gaussia* luciferase in supernatant was measured 24 hours after electroporation of 60,000 cells with 250ng of each RNA.

[0872] Additionally, stability of purified circRNAs with different 5' spacers between the 3' intron fragment and the IRES in Jurkat cells were compared. Briefly, luminescence from secreted *Gaussia* luciferase in supernatant was measured over 2 days after electroporation of 60,000 cells with 250ng of each RNA and normalized to day 1 expression.

[0873] The results are shown in FIGs. 49A and 49B, indicating that adding a spacer can enhance IRES function and the importance of sequence identity and length of the added spacer. A potential explanation is that the spacer is added right before the IRES and likely functions by allowing the IRES to fold in isolation from other structured elements such as the intron fragments.

EXAMPLE 49

[0874] This example describes deletion scanning from 5' or 3' end of the caprine kobuvirus IRES. IRES borders are generally poorly characterized and require empirical analysis, and this example can be used for locating the core functional sequences required for driving translation. Briefly, circular RNA constructs were generated with truncated IRES elements operably linked to a gaussia luciferase coding sequence. The truncated IRES elements had nucleotide sequences of the indicated lengths removed from the 5' or 3' end. Luminescence from secreted gaussia luciferase in supernatant was measured 24 and 48 hours after electroporation of primary human T cells with RNA. Stability of expression was calculated as the ratio of the expression level at the 48-hour time point relative to that at the 24-hour time point.

[0875] As shown in FIG. 50, deletion of more than 40 nucleotides from the 5' end of the IRES reduced expression and disrupted IRES function. Stability of expression was relatively unaffected by the truncation of the IRES element but expression level was substantially reduced by deletion of 141 nucleotides from the 3' end of the IRES, whereas deletion of 57 or 122 nucleotides from the 3' end had a positive impact on the expression level.

[0876] It was also observed that deletion of the 6-nucleotide pre-start sequence reduced the expression level of the luciferase reporter. Replacement of the 6-nucleotide sequence with a classical kozak sequence (GCCACC) did not have a significant impact but at least maintained expression.

EXAMPLE 50

[0877] This example describes modifications (*e.g.*, truncations) of selected IRES sequences, including Caprine Kobuvirus (CKV) IRES, Parabovirus IRES, Apodemus Picornavirus (AP) IRES, Kobuvirus SZAL6 IRES, Crohivirus B (CrVB) IRES, CVB3 IRES, and SAFV IRES. The sequences of the IRES elements are provided in SEQ ID NOs: 348-389. Briefly, circular RNA constructs were generated with truncated IRES elements operably linked to a gaussia luciferase coding sequence. HepG2 cells were transfected with the circular RNAs. Luminescence in the supernatant was assessed 24 and 48 hours after transfection. Stability of expression was calculated as the ratio of the expression level at the 48-hour time point relative to that at the 24-hour time point.

[0878] As shown in FIG. 51, truncations had variable effects depending on the identity of the IRES, which may depend on the initiation mechanism and protein factors used for translation, which often differs between IRESs. 5' and 3' deletions can be effectively combined, for example, in the context of CKV IRES. Addition of a canonical Kozak sequence in some cases significantly improved expression (as in SAFV, Full vs Full+K) or diminished expression (as in CKV, 5d40/3d122 vs 5d40/3d122+K).

EXAMPLE 51

[0879] This example describes modifications of CK-739, AP-748, and PV-743 IRES sequences, including mutations alternative translation initiation sites. Briefly, circular RNA constructs were generated with modified IRES elements operably linked to a gaussia luciferase coding sequence. Luminescence from secreted gaussia luciferase in supernatant was measured 24 and 48 hours after transfection of IC1C7 cells with RNA.

[0880] CUG was the most commonly found alternative start site but many others were also characterized. These triplets can be present in the IRES scanning tract prior to the start codon and can affect translation of correct polypeptides. Four alternative start site mutations were created, with the IRES sequences provided in SEQ ID NOs: 378-380. As shown in FIG. 52, mutations of alternative translation initiation sites in the CK-739 IRES affected translation of correct polypeptides, positively in some instances and negatively in other instances. Mutation of all the alternative translation initiation sites reduced the level of translation.

[0881] Alternative Kozak sequences, 6 nucleotides before start codon, can also affect expression levels. The 6-nucleotide sequence upstream of the start codon were gTcacG, aaagtc, gTcacG, gtcattg, gcaaac, and acaacc, respectively, in CK-739 IRES and Sample Nos. 1-5 in the "6nt Pre-Start" group. As shown in FIG. 52, substitution of certain 6-nucleotide sequences prior to the start codon affected translation.

[0882] It was also observed that 5' and 3' terminal deletions in AP-748 and PV-743 IRES sequences reduced expression. However, in the CK-739 IRES, which had a long scanning tract, translation was relatively unaffected by deletions in the scanning tract.

EXAMPLE 52

[0883] This example describes modifications of selected IRES sequences by inserting 5'

and/or 3' untranslated regions (UTRs) and creating IRES hybrids. Briefly, circular RNA constructs were generated with modified IRES elements operably linked to a gaussia luciferase coding sequence. Luminescence from secreted gaussia luciferase in supernatant was measured 24 and 48 hours after transfection of HepG2 cells with RNA.

[0884] IRES sequences with UTRs inserted are provided in SEQ ID NOs: 390-401. As shown in FIG. 53, insertion of 5' UTR right after the 3' end of the IRES and before the start codon slightly increased the translation from Caprine Kobuvirus (CK) IRES but in some instances abrogated translation from Salivirus SZ1 IRES. Insertion of 3' UTR right after the stop cassette had no impact on both IRES sequences.

[0885] Hybrid CK IRES sequences are provided in SEQ ID NOs: 390-401. CK IRES was used as a base, and specific regions of the CK IRES were replaced with similar-looking structures from other IRES sequences, for example, SZ1 and AV (Aichivirus). As shown in FIG. 53, certain hybrid synthetic IRES sequences were functional, indicating that hybrid IRES can be constructed using parts from distinct IRES sequences that show similar predicted structures while deleting these structures completely abrogates IRES function.

EXAMPLE 53

[0886] This example describes modifications of circular RNAs by introducing stop codon or cassette variants. Briefly, circular RNA constructs were generated with IRES elements operably linked to a gaussia luciferase coding sequence followed by variable stop codon cassettes, which included a stop codon in each frame and two stop codons in the reading frame of the gaussia luciferase coding sequence. 1C1C7 cells were transfected with the circular RNAs. Luminescence in supernatant was assessed 24 and 48 hours after transfection.

[0887] The sequences of the stop codon cassettes are set forth in SEQ ID NOs: 406-412. As shown in FIG. 54, certain stop codon cassettes improved expression levels, although they had little impact on expression stability. In particular, a stop cassette with two frame 1 (the reading frame of the gaussia luciferase coding sequence) stop codons, the first being TAA, followed by a frame 2 stop codon and a frame 3 stop codon, is effective for promoting functional translation.

EXAMPLE 54

[0888] This example describes modifications of circular RNAs by inserting 5' UTR

variants. Briefly, circular RNA constructs were generated with IRES elements with 5' UTR variants inserted between the 3' end of the IRES and the start codon, the IRES being operably linked to a gaussia luciferase coding sequence. IC1C7 cells were transfected with the circular RNAs. Luminescence in supernatant was assessed 24 and 48 hours after transfection.

[0889] The sequences of the 5' UTR variants are set forth in SEQ ID NOs: 402-405. As shown in FIG. 55, a CK IRES with a canonical Kozak sequence (UTR4) was more effective when a 36-nucleotide unstructured/low GC spacer sequence was added (UTR2), suggesting that the GC-rich Kozak sequences may interfere with core IRES folding. Using a higher-GC/structured spacer with a kozak sequence did not show the same benefit (UTR3), possibly due to interference with IRES folding by the spacer itself. Mutating the kozak sequence to gTcacG (UTR1) enhanced translation to the same level as the Kozak+spacer alternative without the need for a spacer.

EXAMPLE 55

[0890] This example describes the impact of miRNA target sites in circular RNAs on expression levels. Briefly, circular RNA constructs were generated with IRES elements operably linked to a human erythropoietin (hEPO) coding sequence, where 2 tandem miR-122 target sites were inserted into the construct. miR-122-expressing Huh7 cells were transfected with the circular RNAs. hEPO expression in supernatant was assessed 24 and 48 hours after transfection by sandwich ELISA.

[0891] As shown in FIG. 56, the hEPO expression level was obrogated where the miR-122 target sites were inserted into the circular RNA. This result demonstrates that expression from circular RNA can be regulated by miRNA. As such, cell type- or tissue-specific expression can be achieved by incorporating target sites of the miRNAs expressed in the cell types in which expression of the recombinant protein is undesirable.

EXAMPLE 56

Example 56A: Expression of two proteins from one circRNA containing a 2A site and functional RNA stability.

[0892] Constructs including anabaena intron / exon regions, varying IRES, a first expression sequence encoding Gaussia luciferase, a 2A self-cleaving peptide, and a second

expression sequence encoding GFP or EGFP are circularized. 20,000-40,000 293 cells are transfected with 40ng of circRNA. Luminescence from secreted Gaussia luciferase in supernatant is measured every 24 hours after transfection with complete media replacements at each time point. Fluorescence from expressed GFP or EGFP is measured 24 hours after transfection.

***Example 56B:** Expression of two proteins from one circRNA containing 2 IRES sequences and functional RNA stability.*

[0893] Constructs including anabaena intron / exon regions, a first IRES, a first expression sequence encoding Gaussia luciferase, optionally a spacer, a second IRES, and a second expression sequence encoding GFP or EGFP, with the first and/or second IRES varying by construct are circularized. 293 cells are electroporated with 1 µg of each circularization reaction. Luminescence/fluorescence from secreted Gaussia luciferase and GFP or EGFP in supernatant is measured 24 hours after electroporation. Functional stability of the IRES constructs in each round of electroporated 293 cells is measured over 3 days. Luminescence/fluorescence from secreted Gaussia luciferase and GFP or EGFP in supernatant is measured every 24 hours after electroporation of cells with 1 µg of each circularization reaction, followed by complete media replacement.

EXAMPLE 57

Expression and functional stability of circular and linear RNA containing 2 expression sequences separated by a cleavage site.

[0894] Constructs including anabaena intron / exon regions, a first expression sequence encoding Gaussia luciferase, a 2A self-cleaving peptide, an expression sequence encoding GFP or EGFP is circularized. mRNA encoding a first expression sequence encoding Gaussia luciferase, a 2A self-cleaving peptide, a second expression sequence encoding EGFP, a ~150 nt polyA tail, and modified to replace 100% of uridine with 1-methyl pseudouridine is produced. Expression of circular and modified mRNA is measured in 293 cells. Luminescence/fluorescence from secreted Gaussia luciferase and GFP or EGFP in supernatant is measured 24 hours after electroporation of cells with 1 µg of each RNA species. Luminescence/fluorescence from secreted Gaussia luciferase and GFP or EGFP in supernatant is measured every 24 hours after electroporation over 3 days in order to compare construct functional stability.

EXAMPLE 58

Example 58A: Comparison of expression and functional stability of dual expression with single expression constructs.

[0895] A construct including an anabaena intron / exon region, an IRES, a *Gaussia* luciferase expression sequence, a 2A self-cleaving peptide, and an GFP or EGFP expression sequence is circularized. Constructs containing a single expression sequence (encoding *Gaussia* luciferase or GFP or EGFP) are also circularized. Linear constructs containing a single expression sequence (encoding *Gaussia* luciferase or GFP or EGFP) are also generated. 293 cells are electroporated with 1 µg of each circularization reaction or an equivalent amount of linear RNA. Fluorescence from secreted *Gaussia* luciferase and GFP or EGFP positivity in supernatant is measured 24 hours after electroporation.

[0896] Functional stability of the constructs in each round of electroporated 293 cells is measured over 3 days. Luminescence from secreted *Gaussia* luciferase in supernatant is measured every 24 hours after electroporation of cells with 1 µg of each circularization reaction or equivalent mRNA, followed by complete media replacement.

Example 58B: Comparison of expression and functional stability of dual expression with single expression constructs.

[0897] A construct including an anabaena intron / exon region, a first IRES, a *Gaussia* luciferase expression sequence, a second IRES, and an GFP or EGFP expression sequence is circularized. Constructs containing a single expression sequence (encoding *Gaussia* luciferase or GFP or EGFP) are also circularized. Linear constructs containing a single expression sequence (encoding *Gaussia* luciferase or GFP or EGFP) are also generated. 293 cells are electroporated with 1 µg of each circularization reaction or an equivalent amount of linear RNA. Fluorescence from secreted *Gaussia* luciferase and GFP or EGFP positivity in supernatant is measured 24 hours after electroporation.

[0898] Functional stability of the constructs in each round of electroporated 293 cells is measured over 3 days. Luminescence from secreted *Gaussia* luciferase in supernatant is measured every 24 hours after electroporation of cells with 1 µg of each circularization reaction or equivalent mRNA, followed by complete media replacement.

EXAMPLE 59

Example 59A: Functional interaction between TCR payloads expressed using an IRES and a

cleavage site.

[0899] Constructs including anabaena intron / exon regions, an IRES, a TCR alpha chain expression sequence, a 2A self-cleaving peptide, and a TCR beta chain expression sequence are circularized.

[0900] Human primary CD3+ T cells are electroporated with the circRNA and co-cultured for 24 hours with cells stably expressing the TCR target, GFP and firefly luciferase. Human primary CD3+ T cells are mock electroporated and co-cultured as a control. Quantification of specific lysis of target cells is determined by detection of firefly luminescence. % Specific lysis defined as $1 - [\text{TCR condition luminescence}] / [\text{mock condition luminescence}]$.

Example 59B: Functional interaction between TCR payloads expressed using 2 IRES.

[0901] Constructs including anabaena intron / exon regions, a first IRES, a TCR alpha chain expression sequence, optionally a spacer, a second IRES, and a TCR beta chain expression sequence are circularized.

[0902] Human primary CD3+ T cells are electroporated with the circRNA and co-cultured for 24 hours with cells stably expressing the TCR target, GFP and firefly luciferase. Human primary CD3+ T cells are mock electroporated and co-cultured as a control. Quantification of specific lysis of target cells is determined by detection of firefly luminescence. % Specific lysis defined as $1 - [\text{TCR condition luminescence}] / [\text{mock condition luminescence}]$.

EXAMPLE 60

Example 60A: CAR payloads expressed using an IRES and a cleavage site.

[0903] Constructs including anabaena intron / exon regions, an IRES, an anti-CD19 CAR expression sequence, a 2A self-cleaving peptide, and an anti-CD20 CAR expression sequence are circularized.

[0904] Human primary CD3+ T cells are electroporated with the circRNA and co-cultured for 24 hours with cells stably expressing the both CAR targets, GFP and firefly luciferase. Human primary CD3+ T cells are mock electroporated and co-cultured as a control. Quantification of specific lysis of target cells is determined by detection of firefly luminescence. % Specific lysis defined as $1 - [\text{TCR condition luminescence}] / [\text{mock condition luminescence}]$.

Example 60B: TCR payloads expressed using 2 IRES.

[0905] Constructs including anabaena intron / exon regions, a first IRES, an anti-CD19 CAR expression sequence, optionally a spacer, a second IRES, and an anti-CD20 CAR

expression sequence are circularized.

[0906] Human primary CD3+ T cells are electroporated with the circRNA and co-cultured for 24 hours with cells stably expressing both CAR targets, GFP and firefly luciferase. Human primary CD3+ T cells are mock electroporated and co-cultured as a control. Quantification of specific lysis of target cells is determined by detection of firefly luminescence. % Specific lysis defined as $1 - [\text{TCR condition luminescence}] / [\text{mock condition luminescence}]$.

EXAMPLE 61

LNP and circular RNA construct containing anti-CD19 CAR reduces B cells in the blood and spleen in vivo.

[0907] Circular RNA constructs encoding an anti-CD19 CAR expression were encapsulated within lipid nanoparticles as described above. For comparison, circular RNAs encoding luciferase expression were encapsulated within separate lipid nanoparticle.

[0908] C57BL/6 mice at 6 to 8 weeks old were injected with either LNP solution every other day for a total of 4 LNP injections within each mouse. 24 hours after the last LNP injection, the mice's spleen and blood were harvested, stained, and analyzed via flow cytometry. As shown in FIG. 57A and FIG. 57B, mice containing LNP-circular RNA constructs encoding an anti-CD19 CAR led to a statistically significant reduction in CD19+ B cells in the peripheral blood and spleen compared to mice treated with LNP-circular RNA encoding a luciferase.

EXAMPLE 62

IRES sequences contained within circular RNA encoding CARs improves CAR expressions and cytotoxicity of T-Cells.

[0909] Activated murine T-cells were electroporated with 200ng of circular RNA constructs containing a unique IRES and a murine anti-CD19 1D3ζ CAR expression sequence. The IRES contained in these constructs were derived either in whole or in part from a Caprine Kobuvirus, Apodemus Picornavirus, Parabovirus, or Salivirus. A Caprine Kobuvirus derived IRES was additionally codon optimized. As a control, a circular RNA containing a wild-type zeta mouse CAR with no IRES was used for comparison. The T-cells were stained for the CD-19 CAR 24 hours post electroporation to evaluate for surface expression and then co-cultured with A20 Fluc target cells. The assay was then evaluated for cytotoxic killing of the Fluc+

A20 cells 24 hours after co-culture of the T-cells with the target cells.

[0910] As seen in FIGs. 58A, 58B, 58C, and 59, the unique IRES were able to increase the frequency that the T-cells expressed the CAR protein and level of CAR expression on the surface of the cells. The increase frequency of expression of the CAR protein and level of CAR expression on the surface of cells lead to an improved anti-tumor response.

EXAMPLE 63

Cytosolic and surface proteins expressed from circular RNA construct in primary human T-cells.

[0911] Circular RNA construct contained either a sequence encoding for a fluorescent cytosolic reporter or a surface antigen reporter. Fluorescent reporters included green fluorescent protein, mCitrine, mWasabi, Tsapphire. Surface reporters included CD52 and Thy1.1^{bio}. Primary human T-cells were activated with an anti-CD3/anti-CD28 antibody and electroporated 6 days post activation of the circular RNA containing a reporter sequence. T-cells were harvested and analyzed via flow cytometry 24 hours post electroporation. Surface antigens were stained with commercially available antibodies (e.g., Biolegend, Miltenyi, and BD).

[0912] As seen in FIG. 60A and FIG. 60B, cytosolic and surface proteins can be expressed from circular RNA encoding the proteins in primary human T-cells.

EXAMPLE 64

Circular RNAs containing unique IRES sequences have improved translation expression over linear mRNA.

[0913] Circular RNA constructs contained a unique IRES along with an expression sequence for Firefly luciferase (FLuc).

[0914] Human T-cells from 2 donors were enriched and stimulated with anti-CD3/anti-CD28 antibodies. After several days of proliferation, activated T cells were harvested and electroporated with equal molar of either mRNA or circular RNA expressing FLuc payloads. Various IRES sequences, including those derived from Caprine Kobuvirus, Apodemus Picornavirus, and Parabovirus, were studied to evaluate expression level and durability of the payload expression across 7 days. Across the 7 days, the T-cells were lysed with Promega Brightglo to evaluate for bioluminescences.

[0915] As shown in FIGs. 61C, 61D, 61E, 61F, and 61G, the presence of an IRES within a circular RNA can increase translation and expression of a cytosolic protein by orders of magnitude and can improve expression compared to linear mRNA. This was found consistent across multiple human T-cell donors.

EXAMPLE 65

Example 65A: LNP-circular RNA encoding anti-CD19 mediates human T-cell killing of K562 cells.

[0916] Circular RNA constructs contained a sequence encoding for anti-CD19 antibodies. Circular RNA constructs were then encapsulated within a lipid nanoparticle (LNP).

[0917] Human T-cells were stimulated with anti-CD3/anti-CD28 and left to proliferate up to 6 days. At day 6, LNP-circular RNA and ApoE3 (1µg/mL) were co-cultured with the T-cells to mediate transfection. 24 hours later, Fluc+ K562 cells were electroporated with 200ng of circular RNA encoding anti-CD19 antibodies and were later co-cultured at day 7. 48 hours post co-culture, the assay was assessed for CAR expression and cytotoxic killing of K562 cells through Fluc expression.

[0918] As shown in FIG. 62A and FIG. 62B, there is T-cell expression of anti-CD19 CAR from the LNP-mediated delivery of a CAR *in vitro* to T-cells and its capability to lyse tumor cells in a specific, antigen dependent manner in engineered K562 cells.

Example 65B: LNP-circular RNA encoding anti-BCMA antibody mediates human T-cell killing of K562 cells.

[0919] Circular RNA constructs contained a sequence encoding for anti-BCMA antibodies. Circular RNA constructs were then encapsulated within a lipid nanoparticle (LNP).

[0920] Human T-cells were stimulated with anti-CD3/anti-CD28 and left to proliferate up to 6 days. At day 6, LNP-circular RNA and ApoE3 (1µg/mL) were co-cultured with the T-cells to mediate transfection. 24 hours later, Fluc+ K562 cells were electroporated with 200ng of circular RNA encoding anti-BCMA antibodies and were later co-cultured at day 7. 48 hours post co-culture, the assay was assessed for CAR expression and cytotoxic killing of K562 cells through Fluc expression.

[0921] As shown in Figure 62B, there is T-cell expression of BCMA CAR from the LNP-mediated delivery of a CAR *in vitro* to T-cells and its capability to lyse tumor cells in a specific, antigen dependent manner in engineered K562 cells.

EXAMPLE 66

Anti-CD19 CAR T-cells exhibit anti-tumor activity in vitro.

[0922] Human T-cells were activated with anti-CD3/anti-CD28 and electroporated once with 200ng of anti-CD19 CAR-expressing circular RNA. Electroporated T-cells were co-cultured with FLuc+ Nalm6 target cells and non-target Fluc+K562 cells to evaluate CAR-mediated killing. After 24 hours post co-culture, the T-cells were lysed and examined for remanent FLuc expression by target and non-target cells to evaluate expression and stability of expression across 8 days total.

[0923] As shown in Figures 63A and 63B, T-cells express circular RNA CAR constructs in specific, antigen-dependent manner. Results also shows improved cytotoxicity of circular RNAs encoding CARs compared to linear mRNA encoding CARs and delivery of a functional surface receptor.

EXAMPLE 67

Effective LNP transfection of circular RNA mediated with ApoE3

[0924] Human T-cells were stimulated with anti-CD3/anti-CD28 and left to proliferate up to 6 days. At day 6, lipid nanoparticle (LNP) was and circular RNA expressing green fluorescence protein solution with or without ApoE3 (1µg/mL) were co-cultured with the T-cells. 24 hours later, the T-cells were stained for live/dead T-cells and the live T-cells were analyzed for GFP expression on a flow cytometer.

[0925] As shown by FIGs. 64A, 64B, 61C, 61D, and 64E, efficient LNP transfection can be mediated by ApoE3 on activated T-cells, followed by significant payload expression. These results were exhibited across multiple donors.

INCORPORATION BY REFERENCE

[0926] 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 as being incorporated by reference herein.

WHAT IS CLAIMED IS:

1. A circular RNA polynucleotide comprising, in the following order, a post splicing 3' group I intron fragment, an Internal Ribosome Entry Site (IRES), a first expression sequence, a second expression sequence, and a post splicing 5' group I intron fragment.
2. The circular RNA polynucleotide of claim 1, comprising a polynucleotide sequence encoding a cleavage site between the first expression sequence and the second expression sequence.
3. The circular RNA polynucleotide of claim 2, wherein the cleavage site is a self-cleaving spacer.
4. The circular RNA polynucleotide of claim 3, wherein the self-cleaving spacer is a 2A self-cleaving peptide.
5. The circular RNA polynucleotide of claim 1, comprising a second IRES between the first expression sequence and the second expression sequence.
6. The circular RNA of claim 5, wherein the first IRES consists of or comprises a sequence according to any of SEQ ID NO: 1-72.
7. The circular RNA of claim 5 or 6, wherein the second IRES consists of or comprises a sequence according to any of SEQ ID NO: 1-72.
8. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence encodes a first therapeutic protein, and the second expression sequence encodes a second therapeutic protein.
9. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence or the second expression sequence encodes an antibody.
10. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence or the second expression sequence encodes a chimeric antigen receptor.
11. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence or the second expression sequence encodes a transcription factor.
12. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence or the second expression sequence encodes a cytokine.
13. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence or the second expression sequence encodes an immune inhibitory molecule.
14. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence or the second expression sequence encodes an agonist of a costimulatory molecule.

15. The circular RNA polynucleotide of any preceding claim, wherein the first expression sequence or the second expression sequence encodes an inhibitor of an immune checkpoint molecule.
16. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes the alpha chain of a T cell receptor (TCR) and the second expression sequence encodes the beta chain of a T cell receptor (TCR).
17. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes the beta chain of a T cell receptor (TCR) and the second expression sequence encodes the alpha chain of a T cell receptor (TCR).
18. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes the gamma chain of a T cell receptor (TCR) and the second expression sequence encodes the delta chain of a T cell receptor (TCR).
19. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes the delta chain of a T cell receptor (TCR) and the second expression sequence encodes the gamma chain of a T cell receptor (TCR).
20. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a T cell receptor (TCR) and the second expression sequence encodes a cytokine.
21. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a cytokine and the second expression sequence encodes a T cell receptor (TCR).
22. The circular RNA polynucleotide of any of claims 20-21, wherein the cytokine is selected from IL-2, IL-7, IL-12, and IL-15.
23. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a T cell receptor (TCR) and the second expression sequence encodes a chemokine.
24. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for a chemokine and the second expression sequence encodes for a T cell receptor (TCR).
25. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for a T cell receptor (TCR) and the second expression sequence encodes for a transcription factor.

26. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for a transcription factor and the second expression sequence encodes for a T cell receptor (TCR).
27. The circular RNA polynucleotide of any of claims 25-26, wherein the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt, and cd25.
28. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a chimeric antigen receptor (CAR) and the second expression sequence encodes a PD1 or PDL1 antagonist.
29. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a PD1 or PDL1 antagonist and the second expression sequence encodes a chimeric antigen receptor (CAR).
30. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a chimeric antigen receptor (CAR) and the second expression sequence encodes a cytokine.
31. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a cytokine and the second expression sequence encodes a chimeric antigen receptor (CAR).
32. The circular RNA polynucleotide of any of claims 30-31, wherein the cytokine is selected from IL-2, IL-7, IL-12, and IL-15.
33. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for a chimeric antigen receptor (CAR), and the second expression sequence encodes a chemokine.
34. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for a chemokine, and the second expression sequence encodes for a chimeric antigen receptor (CAR).
35. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a transcription factor and the second expression sequence encodes a cytokine.
36. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a cytokine and the second expression sequence encodes a transcription factor.
37. The circular RNA polynucleotide of any of claims 35-36, wherein the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt, and cd25.

38. The circular RNA polynucleotide of any of claims 35-37, wherein the cytokine is selected from IL-10, IL-12, and TGF β .
39. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a transcription factor and the second expression sequence encodes a chemokine.
40. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a chemokine and the second expression sequence encodes a transcription factor.
41. The circular RNA of any of claims 39-40, wherein the transcription factor is selected from FOXP3, STAT5B, and HELIOS.
42. The circular RNA polynucleotide of any of claims 39-41, wherein the chemokine is a CC chemokine, CXC chemokine, C chemokine, or a CX3C chemokine.
43. The circular RNA polynucleotide of any of claims 39-42, wherein the chemokine is selected from CCL1, CCL2, CCL3, CCL4, CCL5, CCL6, CCL7, CCL8, CCL9/CCL10, CCL11, CCL12, CCL13, CCL14, CCL15, CCL16, CCL17, CCL18, CCL19, CCL20, CCL21, CCL22, CCL23, CCL24, CCL25, CCL26, CCL27, CCL28, CXCL 1, CXCL2, CXCL3, CXCL4, CXCL5, CXCL6, CXCL7, CXCL8, CXCL9, CXCL10, CXCL11, CXCL12, CXCL13, CXCL14, CXCL15, CXCL16, CXCL17, XCL1, XCL2, and CX3CL1.
44. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a tumor antigen and the second expression sequence encodes a cytokine.
45. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a cytokine and the second expression sequence encodes a tumor antigen.
46. The circular RNA polynucleotide of any of claims 44-45, wherein the antigen is a neoantigen.
47. The circular RNA polynucleotide of any of claims 44-46, wherein the cytokine is IFN γ .
48. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a CAR and the second expression sequence encodes a CAR.
49. The circular RNA polynucleotide of any of claims 1-8, wherein the first expression sequence encodes a cytokine and the second expression sequence encodes a cytokine.
50. The circular RNA polynucleotide of claim 49, wherein the first or second expression sequence encodes a cytokine selected from IL-10, TGF β , and IL-35.

51. The circular RNA polynucleotide of any of claims 49-50, wherein the first or second expression sequence encodes a cytokine selected from IFN γ , IL-2, IL-7, IL-15, and IL-18.
52. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a T cell receptor (TCR) and the second expression sequence encodes a T cell receptor (TCR).
53. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for a chemokine and the second expression sequence encodes for a chemokine.
54. The circular RNA polynucleotide of any one of claims 1-8, wherein the first or second expression sequence encodes for an immunosuppressive enzyme.
55. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a rate limiting enzyme and the second expression sequence encodes a flux-limiting enzyme.
56. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a flux-limiting enzyme and the second expression sequence encodes a rate limiting enzyme.
57. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a transcription factor and the second expression sequence encodes a survival factor.
58. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a survival factor and the second expression sequence encodes a transcription factor.
59. The circular RNA polynucleotide of any one of claims 57-58, wherein the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, ROR γ t, and cd25.
60. The circular RNA polynucleotide of any one of claims 57-59, wherein the survival factor is selected from BCL-XL.
61. The circular RNA polynucleotide of any one of claims 1-8, wherein the first or second expression sequence encodes for a chaperone protein or complex.
62. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a transcription factor and the second expression sequence encodes a chaperone protein or complex.

63. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a chaperone protein or complex and the second expression sequence encodes for a transcription factor.
64. The circular RNA polynucleotide of any one of claims 61-63, wherein the chaperone protein or complex is selected from Skp, Spy, FkpA, SurA, Hsp60, Hsp70, GroEL, GroES, Hsp90, HtpG, Hsp100, ClpA, ClpX, ClpP, and Hsp104.
65. The circular RNA polynucleotide of any one of claims 61-64, wherein the transcription factor is selected from FOXP3, STAT5B, HELIOS, Tbet, GATA3, RORgt, and cd25.
66. The circular RNA polynucleotide of any one of claims 1-8, wherein one or both expression sequences encode a signaling protein.
67. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for an enzyme and the second expression encodes for the first expression sequence's negative regulatory inhibitor.
68. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes for a negative regulatory inhibitor protein of an enzyme encoded in the second expression sequence.
69. The circular RNA polynucleotide of any one of claims 67-68, wherein the negative regulatory inhibitor is selected from a p57kip2, BAX inhibitor, and TIPE2.
70. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a dominant negative protein and the second expression sequence encodes an immune protein.
71. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes an immune protein and the second expression sequence encodes a dominant negative protein.
72. The circular RNA polynucleotide of any one of claims 1-8, wherein the first or second expression sequence encodes for an anti-inflammatory protein.
73. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence encodes a transcription factor and the second expression sequence is capable of converting 5-fluorocytosine (5-FC) into 5-fluorouracil (5-FU).
74. The circular RNA polynucleotide of any one of claims 1-8, wherein the first expression sequence is capable of converting 5-fluorocytosine (5-FC) into 5-fluorouracil (5-FU) and the second expression sequence is a transcription factor.

75. The circular RNA polynucleotide of any one of claims 73-74, wherein the expression sequence capable of converting 5-fluorocytosine (5-FC) into 5-fluorouracil (5-FU) is cytosine deaminase.
76. The circular RNA polynucleotide of any preceding claim, comprising a first spacer between the 5' duplex forming region and the post splicing 3' group I intron fragment, and a second spacer between the post splicing 5' group I intron fragment and the 3' duplex forming region.
77. The circular RNA polynucleotide of claim 76, wherein the first and second spacers each have a length of about 10 to about 60 nucleotides.
78. The circular RNA polynucleotide of any one of the preceding claims, wherein the first and second duplex forming regions each have a length of about 9 to about 19 nucleotides.
79. The circular RNA polynucleotide of any one of claims 77-78, wherein the first and second duplex forming regions each have a length of about 30 nucleotides.
80. The circular RNA polynucleotide of any one of the preceding claims, wherein the IRES has a sequence of an IRES from Taura syndrome virus, Triatoma virus, Theiler's encephalomyelitis virus, Simian Virus 40, Solenopsis invicta virus 1, Rhopalosiphum padi virus, Reticuloendotheliosis virus, Human poliovirus 1, Plautia stali intestine virus, Kashmir bee virus, Human rhinovirus 2, Homalodisca coagulata virus- 1, Human Immunodeficiency Virus type 1, Homalodisca coagulata virus- 1, Himetobi P virus, Hepatitis C virus, Hepatitis A virus, Hepatitis GB virus, Foot and mouth disease virus, Human enterovirus 71, Equine rhinitis virus, Ectropis obliqua picorna-like virus, Encephalomyocarditis virus, Drosophila C Virus, Human coxsackievirus B3, Crucifer tobamovirus, Cricket paralysis virus, Bovine viral diarrhea virus 1, Black Queen Cell Virus, Aphid lethal paralysis virus, Avian encephalomyelitis virus, Acute bee paralysis virus, Hibiscus chlorotic ringspot virus, Classical swine fever virus, Human FGF2, Human SFTPA1, Human AML1/RUNX1, Drosophila antennapedia, Human AQP4, Human AT1R, Human BAG-1, Human BCL2, Human BiP, Human c-IAP1, Human c-myc, Human eIF4G, Mouse NDST4L, Human LEF1, Mouse HIF1 alpha, Human n.myc, Mouse Gtx, Human p27kipl, Human PDGF2/c-sis, Human p53, Human Pim-1, Mouse Rbm3, Drosophila reaper, Canine Scamper, Drosophila Ubx, Human UNR, Mouse UtrA, Human VEGF-A, Human XIAP, Drosophila hairless, S. cerevisiae TFIID, S. cerevisiae YAP1, tobacco etch virus, turnip crinkle virus, EMCV-A, EMCV-B, EMCV-Bf, EMCV-Cf, EMCV pEC9, Picobirnavirus, HCV QC64, Human Cosavirus

- E/D, Human Cosavirus F, Human Cosavirus JMY, Rhinovirus NAT001, HRV14, HRV89, HRVC-02, HRV-A21, Salivirus A SH1, Salivirus FHB, Salivirus NG-J1, Human Parechovirus 1, Crohivirus B, Yc-3, Rosavirus M-7, Shanbavirus A, Pasivirus A, Pasivirus A 2, Echovirus E14, Human Parechovirus 5, Aichi Virus, Hepatitis A Virus HA16, Phopivirus, CVA10, Enterovirus C, Enterovirus D, Enterovirus J, Human Pegivirus 2, GBV-C GT110, GBV-C K1737, GBV-C Iowa, Pegivirus A 1220, Pasivirus A 3, Sapelovirus, Rosavirus B, Bakunsa Virus, Tremovirus A, Swine Pasivirus 1, PLV-CHN, Pasivirus A, Sicinivirus, Hepacivirus K, Hepacivirus A, BVDV1, Border Disease Virus, BVDV2, CSFV-PK15C, SF573 Dicistrovirus, Hubei Picorna-like Virus, CRPV, Salivirus A BN5, Salivirus A BN2, Salivirus A 02394, Salivirus A GUT, Salivirus A CH, Salivirus A SZ1, Salivirus FHB, CVB3, CVB1, Echovirus 7, CVB5, EVA71, CVA3, CVA12, EV24 or an aptamer to eIF4G.
81. The circular RNA polynucleotide of any one of the preceding claims, consisting of natural nucleotides.
 82. The circular RNA polynucleotide of any one of the preceding claims, wherein the expression sequence is codon optimized.
 83. The circular RNA polynucleotide of any one of the preceding claims, optimized to lack at least one microRNA binding site present in an equivalent pre-optimized polynucleotide.
 84. The circular RNA polynucleotide of any one of the preceding claims, optimized to lack at least one endonuclease susceptible site present in an equivalent pre-optimized polynucleotide.
 85. The circular RNA polynucleotide of any one of the preceding claims, optimized to lack at least one RNA-editing susceptible site present in an equivalent pre-optimized polynucleotide.
 86. The circular RNA polynucleotide of any one of the preceding claims, wherein the circular RNA polynucleotide is from about 100 nucleotides to about 15 kilobases in length.
 87. The circular RNA polynucleotide of any one of the preceding claims, having an *in vivo* duration of therapeutic effect in humans of at least about 20 hours.
 88. The circular RNA polynucleotide of any one of the preceding claims, having a functional half-life of at least about 20 hours.

89. The circular RNA polynucleotide of any one of the preceding claims, having a duration of therapeutic effect in a human cell greater than or equal to that of an equivalent linear RNA polynucleotide comprising the same expression sequence.
90. The circular RNA polynucleotide of any one of the preceding claims, having a functional half-life in a human cell greater than or equal to that of an equivalent linear RNA polynucleotide comprising the same expression sequence.
91. The circular RNA polynucleotide of any one of the preceding claims, having an *in vivo* duration of therapeutic effect in humans greater than that of an equivalent linear RNA polynucleotide having the same expression sequence.
92. The circular RNA polynucleotide of any one of the preceding claims, having an *in vivo* functional half-life in humans greater than that of an equivalent linear RNA polynucleotide having the same expression sequence.
93. A pharmaceutical composition comprising a circular RNA polynucleotide of any one of the preceding claims, a nanoparticle, and optionally, a targeting moiety operably connected to the nanoparticle.
94. The pharmaceutical composition of claim 93, wherein the nanoparticle is a lipid nanoparticle, a core-shell nanoparticle, a biodegradable nanoparticle, a biodegradable lipid nanoparticle, a polymer nanoparticle, or a biodegradable polymer nanoparticle.
95. The pharmaceutical composition of claim 93 or 94, comprising a targeting moiety, wherein the targeting moiety mediates receptor-mediated endocytosis or direct fusion into selected cells of a selected cell population or tissue in the absence of cell isolation or purification.
96. The pharmaceutical composition of any one of claims 93-95, comprising a targeting moiety operably connected to the nanoparticle.
97. The pharmaceutical composition of any one of claims 93-96, wherein the targeting moiety is a scFv, nanobody, peptide, minibody, polynucleotide aptamer, heavy chain variable region, light chain variable region or fragment thereof.
98. The pharmaceutical composition of any one of claims 93-97, wherein less than 1%, by weight, of the polynucleotides in the composition are double stranded RNA, DNA splints, or triphosphorylated RNA.
99. The pharmaceutical composition of any one of claims 93-98, wherein less than 1%, by weight, of the polynucleotides and proteins in the pharmaceutical composition are double stranded RNA, DNA splints, triphosphorylated RNA, phosphatase proteins, protein ligases, and capping enzymes.

100. A method of treating a subject in need thereof comprising administering a therapeutically effective amount of a composition comprising the circular RNA polynucleotide of any one of claims 1-92, a nanoparticle, and optionally, a targeting moiety operably connected to the nanoparticle.
101. The method of claim 100, wherein the targeting moiety is an scFv, nanobody, peptide, minibody, heavy chain variable region, light chain variable region or fragment thereof.
102. The method of any one of claims 100-101, wherein the nanoparticle is a lipid nanoparticle, a core-shell nanoparticle, or a biodegradable nanoparticle.
103. The method of any one of claims 100-102, wherein the nanoparticle comprises one or more cationic lipids, ionizable lipids, or poly β -amino esters.
104. The method of any one of claims 100-103, wherein the nanoparticle comprises one or more non-cationic lipids.
105. The method of any one of claims 100-104, wherein the nanoparticle comprises one or more PEG-modified lipids, polyglutamic acid lipids, or hyaluronic acid lipids.
106. The method of any one of claims 100-105, wherein the nanoparticle comprises cholesterol.
107. The method of any one of claims 100-106, wherein the nanoparticle comprises arachidonic acid or oleic acid.
108. The method of any one of claims 100-107, wherein the composition comprises a targeting moiety, wherein the targeting moiety mediates receptor-mediated endocytosis selectively into cells of a selected cell population in the absence of cell selection or purification.
109. The method of any one of claims 100-108, wherein the nanoparticle comprises more than one circular RNA polynucleotide.
110. The method of any one of claims 100-109, wherein the subject has a cancer selected from the group consisting of acute lymphocytic leukemia; acute myeloid leukemia (AML); alveolar rhabdomyosarcoma; B cell malignancies; bladder cancer (*e.g.*, bladder carcinoma); bone cancer; brain cancer (*e.g.*, medulloblastoma); breast cancer; cancer of the anus, anal canal, or anorectum; cancer of the eye; cancer of the intrahepatic bile duct; cancer of the joints; cancer of the neck; gallbladder cancer; cancer of the pleura; cancer of the nose, nasal cavity, or middle ear; cancer of the oral cavity; cancer of the vulva; chronic lymphocytic leukemia; chronic myeloid cancer; colon cancer; esophageal cancer, cervical cancer; fibrosarcoma; gastrointestinal carcinoid tumor; head and neck cancer (*e.g.*, head and neck squamous cell carcinoma); Hodgkin lymphoma;

- hypopharynx cancer; kidney cancer; larynx cancer; leukemia; liquid tumors; liver cancer; lung cancer (*e.g.*, non-small cell lung carcinoma and lung adenocarcinoma); lymphoma; mesothelioma; mastocytoma; melanoma; multiple myeloma; nasopharynx cancer; non-Hodgkin lymphoma; B-chronic lymphocytic leukemia; hairy cell leukemia; acute lymphocytic leukemia (ALL); Burkitt's lymphoma; ovarian cancer; pancreatic cancer; cancer of the peritoneum; cancer of the omentum; mesentery cancer; pharynx cancer; prostate cancer; rectal cancer; renal cancer; skin cancer; small intestine cancer; soft tissue cancer; solid tumors; synovial sarcoma; gastric cancer; testicular cancer; thyroid cancer; and ureter cancer.
111. The method of any one of claims 100-109, wherein the subject has an autoimmune disorder selected from scleroderma, Grave's disease, Crohn's disease, Sjorgen's disease, multiple sclerosis, Hashimoto's disease, psoriasis, myasthenia gravis, autoimmune polyendocrinopathy syndromes, Type I diabetes mellitus (T1DM), autoimmune gastritis, autoimmune uveoretinitis, polymyositis, colitis, thyroiditis, and the generalized autoimmune diseases typified by human Lupus.
112. A vector for making a circular RNA polynucleotide, comprising, in the following order, a 5' duplex forming region, a 3' Group I intron fragment, an Internal Ribosome Entry Site (IRES), a first expression sequence, a second expression sequence, a 5' Group I intron fragment, and a 3' duplex forming region.
113. The vector of claim 112, comprising a polynucleotide sequence encoding a cleavage site between the first expression sequence and the second expression sequence.
114. The vector of claim 112 or 113, wherein the cleavage site is a self-cleaving spacer.
115. The vector of any one of claims 112-114, wherein the self-cleaving spacer is a 2A self-cleaving peptide.
116. The vector of any one of claims 112-115, comprising a first spacer between the 5' duplex forming region and the 3' group I intron fragment, and a second spacer between the 5' group I intron fragment and the 3' duplex forming region.
117. The vector of any one of claims 112-116, wherein the first and second spacers each have a length of about 5 to about 60 nucleotides.
118. The vector of any one of claims 112-117, wherein the first and second spacers each comprise an unstructured region at least 5 nucleotides long.
119. The vector of any one of claims 112-118, wherein the first and second spacers each comprise a structured region at least 7 nucleotides long.

120. The vector of any one of claims 112-119, wherein the first and second duplex forming regions each have a length of about 9 to 50 nucleotides.
121. The vector of any one of claims 112-120 wherein the vector is codon optimized.
122. The vector of any one of claims 112-121, lacking at least one microRNA binding site present in an equivalent pre-optimization polynucleotide.
123. A eukaryotic cell comprising a circular RNA polynucleotide according to any of claims 1-92.
124. The eukaryotic cell of claim 123, wherein the eukaryotic cell is a human cell.
125. The eukaryotic cell of claim 124, wherein the eukaryotic cell is an immune cell.
126. The eukaryotic cell of claim 125, wherein the eukaryotic cell is a T cell.